

Model Report: RandomForest

18/12/2024

Introduction

This is an automated report for the Test on Tetouan City consumption data; the RandomForest model.

This report will first introduce the model setup, including the hyperparameters and search algorithms used. Hereafter the base dataset will be described, and the differently created training datasets will be listed. After that, the results for the different forecast types will be presented, and the best results will be shown in plots.

Experiment description:

First experiment on the dataset with Tetouan City dataset.

Model setup

The model has been used for the following forecast purposes:

- ☐ one_step
- ☐ multistep
- ☐ recursive

The model has been optimized using the following hyperparameters:

- max_depth: [2, 4, 6, 8, 10, 15, 20]
- criterion: ['squared_error']
- random_state: [42]
- min_samples_split: [5, 10, 50, 150, 200, 250]
- min_samples_leaf: [5, 10, 25, 50, 100]
- scaler: [None, StandardScaler(), MinMaxScaler(), RobustScaler(), PowerTransformer()]

And with the following search algorithms:

- ☐ grid
- ☐ random

The used performance measure is the r2 measure.

Dataset setup

The baseline dataset used for these forecasts is

the 'Power consumption of Tetouan City between 2017-01-01 and 2017-12-31 in 10 minutes resolution.' dataset: *'Power consumption of Tetouan City between 2017-01-01 and 2017-12-31 in 10 minutes resolution. Has the weather data for temperature, humidity, wind speed and diffuse flows, as well as the consumption data for 3 different zones in the city. For simplicity, we will only use one of the zones.'*

The test size used for the forecasts is 0.2.

○ Dataset 1

- name: univariate_temporal
- dataset_type: univariate
- prediction_type: one_step
- components: ['one_step_target', 'temporal_features']

○ Dataset 2

- name: univariate_lagged
- dataset_type: univariate
- prediction_type: one_step
- components: ['one_step_target', 'lagged_target']

○ Dataset 3

- name: univariate_temporal_and_lagged
- dataset_type: univariate
- prediction_type: one_step
- components: ['one_step_target', 'temporal_features', 'lagged_target']

○ Dataset 4

- name: univariate_multistep_temporal
- dataset_type: univariate
- prediction_type: multistep
- components: ['multistep_target', 'temporal_features']

Dataset 5

- name: univariate_multistep_lagged
- dataset_type: univariate
- prediction_type: multistep
- components: ['multistep_target', 'lagged_target']

○ Dataset 6

- name: univariate_multistep_temporal_and_lagged
- dataset_type: univariate
- prediction_type: multistep
- components: ['multistep_target', 'temporal_features', 'lagged_target']

○ Dataset 7

- name: multivariate_temporal
- dataset_type: multivariate
- prediction_type: one_step
- components: ['one_step_target', 'feature_columns', 'temporal_features']

○ Dataset 8

- name: multivariate_lagged
- dataset_type: multivariate
- prediction_type: one_step
- components: ['one_step_target', 'feature_columns', 'lagged_target']

○ Dataset 9

- name: multivariate_temporal_and_lagged
- dataset_type: multivariate
- prediction_type: one_step
- components: ['one_step_target', 'feature_columns', 'temporal_features', 'lagged_target']

○

Dataset 10

- name: multivariate_multistep_temporal
- dataset_type: multivariate
- prediction_type: multistep
- components: ['multistep_target', 'feature_columns', 'temporal_features']

○ Dataset 11

- name: multivariate_multistep_lagged
- dataset_type: multivariate
- prediction_type: multistep
- components: ['multistep_target', 'feature_columns', 'lagged_target']

○ Dataset 12

- name: multivariate_multistep_temporal_and_lagged
- dataset_type: multivariate
- prediction_type: multistep
- components: ['multistep_target', 'feature_columns', 'temporal_features', 'lagged_target']

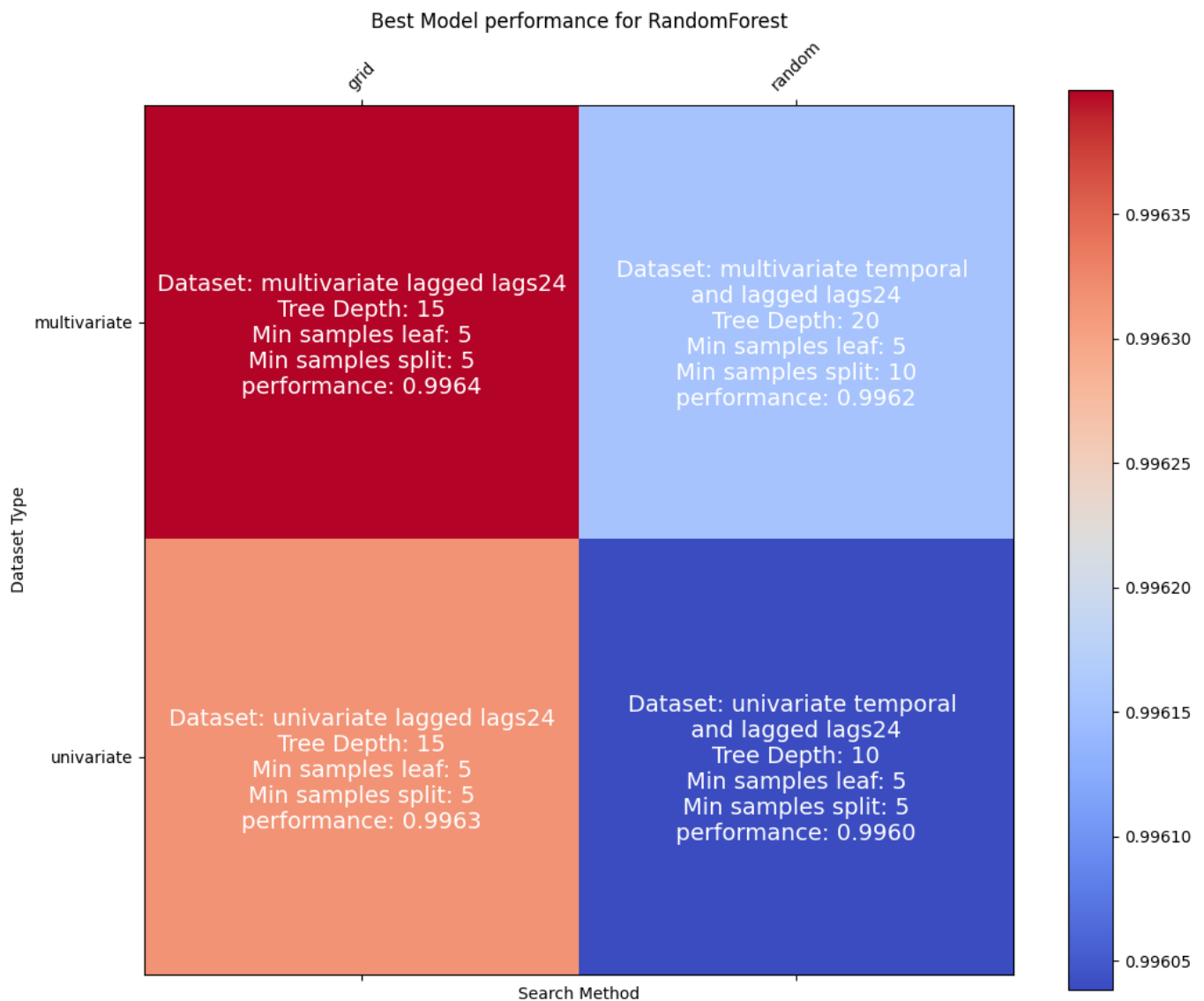
Results: RandomForest

The presentation of the results follows this system: For each prediction type, the best and worst results for each combination of search method and dataset type are presented in heat plots along with the corresponding model setup.

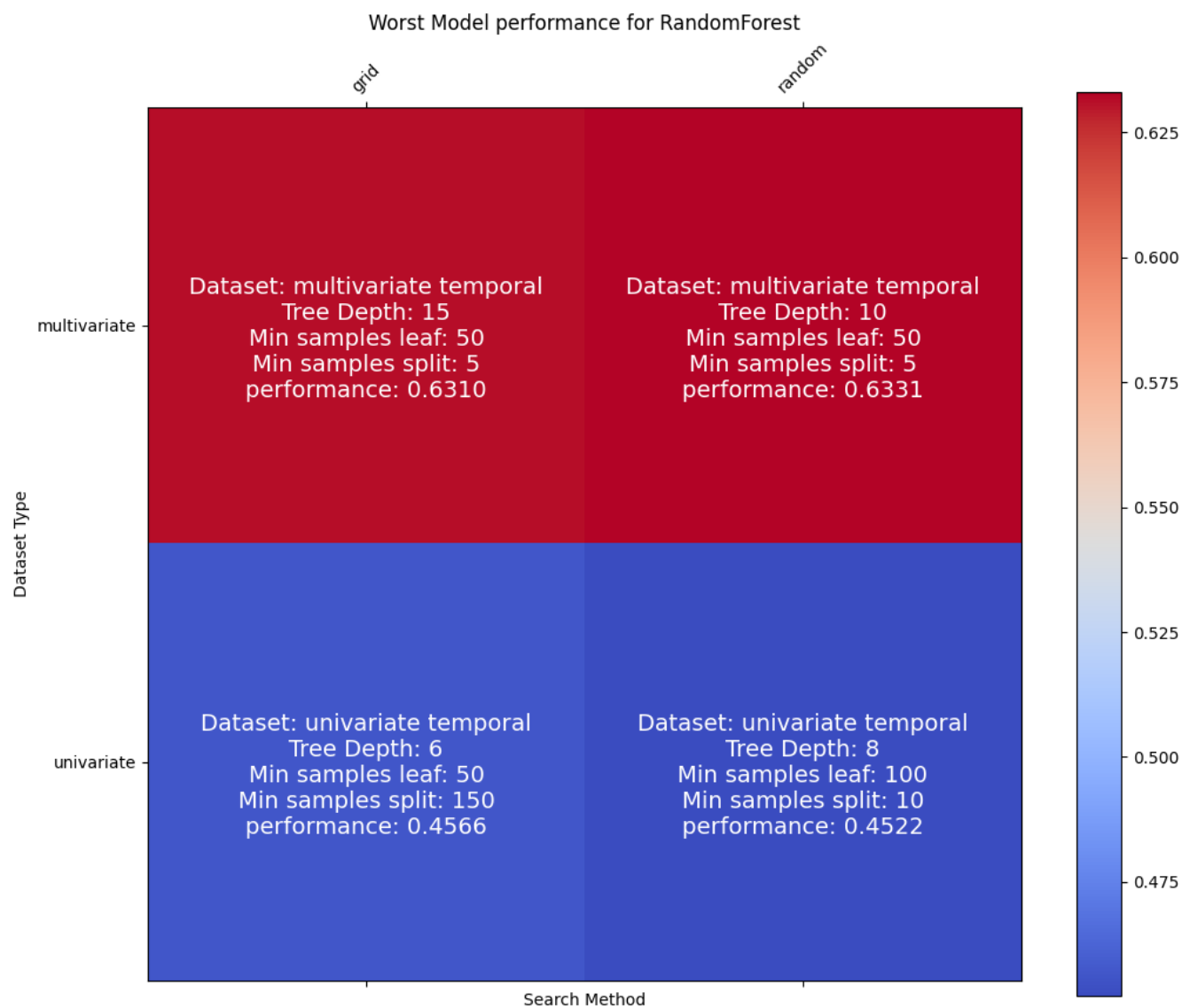
- Then, if the prediction type is one-step forecasts, the best prediction over time is visualized in a line plot.
- If the prediction type is a multi-step forecast, either direct or recursive, the model with the average best r^2 score is chosen, and the three best and worst predictions are visualized in a line plot. Furthermore, three steps of the forecasts are plotted.

Results for the one_step forecast.

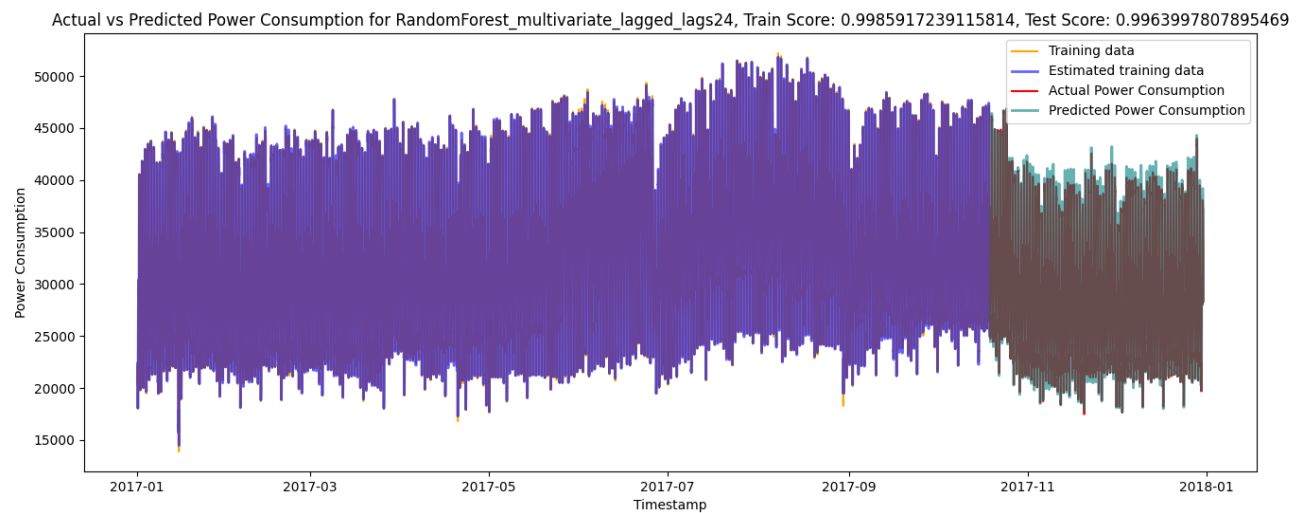
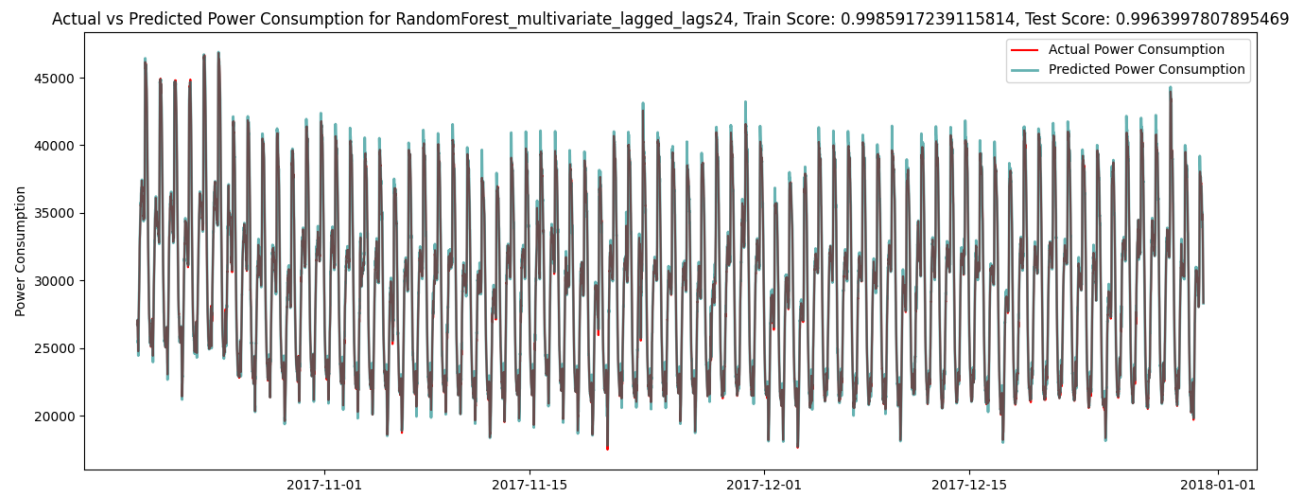
The best results (one_step) for the different setup combinations are as follows:



The worst results (one_step) for the different setup combinations are as follows:

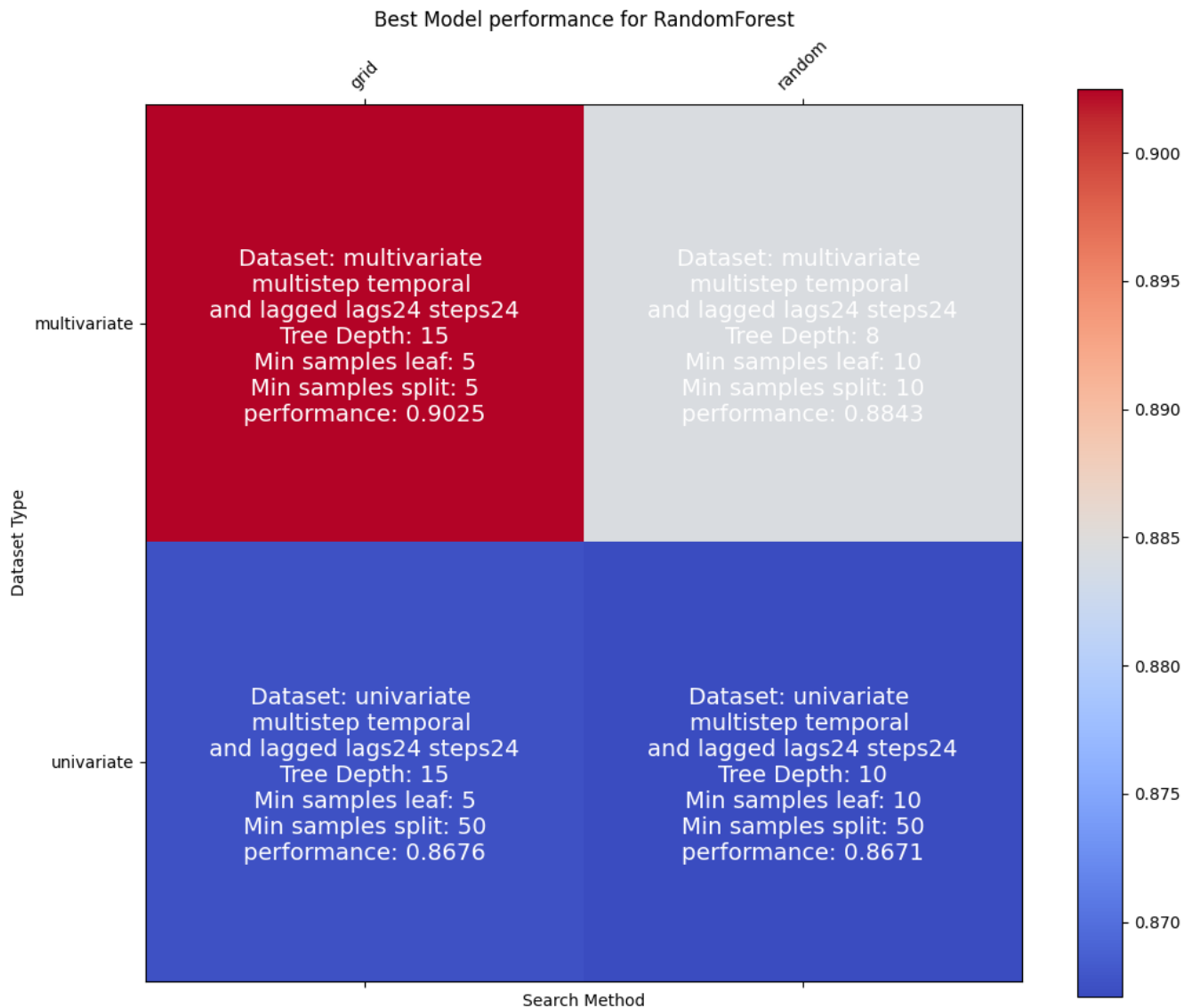


Best one_step forecast over time

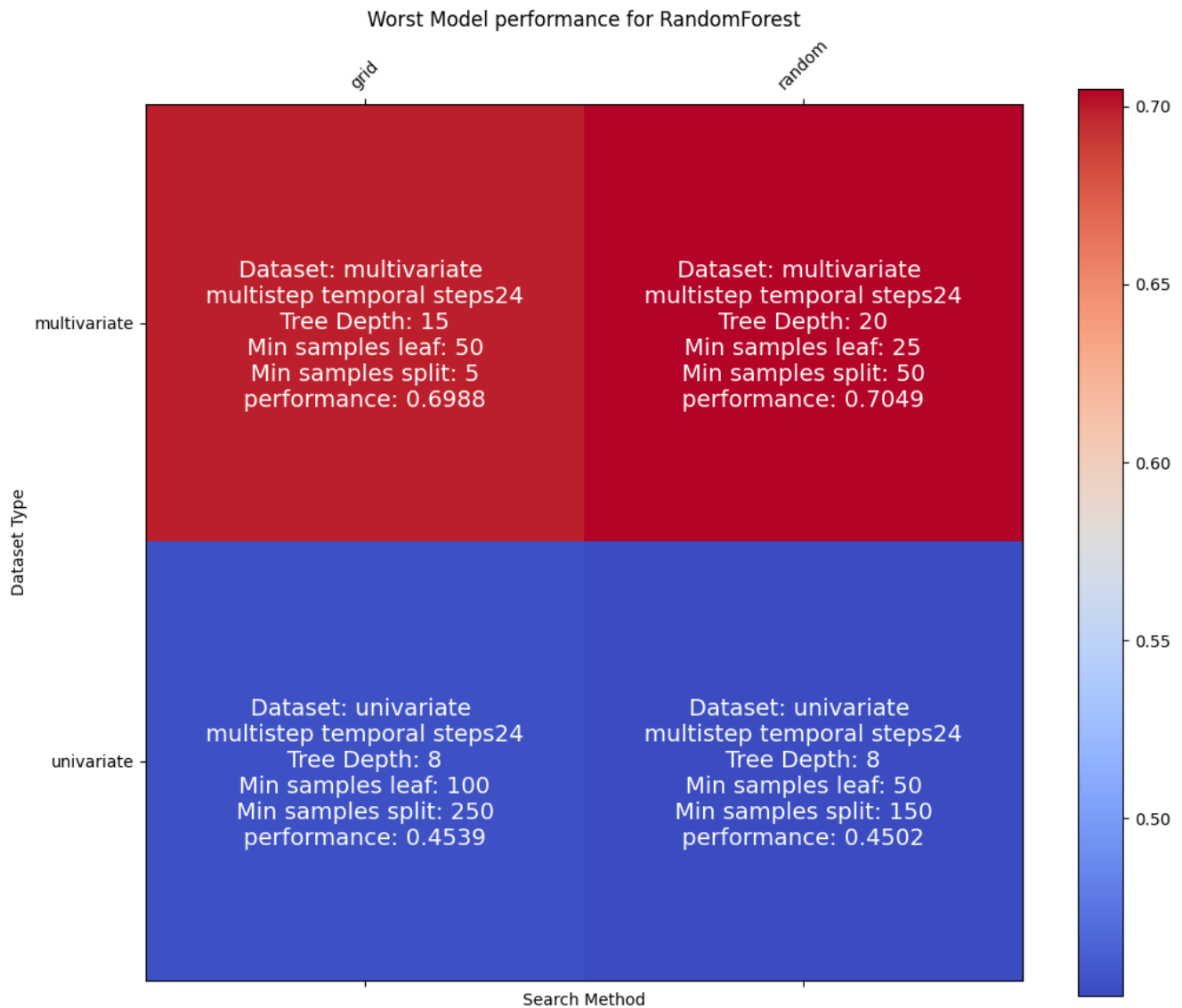


Results for the multistep forecast.

The best results (multistep) for the different setup combinations are as follows:

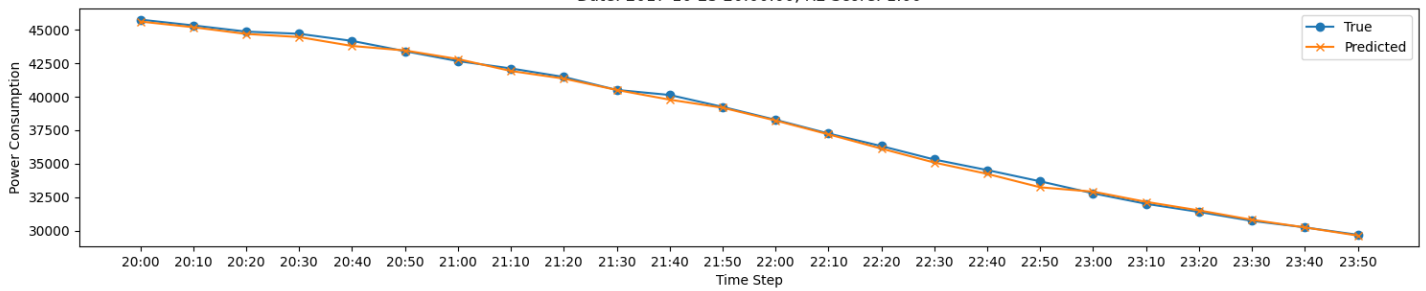


The worst results (multistep) for the different setup combinations are as follows:

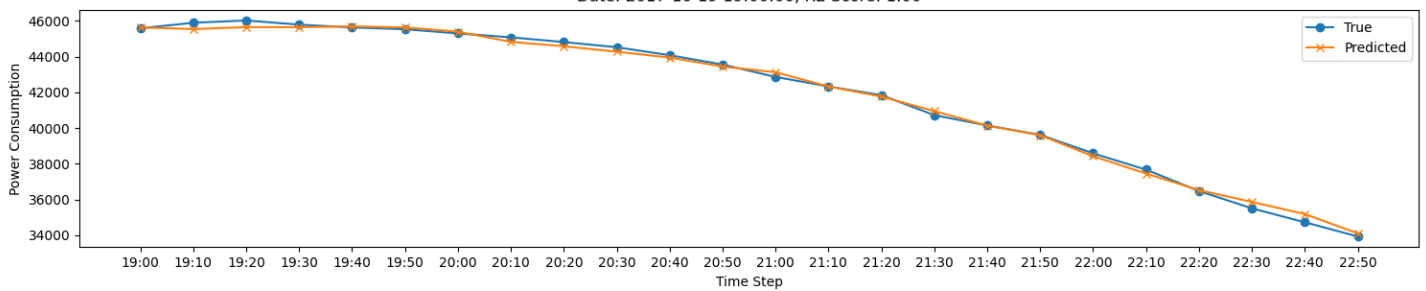


Best predicted days for RandomForest.

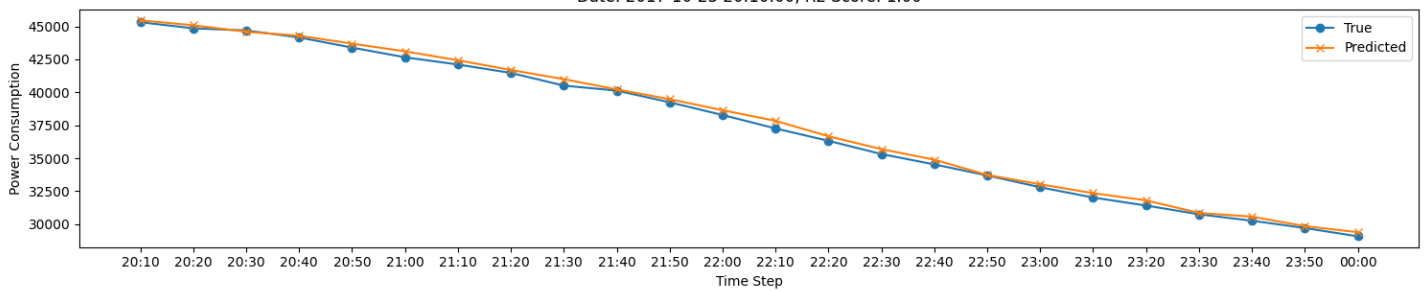
Date: 2017-10-23 20:00:00, R2 Score: 1.00



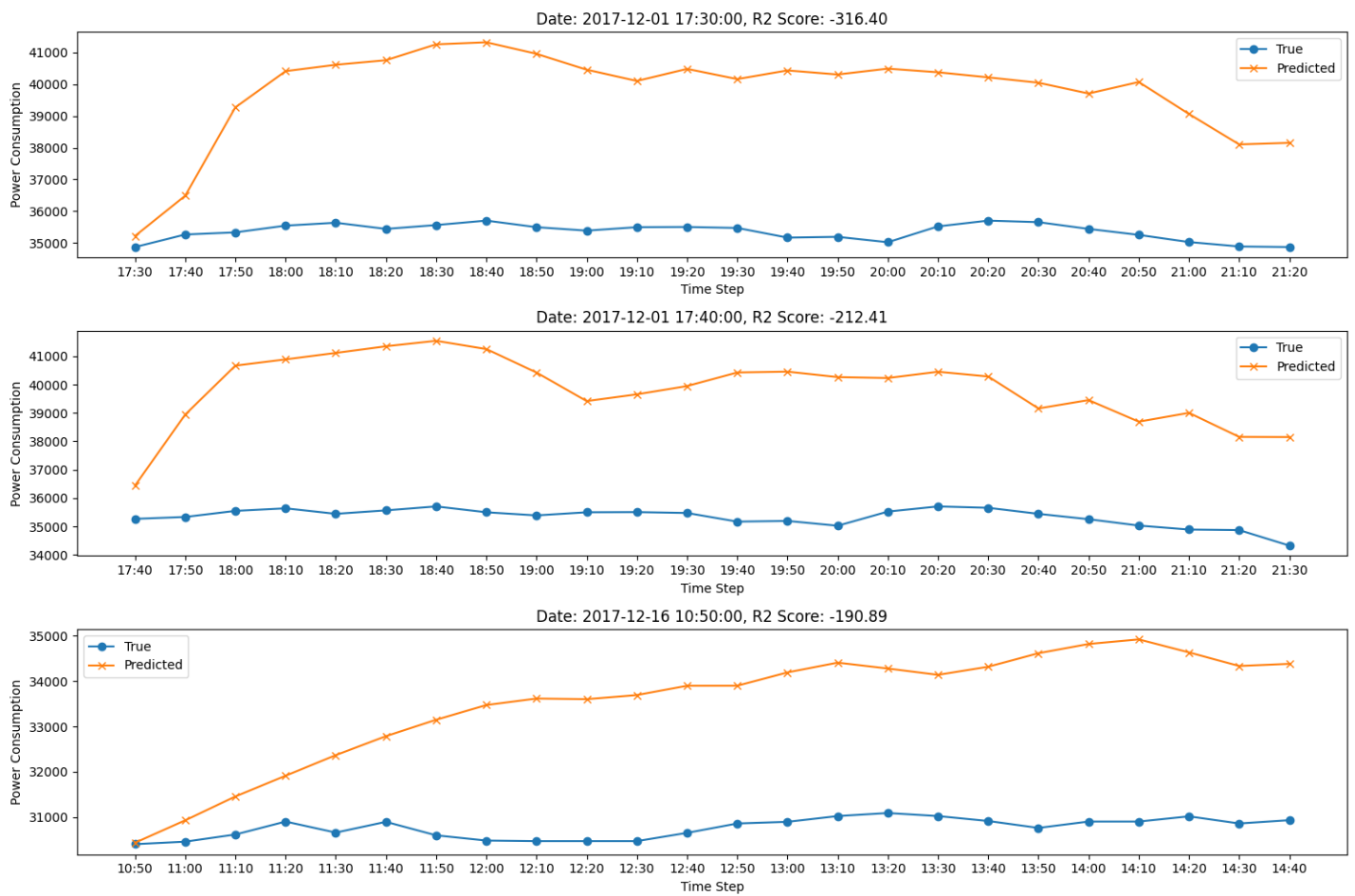
Date: 2017-10-19 19:00:00, R2 Score: 1.00



Date: 2017-10-23 20:10:00, R2 Score: 1.00

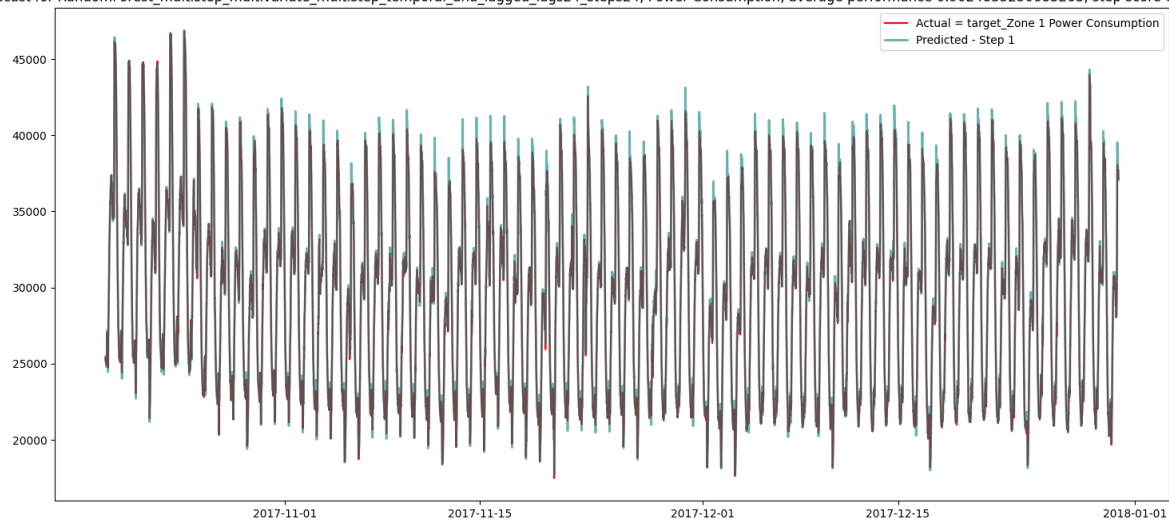


Worst predicted days for RandomForest.

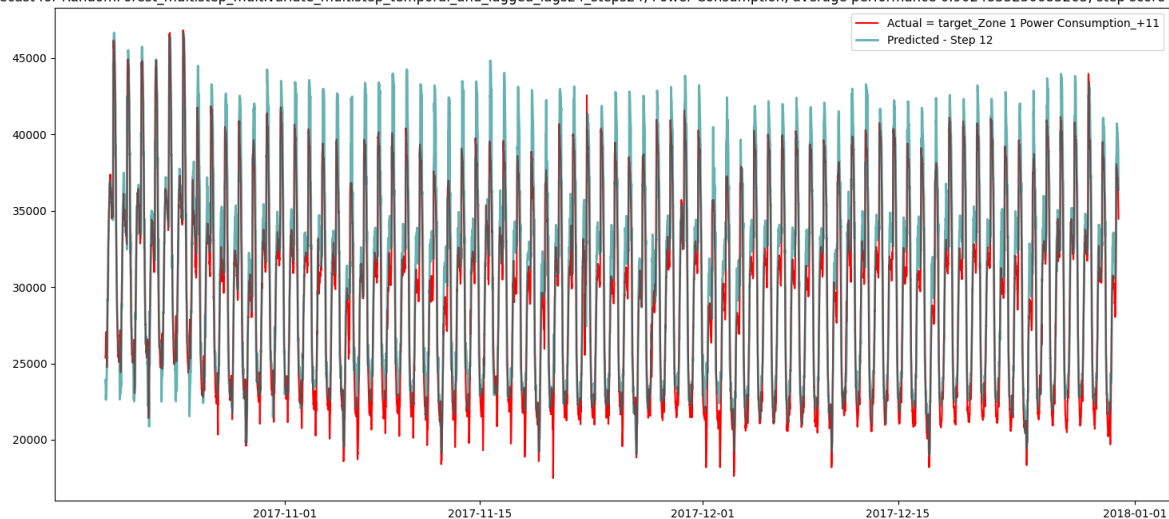


Steps plots for RandomForest forecasts over time

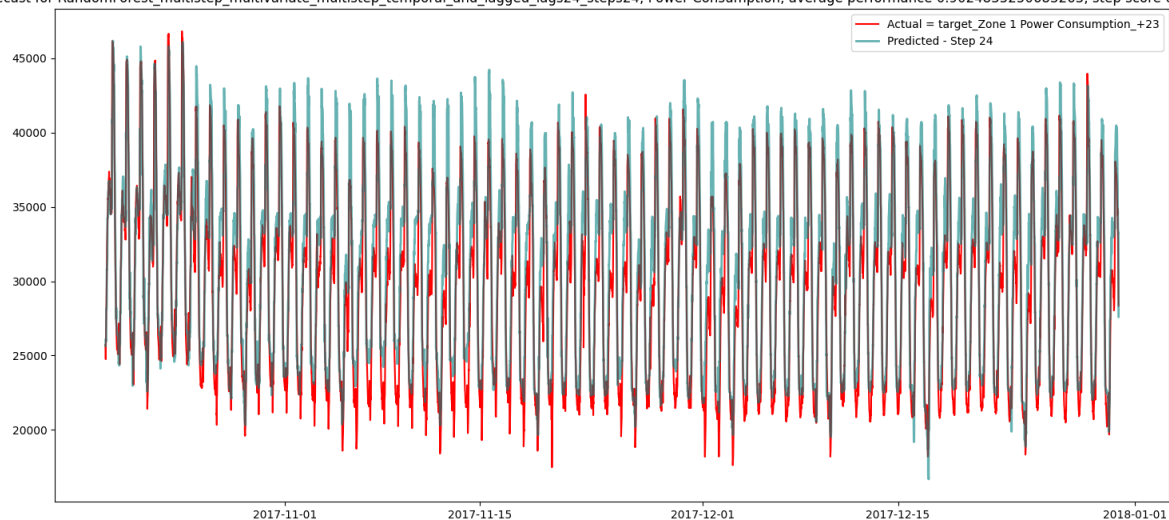
Multistep forecast for RandomForest_multistep_multivariate_multistep_temporal_and_lagged_lags24_steps24, Power Consumption, average performance 0.9024835250683263, step score 0.9961525279321547



Multistep forecast for RandomForest_multistep_multivariate_multistep_temporal_and_lagged_lags24_steps24, Power Consumption, average performance 0.9024835250683263, step score 0.870209628631891

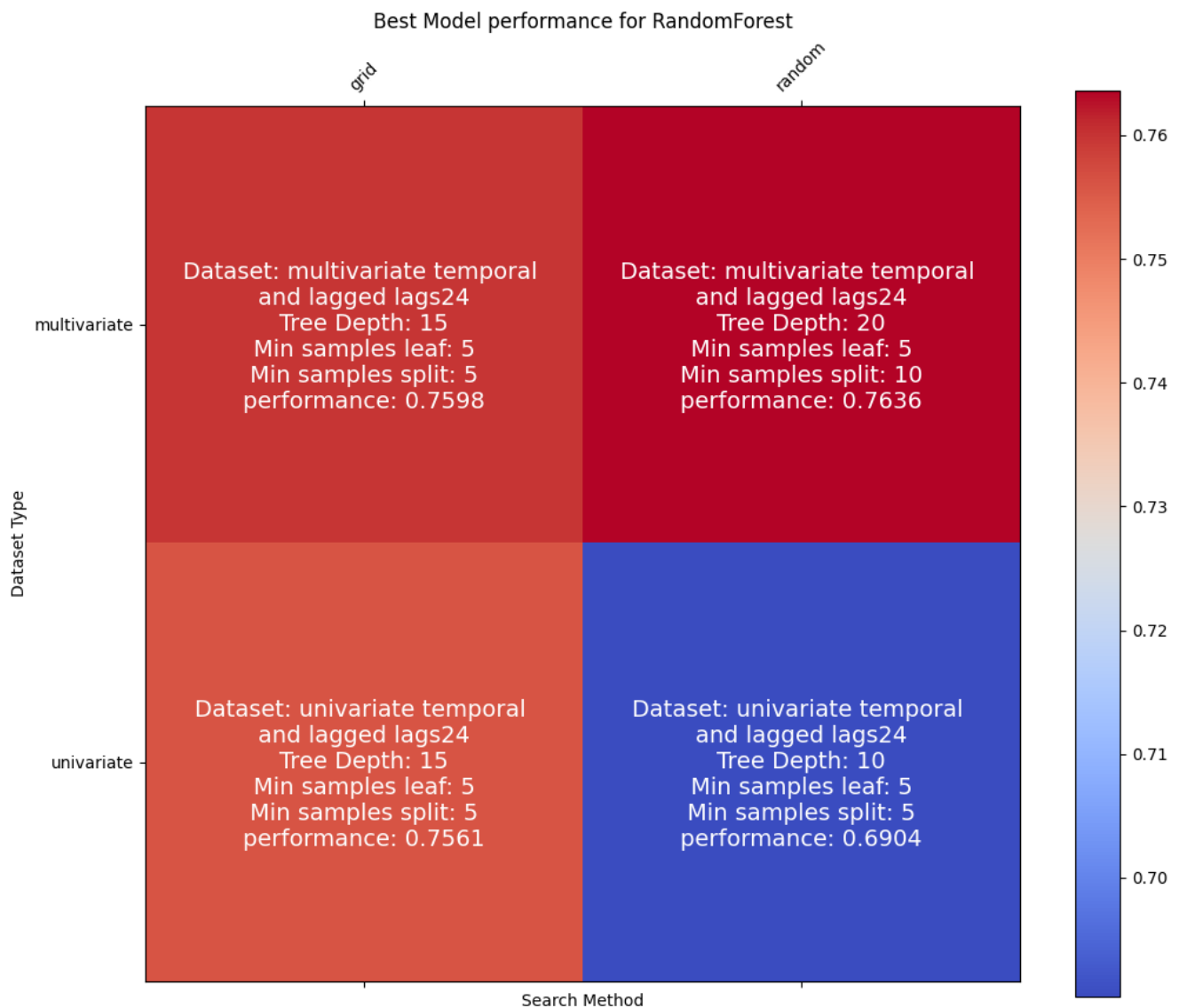


Multistep forecast for RandomForest_multistep_multivariate_multistep_temporal_and_lagged_lags24_steps24, Power Consumption, average performance 0.9024835250683263, step score 0.8528379066438662

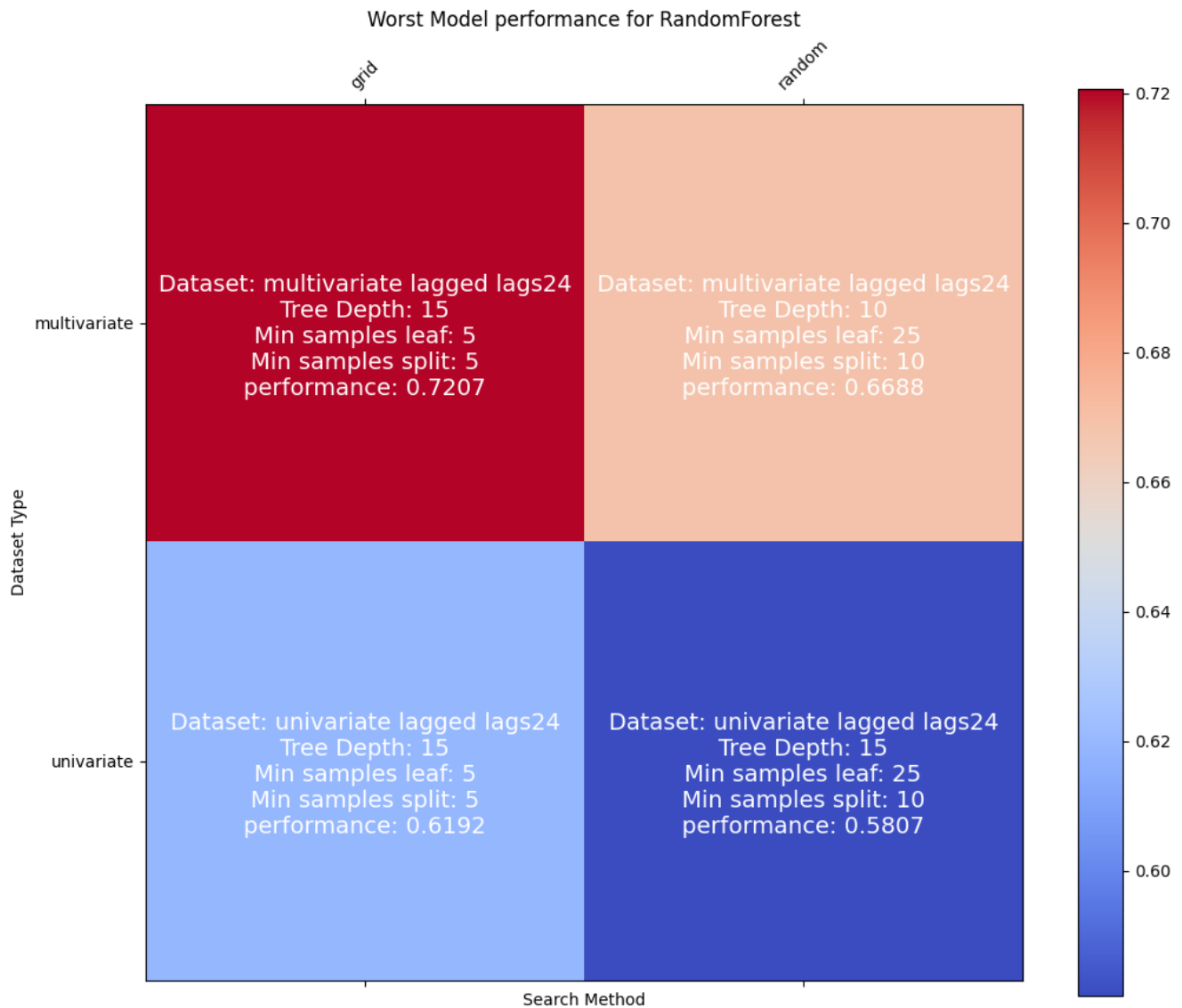


Results for the recursive forecast.

The best results (recursive) for the different setup combinations are as follows:

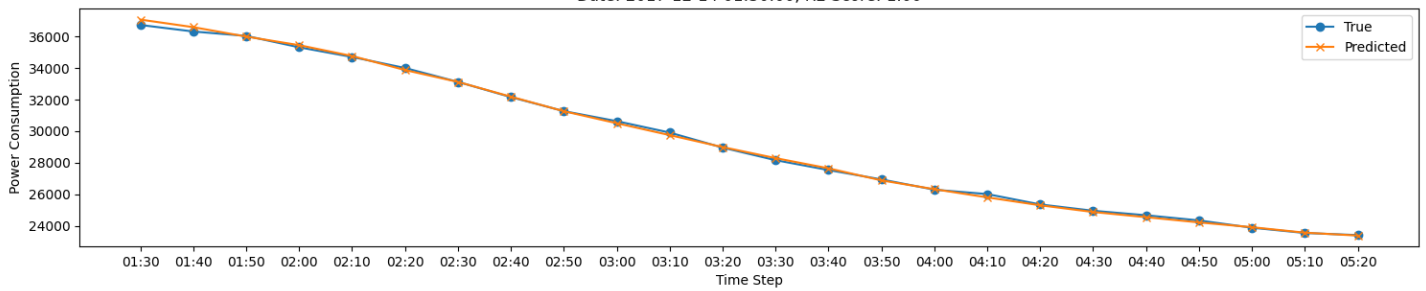


The worst results (recursive) for the different setup combinations are as follows:

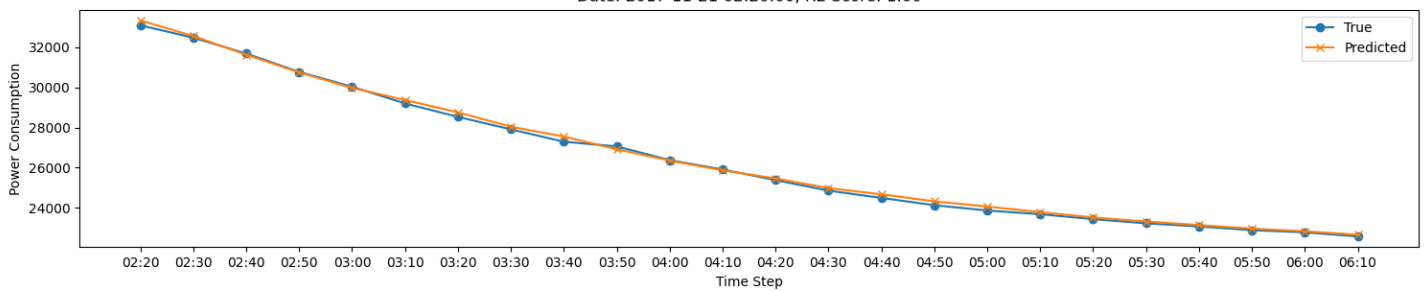


Best predicted days for RandomForest.

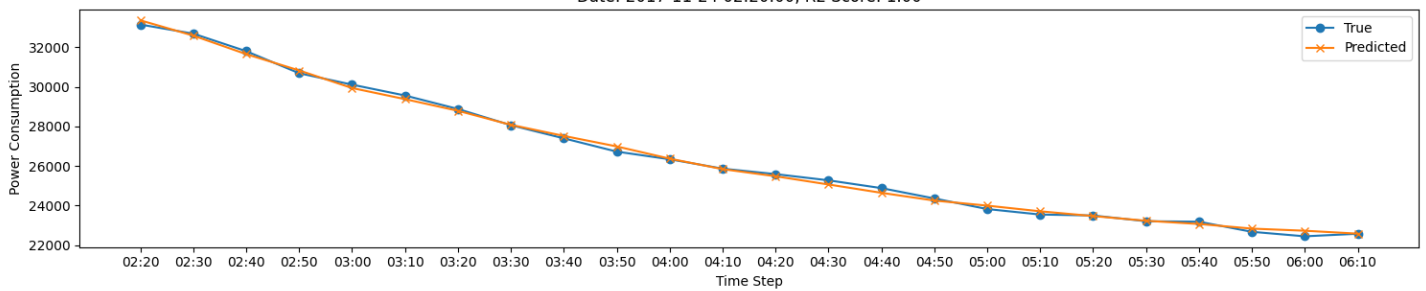
Date: 2017-12-14 01:30:00, R2 Score: 1.00



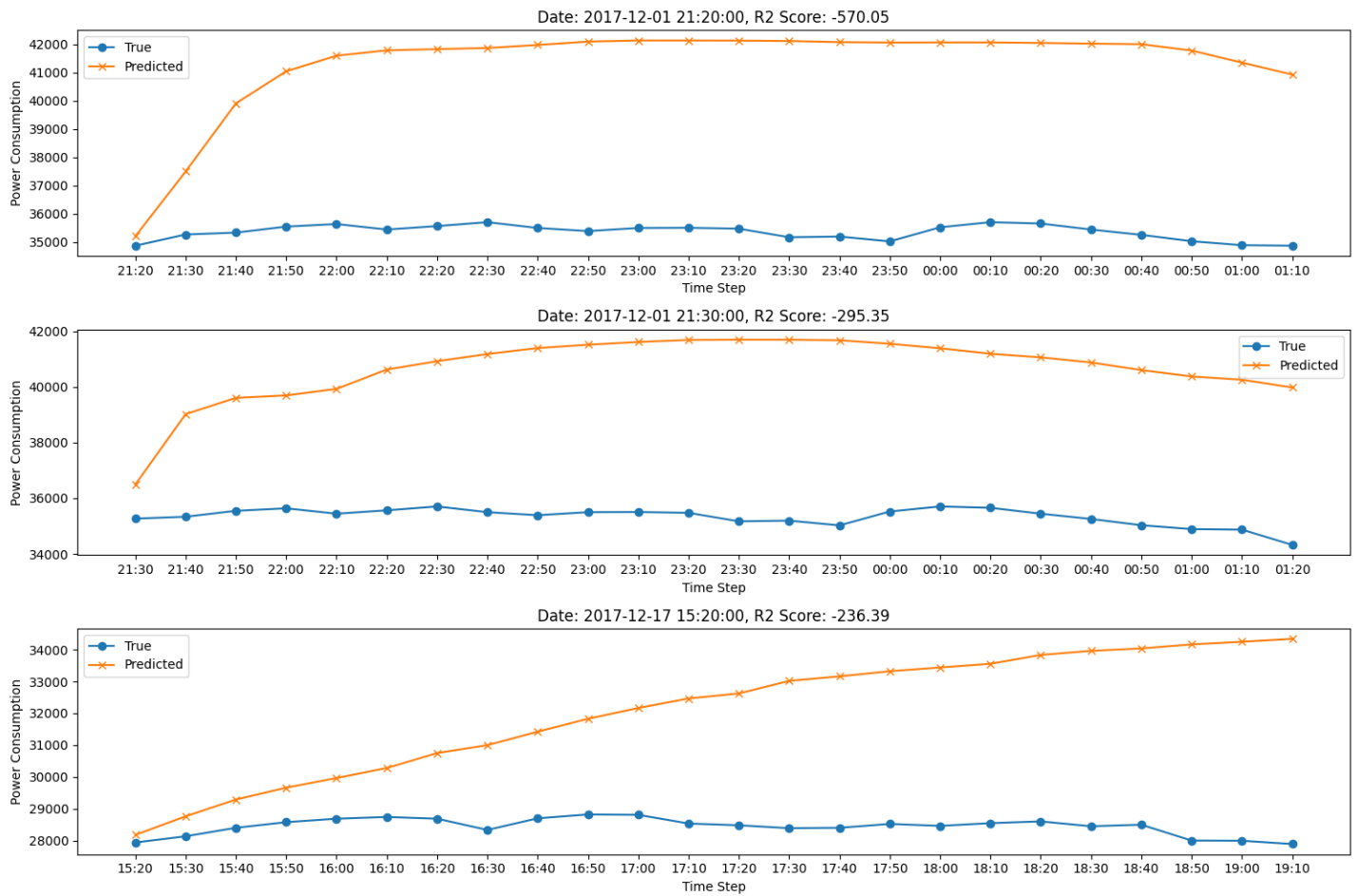
Date: 2017-11-21 02:20:00, R2 Score: 1.00



Date: 2017-11-24 02:20:00, R2 Score: 1.00

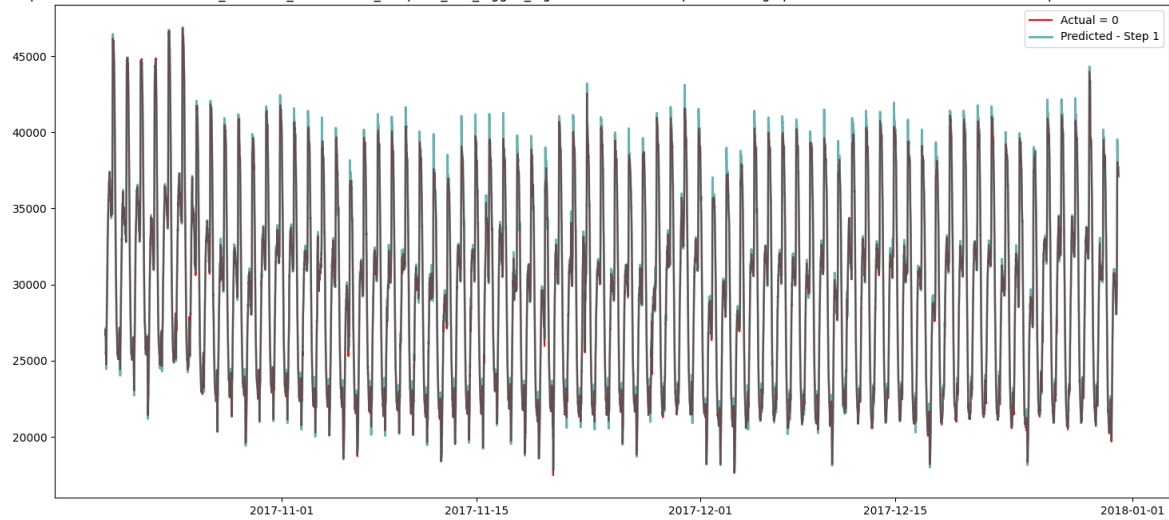


Worst predicted days for RandomForest.

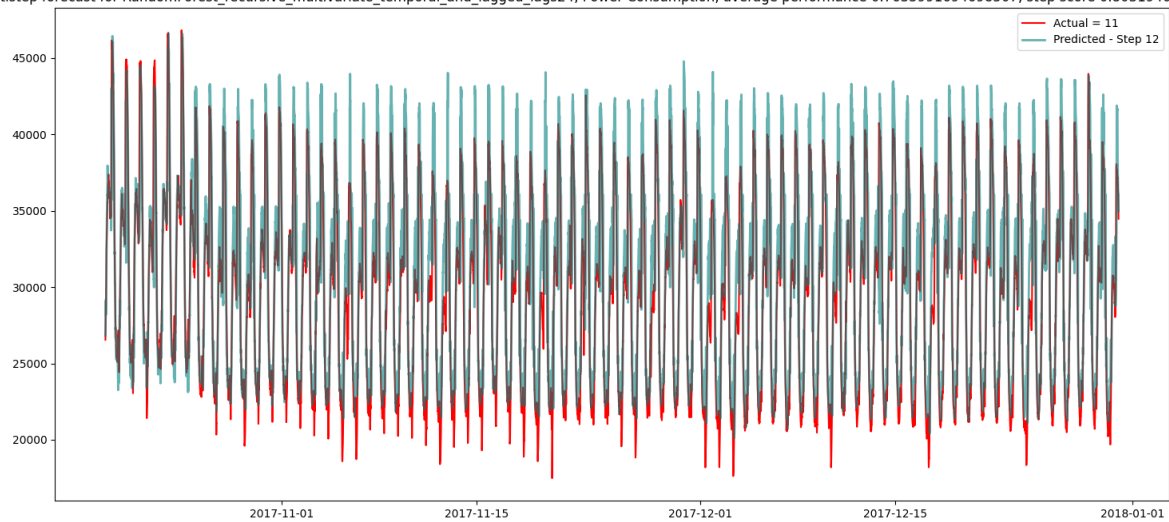


Steps plots for RandomForest forecasts over time

Multistep forecast for RandomForest_recursive_multivariate_temporal_and_lagged_lags24, Power Consumption, average performance 0.7635991694698507, step score 0.9961547674405665



Multistep forecast for RandomForest_recursive_multivariate_temporal_and_lagged_lags24, Power Consumption, average performance 0.7635991694698507, step score 0.8031940126261123



Multistep forecast for RandomForest_recursive_multivariate_temporal_and_lagged_lags24, Power Consumption, average performance 0.7635991694698507, step score 0.42448394564897174

