**Results for M3 Monthly Micro Data vs. M4 Weekly Finance Data**

*Important Points:*

* These results are for a 1-step (h=1) forecast horizon.
* The m4 weekly finance data was chosen to obtain initial results since it contains a relatively low number of series (164) and would enable quick code implementation.
* The m3 monthly micro data was chosen for comparison because machine learning methods perform well on this data, and it provides a reasonably high number of time series (474).
* The pre-processing performed for the exponential smoothing models is relatively minimal – we allow the models to capture the important aspects of the time series.
* We perform pre-processing for the LGBM models that is consistent with best practice.

Steps in Analysis:

1. Obtain original series
2. Create protected dataset
3. Perform pre-processing on and :
   1. For SES, DES, TES:
      1. As needed for , Truncate values < 1 to 1 (enables taking the log)
      2. Take the log (removes multiplicative effects and stabilizes variance) – helps with model convergence.
   2. For LGBM (Hewamalage et al., 2022; Makridakis et al., 2018):
      1. As needed for , Truncate values < 1 to 1 (enables taking the log)
      2. Mean normalize to put data on same scale across series
      3. Take the log (removes multiplicative effects and stabilizes variance)
      4. Perform conditional deseasonalization (conditional on 90% confidence auto-correlation test): - *I believe this is the equation used, which is described in the M4 competition benchmark documentation. I need to verify with python package developers that this is the case.*
      5. Convert time series into windows
      6. Normalize each window by removing the trend of each window
4. Train each forecasting model:
   1. SES, DES(additive trend), TES(additive trend and seasonality), LGBM(window length= 20)
   2. Generate forecasts for 1-step horizon for and
5. Post process the forecasts
   1. Reverse the pre-processing, including bias correction for exponential transformation
6. Compare forecast metrics:
   1. MAPE (equal to MdAPE under one-step horizon)
   2. **See Results Below**

**Results for Monthly M3 Micro Data**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Error Metric | Forecasting Model | Original | DP () | Add. Noise (1 SD) | Top Coding (10%) | Bottom Coding (10%) |
| MAPE | SES | 0.262 | **0.336** | **0.298** | 0.257 | 0.281 |
| DES | 0.252 | 0.405 | 0.354 | 0.247 | 0.268 |
| TES | 0.230 | 0.514 | 0.492 | **0.223** | **0.249** |
| LGBM | **0.227** | 0.505 | 0.380 | **0.223** | 0.261 |

* *Compare values within columns first (most accurate are highlighted)*
  + LGBM is best on original data
  + SES is best under DP and additive noise (by a lot)
  + LGBM/TES compete under top and bottom coding
* *Comparing values across columns*
  + All methods have worse accuracy under DP, additive noise, and bottom coding
  + All methods have improved accuracy under top coding relative to the original data (we will see if we can relate this to downward adjusting forecasts in the judgmental forecasting literature)

**Results for Weekly M4 Finance Data**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Error Metric | Forecasting Model | Original | DP () | Add. Noise (1 SD) | Top Coding (10%) | Bottom Coding (10%) |
| MAPE | SES | **0.023** | 0.251 | **0.184** | 0.120 | 0.072 |
| DES | **0.023** | **0.248** | 0.204 | **0.119** | **0.071** |
| TES | 0.026 | 0.344 | 0.395 | 0.120 | 0.074 |
| LGBM | 0.032 | 0.382 | 0.477 | 0.125 | 0.078 |

* Compare values within columns first
  + SES/DES are best on original data and under data protection
* Comparing values across columns
  + All methods have worse accuracy under any data protection

\*We will explain the differences between these two datasets using time series features (trend, seasonality, autocorrelation, etc.) and how these affect the accuracy of forecasting models after data protection.

**References**

Hewamalage, H., Bergmeir, C., & Bandara, K. (2022). Global models for time series forecasting: A Simulation study. *Pattern Recognition*, *124*, 108441. https://doi.org/10.1016/j.patcog.2021.108441

Makridakis, S., Spiliotis, E., & Assimakopoulos, V. (2018). Statistical and Machine Learning forecasting methods: Concerns and ways forward. *PLOS ONE*, *13*(3), e0194889. https://doi.org/10.1371/journal.pone.0194889