Experiment

Creator **Olusanmi Hundogan** Created **Feb 29, 2024, 21:26** Last updated **Apr 25, 2024, 13:08**

Files

- experiment_bucket_optimum.py
- analyze_bucket_optimumpy

Motivation

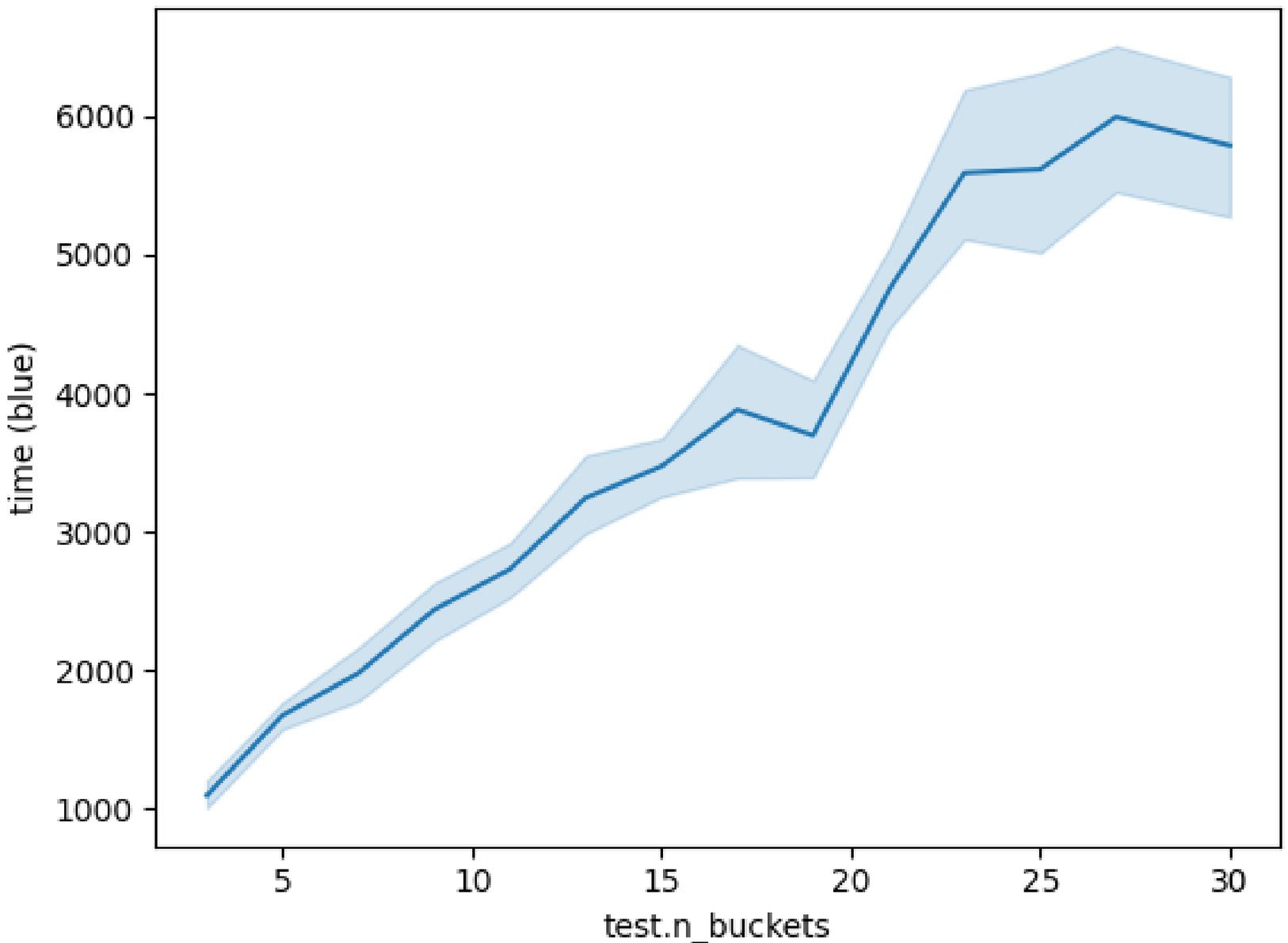
The MMA model uses a two-step approach to prediction. The model trains a VotingRegression model for different buckets of scope values. For this, we need a reliable method to classify an expected bucket of a datapoint. We solve this by also training a classification model before prediction. The number of buckets is a hyperparameter that determines the training time, size and performance of our model, as each bucket requires an extra set of models that vote on the regressed value. In this experiment we investigate the interplay between time and performance of the model.

Design

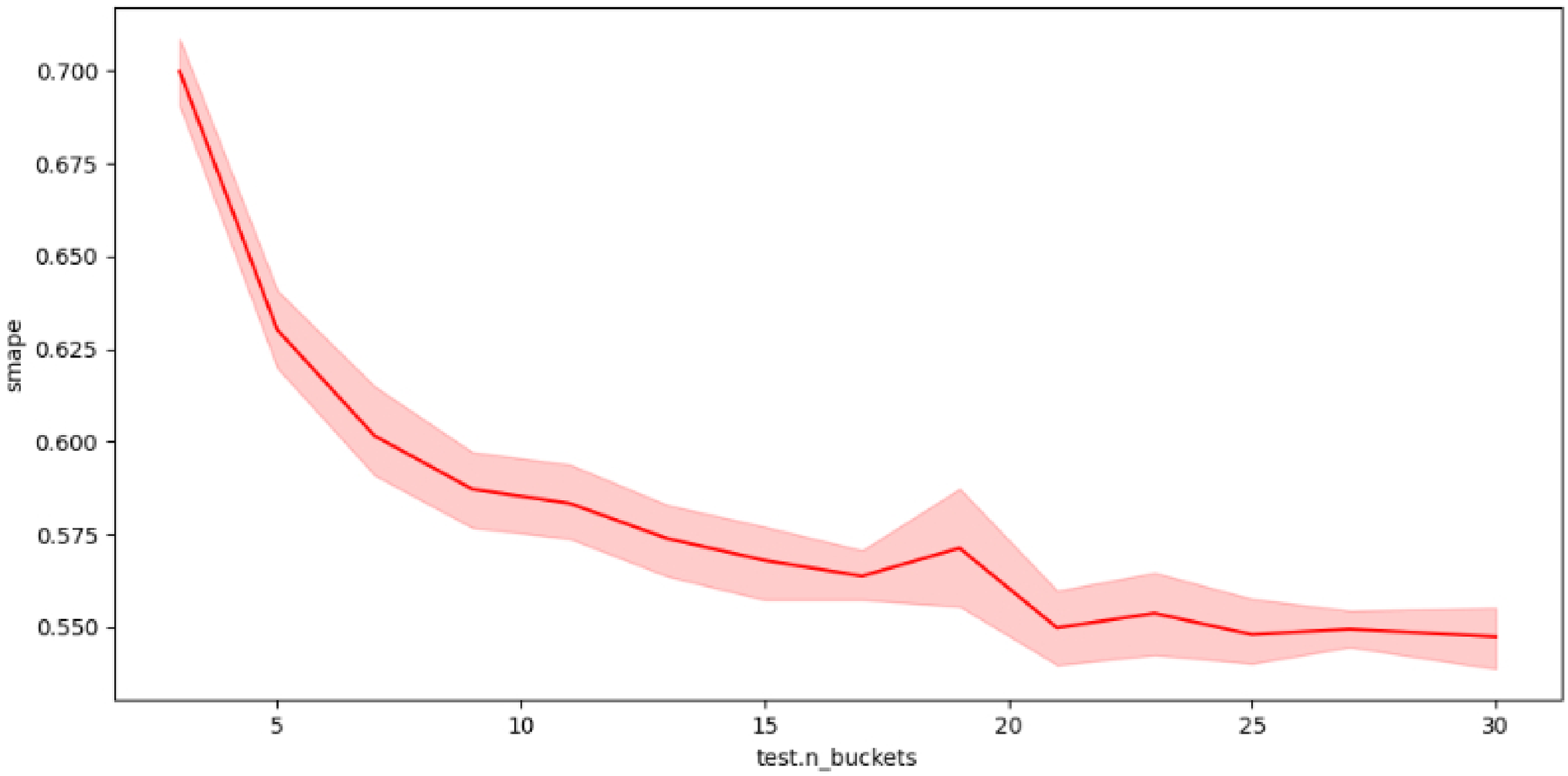
For this experiment, we train an MMA model for a different number of buckets and evaluate the sMAPEvalue on scope 1. The range of buckets we investigate starts with 3 and goes up to 30 in steps of 2. We run the experiment for 10 repetitions for each configuration. We run the experiments without dimensionality reduction.

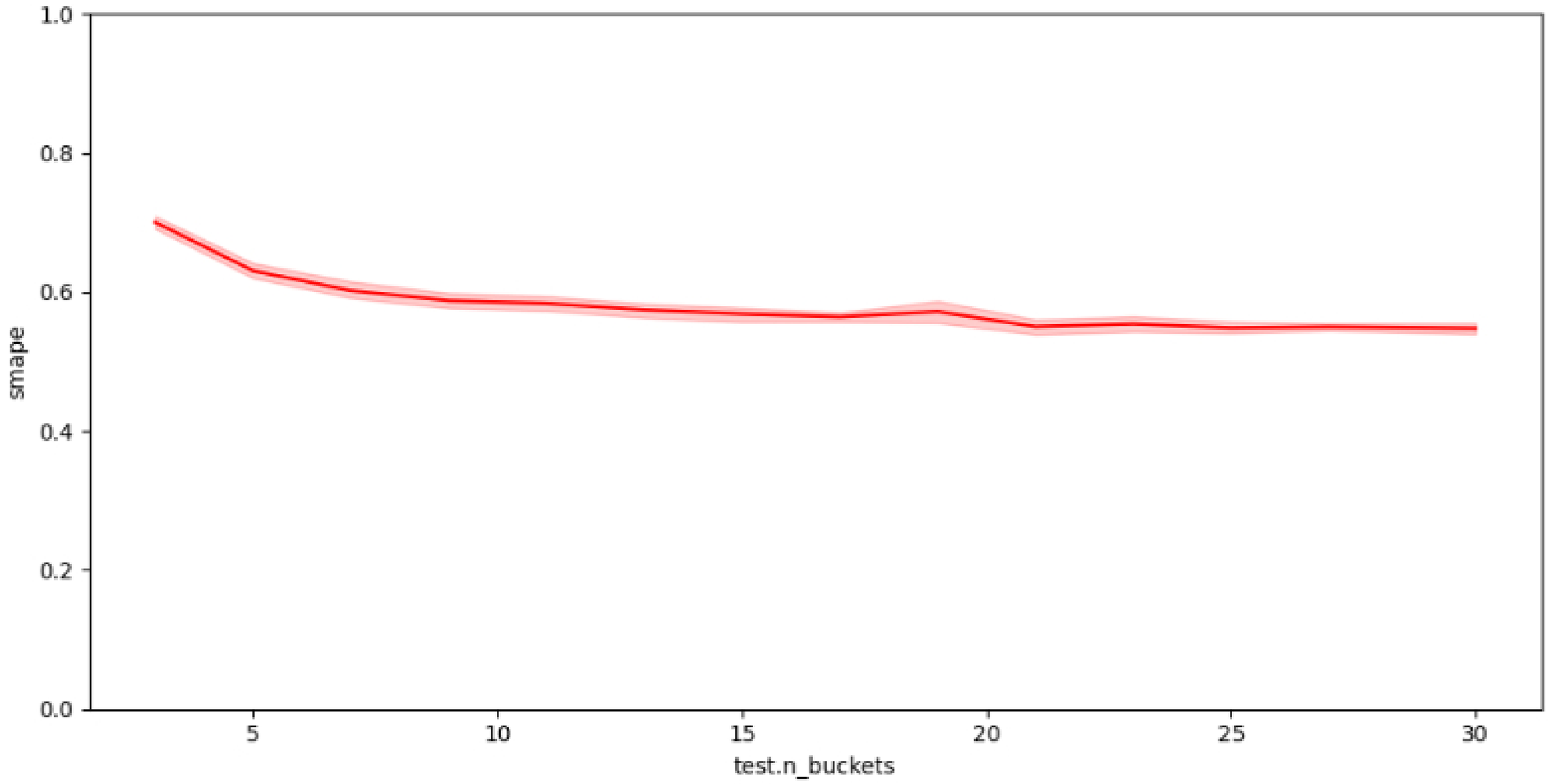
Results and Insights

The bucket number displays a mostly linear relationship to the overall fitting time. (The processing times were slightly contaminated by parallel experiments running on the same machine.) This result was expected.

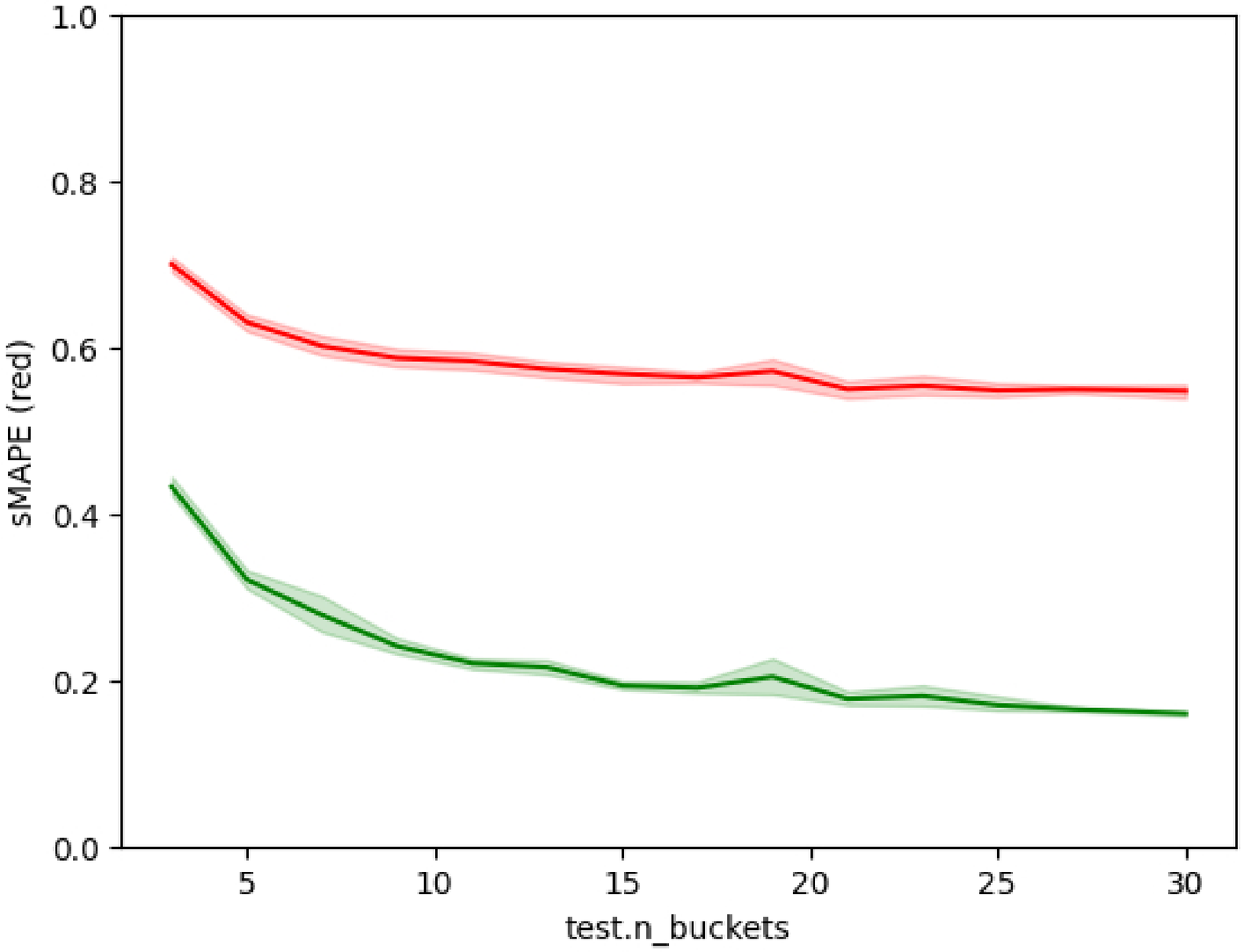


In terms of performance, the results show an equally expected behaviour. The bucket number directly contributes to the sMAPEperformance. However, we also see a plateau for higher numbers of buckets. The plot shows the results on a "zoomed" scale and one on a "widened".





We also investigate the overfitting behavior of the model. The results show the model overfits more if the number of buckets is increased. However, the model seems to be overfitting from the start. The green plot shows the training sMAPEresults and the red the test sMAPE.



Decision

Update 29.02.24

Based on the results, it is advisable to choose a bucket number which is providing significant results while being reasonably fast to train. We opt for 10 as a reasonable number. Everything further does not warrant the increase in time.

Update 25.03.24

Given an updated experiment running with the new data the reading changes a little. The new results suggest 15 being the optimal value. However, this was run on a small sample of the dataset. As a takeway, we can conclude that the optimal value is between 10-15 buckets.

Image failed to load