Data
Lab for
Social Good
Cardiff University, UK

Cardiff University
Prifysgol Caerdydd
Cardiff Business School
Ysgol Busnes Caerdydd

To aggregate or not to aggregate: Forecasting of finiteautocorrelated series

JORS Research in Focus Webinar - 25 July 2025

Bahman Rostami-Tabar, Professor of Analytics
Director of Data Lab for Social Good Reserach Group, Cardiff University
Read paper:https://doi.org/10.1080/01605682.2022.2118631
Slides: https://bahmanrt.netlify.app/talk/

Outline

- Introduction & Framing
- Literature Background & Contributions
- Model Setup
- Results: Analytical & Empirical
- Conclusion, Limitations & Future Work

Outline

- Introduction & Framing
- Literature Background & Contributions
- Model Setup
- Results: Analytical & Empirical
- Conclusion, Limitations & Future Work

JOURNAL OF THE OPERATIONAL RESEARCH SOCIETY
2023, VOL. 74, NO. 8, 1840–1859
https://doi.org/10.1080/01605682.2022.2118631

RESEARCH ARTICLE

To aggregate or not to aggregate: Forecasting of finite autocorrelated demand

Bahman Rostami-Tabar[a], M. Zied Babai[b] and Aris Syntetos[a]

[a]Cardiff Business School, Cardiff University, Cardiff, UK; [b]Kedge Businesss School, Talence, France

ABSTRACT

Temporal aggregation is an intuitively appealing approach to deal with demand uncertainty. There are two types of temporal aggregation: non-overlapping and overlapping. Most of the supply chain forecasting literature has focused so far on the former and there is no research that analyses the latter for auto-correlated demands. In addition, most of the analytical research to-date assumes infinite demand series' lengths whereas, in practice, forecasting is based on finite demand histories. The length of the demand history is an important determinant of the comparative performance of the two approaches but has not been given sufficient attention in the literature. In this article, we examine the effectiveness of temporal aggregation for forecasting finite auto-correlated demand. We do so by means of an analytical study of the forecast accuracy of aggregation and non-aggregation approaches based on mean-squared error. We complement this with a numerical analysis to explore the impact of demand parameters and the length of the series on (comparative) performance. We also conduct an empirical evaluation to validate the analytical results using monthly time series of the M4-competition dataset. We find the degree of auto-correlation, the forecast horizon

ARTICLE HISTORY
Received 2 April 2021
Accepted 12 August 2022

KEYWORDS
Non-overlapping temporal aggregation; overlapping temporal aggregation; time series forecasting; auto-correlated demand; exponential smoothing

Should we forecast at hourly or weekly frequency?

Motivating example: hospital admissions

- Hourly data on patient admissions is available.
- The planning task requires weekly forecasts to allocate resources.
- Key decision: Forecast using hourly data directly and then aggregate forecast or aggregate to weekly and then forecast?

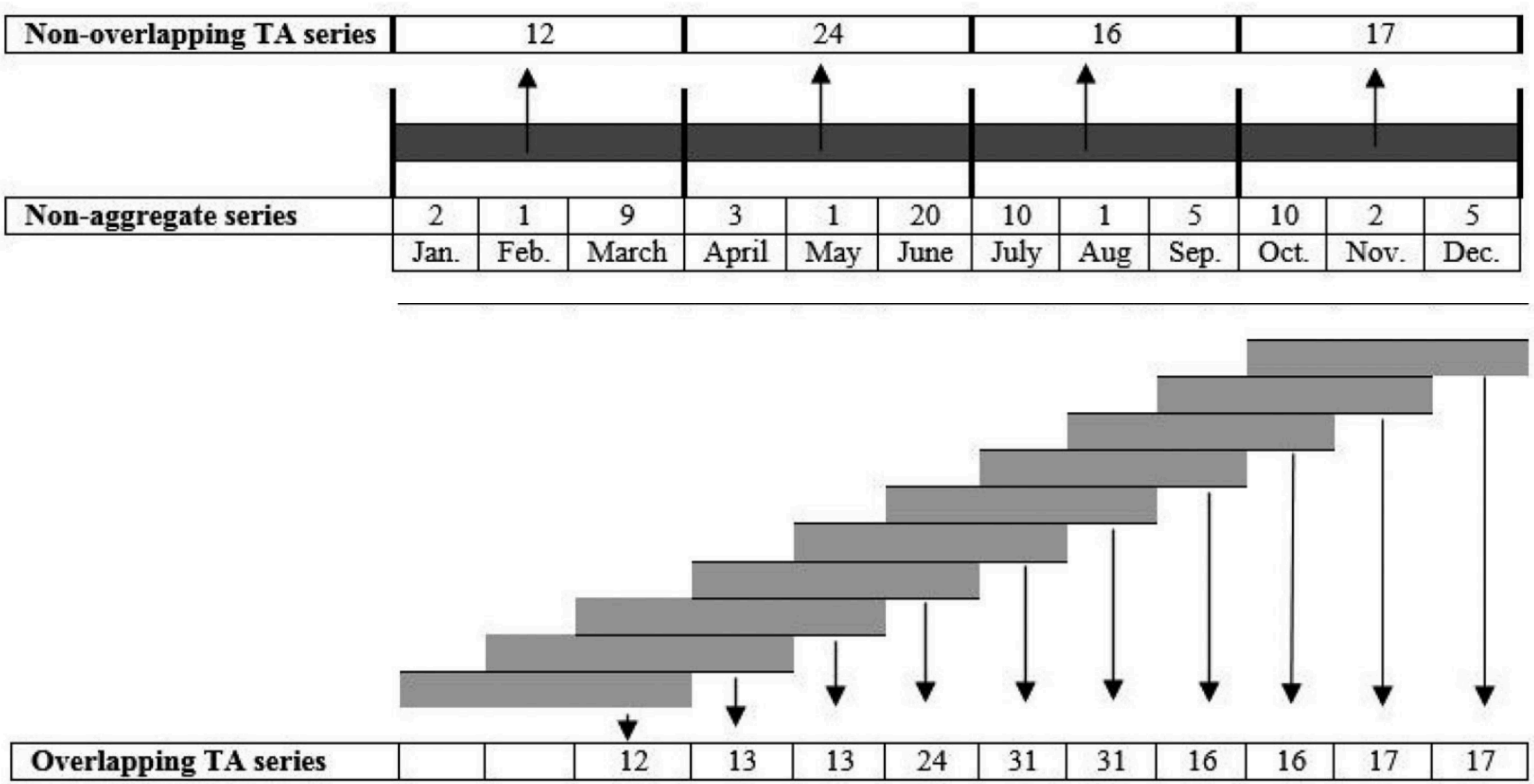
Common across domains and time granulrity

Data is often collected at higher frequency (e.g., Minutes, hours) than the forecasting target frequency (e.g., day, week, month, quarter).

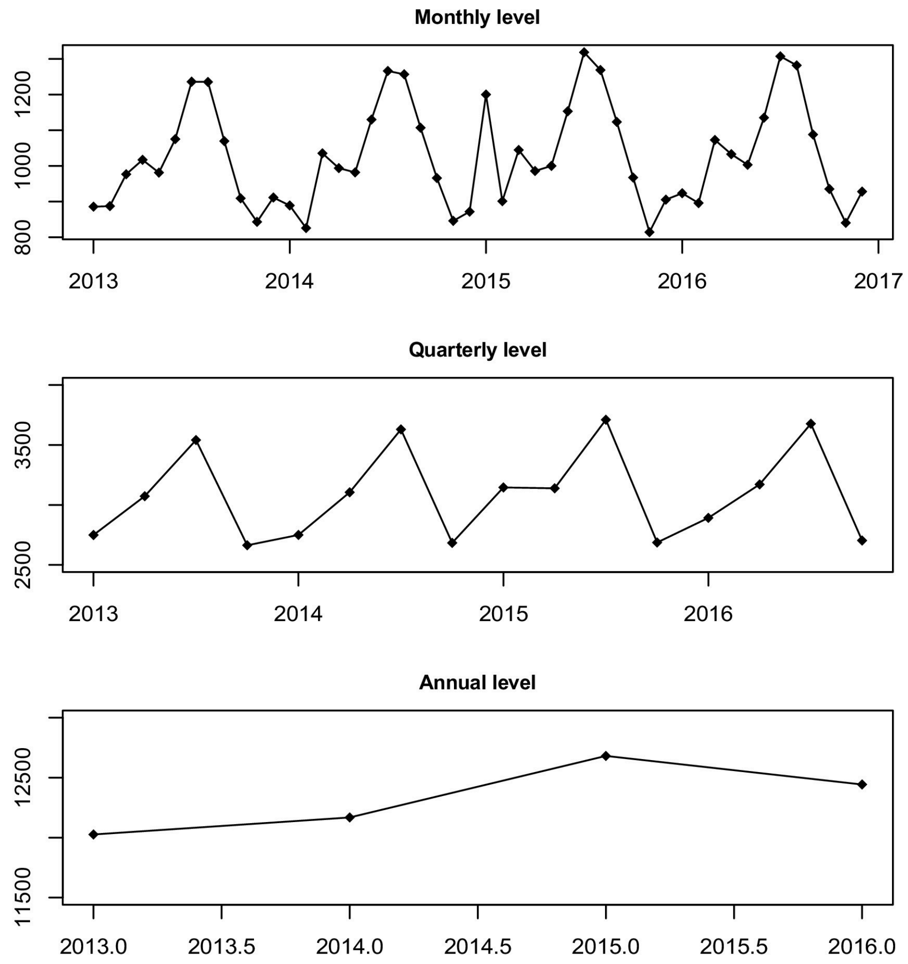
- Retail
- Energy
- Transportation
- Finance
- Manufacturing
- Agriculture
- and more

Temporal aggregation

Transforming higher-frequency data into lower-frequency data



Temporal aggregation - a tardeoff

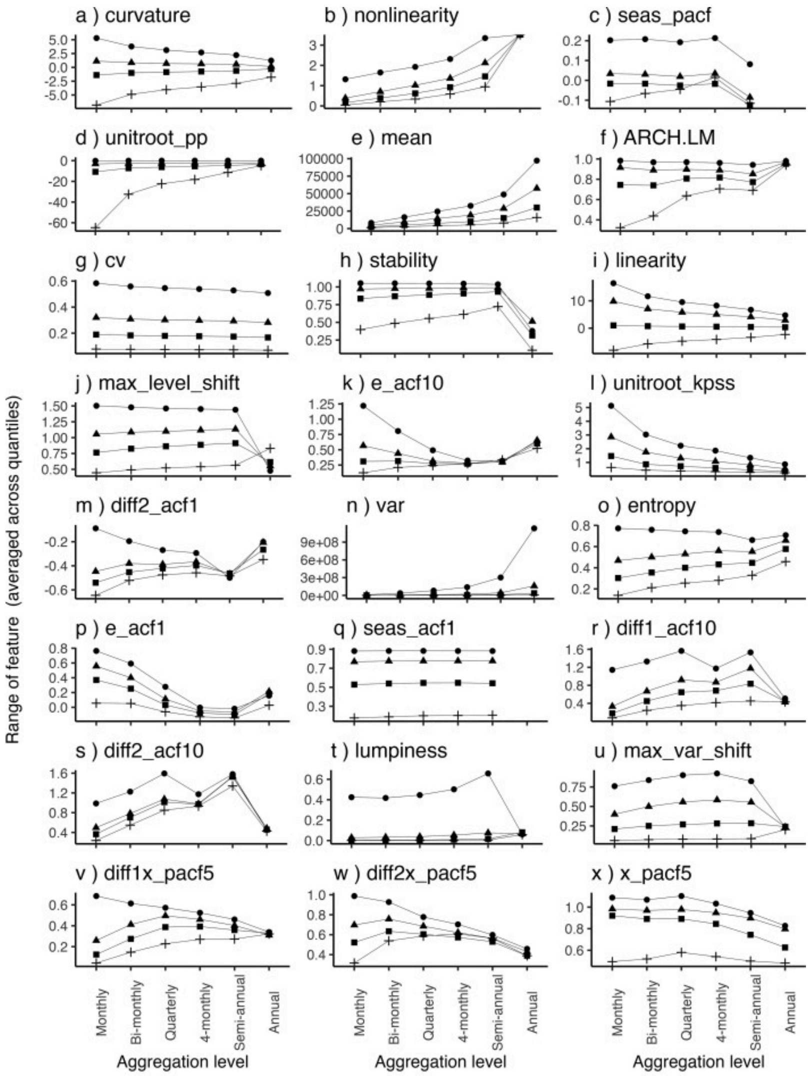


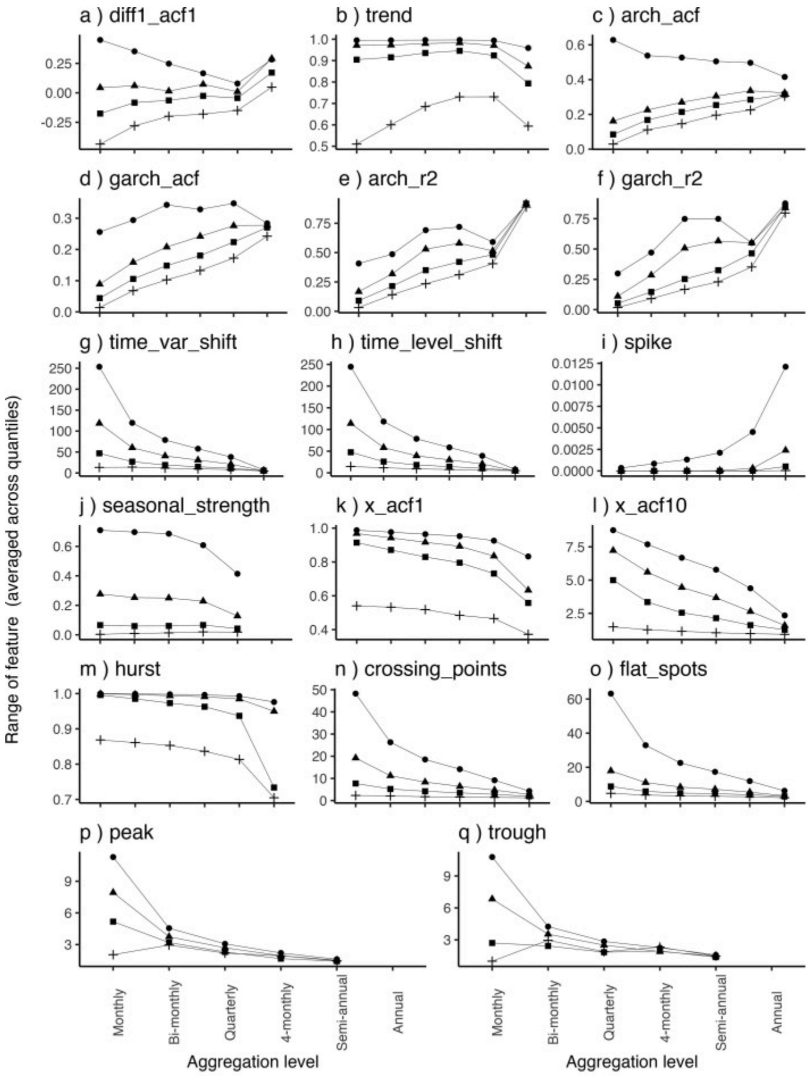
- Signal-to-noise
- Model complexity

Kourentzes, Nikolaos, Bahman Rostami-Tabar, and Devon K. Barrow. "series forecasting by temporal aggregation: Using optimal or multiple aggregation levels?." Journal of Business Research 78 (2017): 1-9.

How TA affects time series features

Rostami-Tabar & Mircetic. Neurocomputing 548 (2023): 126376.





Outline

- Introduction & Framing
- Literature Background & Contributions
- Model Setup
- Results: Analytical & Empirical
- Conclusion, Limitations & Future Work

Temporal aggregation - a very brief history

- The term temporal aggregation (TA) emerged in the context of econometrics and time series analysis, dating back to the 1970s.
- "TA affects the specification of models, estimation of parameters an efficiency of forecasting" (Brewer (1973), Wei (1979).

The term became particularly important in studies on:

- Macroeconomic modeling
- Autoregressive Integrated Moving Average (ARMA) processes.
- Forecasting

Brewer, K.R.W. (1973). Some consequences of temporal aggregation and systematic sampling for ARIMA and ARMAX models. J. Econometrics 1, 133-154.

Wei, W.W.S. (1979). Some consequences of temporal aggregation in seasonal time series models. In Seasonal Analysis of Economic Time Series, Ed. A. Zellner, pp. 433-444. Washington, D.C.; U.S. Department of Commerce, Bureau of the Census

Two distinct approaches to forecasting with temporal aggregation

Approach 1: understanding and optimizing

- Investigates how and when TA improves forecast accuracy.
- Focuses on finding the optimal aggregation level for a given forecasting task.
- Evaluates trade-offs between noise reduction and information loss.

Approach 2: Combining information across temporal levels

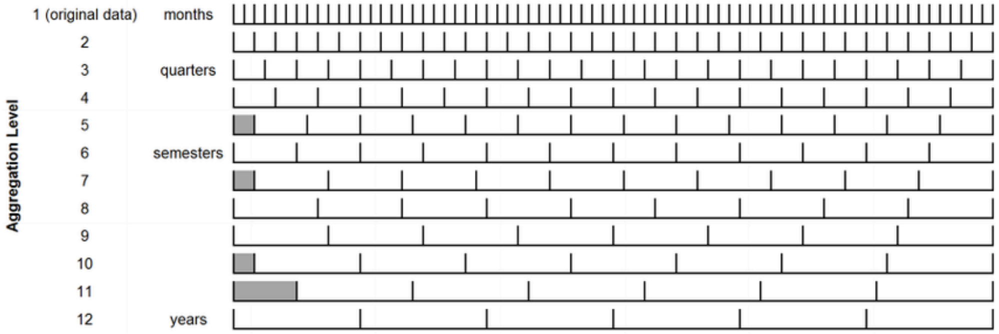
- Leverages data from multiple levels of aggregation simultaneously (e.g., hourly + daily + weekly).
- Aims to improve forecast performance through multi-scale modeling or reconciliation.
- Reflects the hierarchical nature of many real-world decision processes.

How and when TA is useful? Finding optmial aggregation level

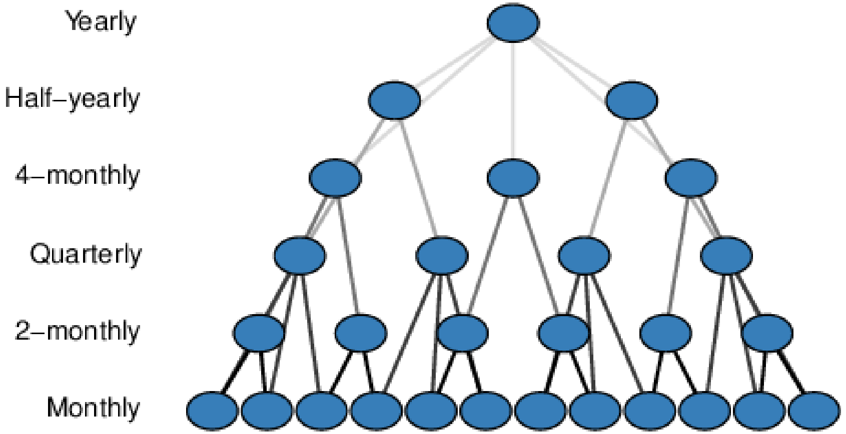
- Nikolopoulos, Konstantinos, et al. JORS 62.3 (2011): 544-554.
    - Empirical evaluation on intermittent series
    - TA can improve accuracy of forecasts
    - There might be an optimal aggregation level
- Rostami-Tabar, Bahman, et al. Naval Research Logistics (NRL) 60.6 (2013): 479-498.
    - Assuming autocorrelated series, AR processes, and SES
    - Analytical MSE expressions for non-aggregated and non-overlapping aggregated series
    - We also provided an analytical proof showing when the non-overlapping TA approach outperforms the non-aggregated alternative.

Combining information from different levels

- Kourentzes, Nikolaos, et al. International Journal of Forecasting 30.2 (2014): 291-302.
  - Multiple temporal aggregation levels



- Athanasopoulos, George, et al. EJOR 262.1 (2017): 60-74.
  - Temporal Hierarchies



Motivation for this paper

- Previous research focused solely on non-overlapping temporal aggregation
- Previous research assumed infinite history length
- This paper considers both overlapping and non-overlapping temporal aggregation and compare with non-aggregation approach.

Objectives

[1.] We derive analytical MSE expressions under the three approaches when a finite history length is used.

[2.] We evaluate the performance of the three approaches by analysing the impact of the length of the series, the aggregation level and the process parameters on the forecast performance.

[3.] Using monthly time series from the M4 competition, we empirically evaluate the performance of the three approaches.

Outline

- Introduction & Framing
- Literature Background & Contributions
- Model Setup
- Results: Analytical & Empirical
- Conclusion, Limitations & Future Work

Assumption about data

We assume that the non-aggregated series d_t, follows an ARMA(1,1) process:

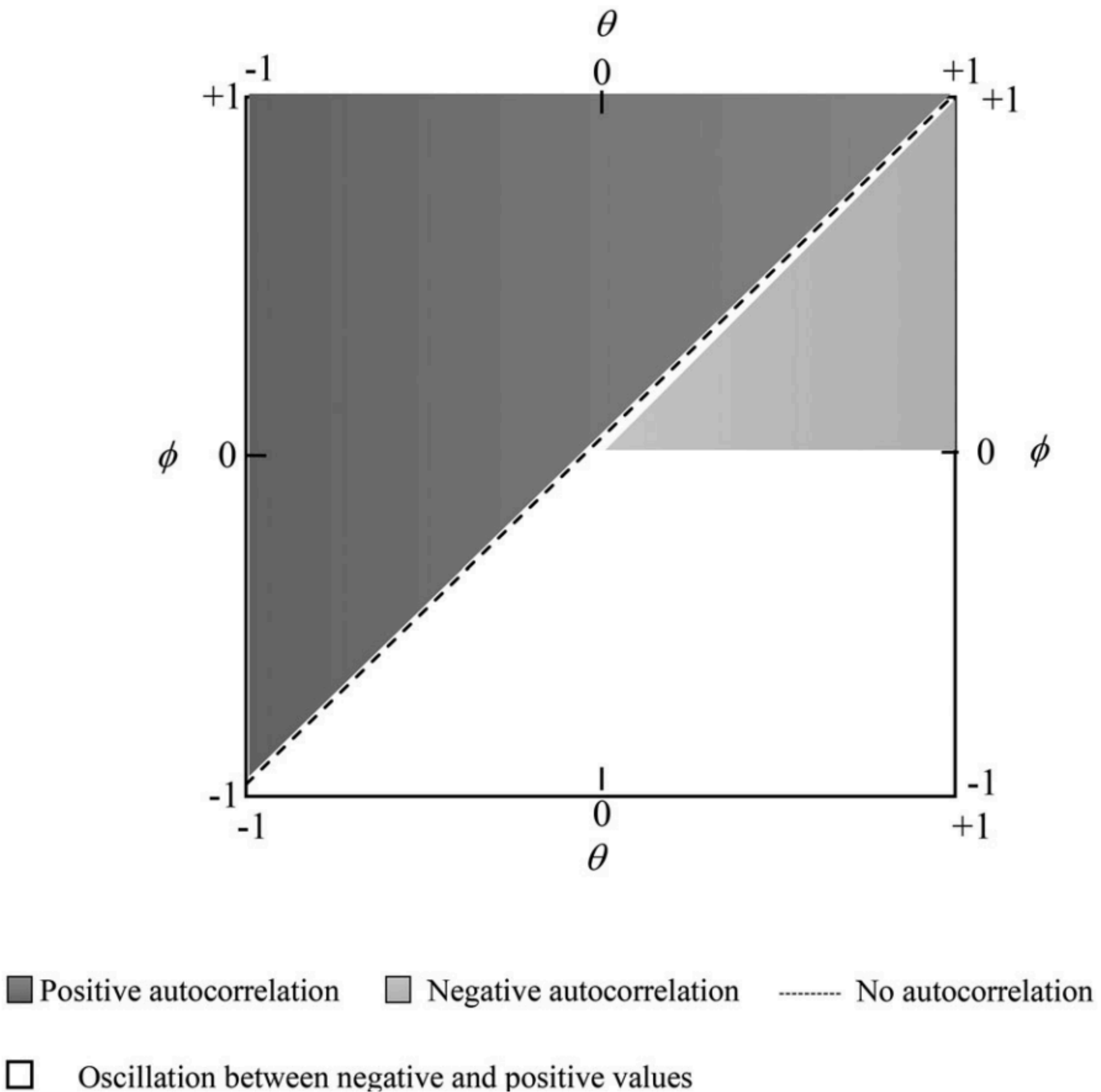
$$d_t = C + \epsilon_t + \phi d_{t-1} - \theta \epsilon_{t-1} \quad \text{where } |\theta| \leq 1,\ |\phi| \leq 1$$

with a constant C, autoregressive coefficient ϕ, and moving average coefficient θ, and $\epsilon_t \sim \mathcal{N}(0, \sigma^2)$

$\gamma_k = \text{Cov}(d_t, d_{t-k})$ =

$$\begin{cases} \frac{(1 - 2\phi\theta + \theta^2)}{1 - \phi^2}\sigma^2 & k = 0, \\ \frac{(\phi - \theta)(1 - \phi\theta)}{1 - \phi^2}\sigma^2 & |k| = 1, \\ \phi^{|k|-1}\gamma_1 & |k| > 1. \end{cases}$$

Autocorrelation associated with an ARMA(1,1) process



Forecast

- Simple Exponential Smoothing (SES) is used
- We aim to forecast the cumulative (aggregated) series, written as follows:

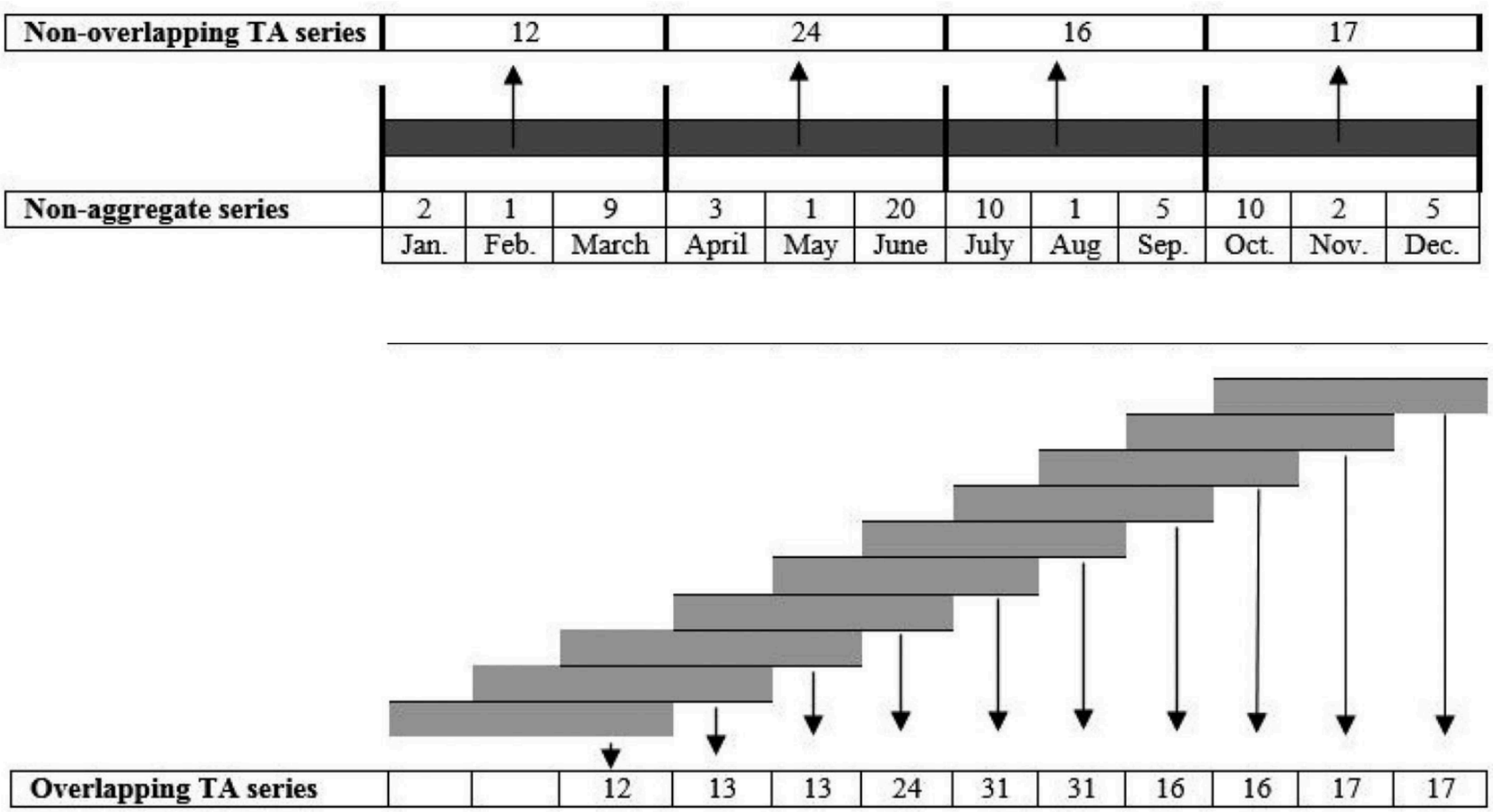
$$D_T = d_t + d_{t+1} + \cdots + d_{t+m-1}.$$

$$f_t = \sum_{k=1}^{N} \alpha(1-\alpha)^{k-1} d_{t-k} + (1-\alpha)^N f_0$$

$$F_{T,\mathrm{NOA}} = \sum_{k=1}^{\lceil \frac{N}{m} \rceil} \beta_N (1-\beta_N)^{k-1} D_{T-k,\mathrm{NOA}} + (1-\beta_N)^{\lceil \frac{N}{m} \rceil} F_{0,\mathrm{NOA}}$$

$$F^1_{T,\mathrm{OA}} = \sum_{k=1}^{N-m+1} \beta_0 (1-\beta_0)^{k-1} D_{T-k,\mathrm{OA}} + (1-\beta_0)^{N-m+1} F_{0,\mathrm{OA}}$$

Compare three approaches



MSE for three approaches

$$\text{MSE}_{\text{NA}} = \text{var}(D_T - f_t^m),$$

$$\text{MSE}_{\text{NOA}} = \text{var}\left(D_T - F_{T,\text{NOA}}^1\right),$$

$$\text{MSE}_{\text{OA}} = \text{var}\left(D_T - F_{T,\text{OA}}^1\right).$$

- NA: forecasting method applied to non-aggregated data, then summed over horizon m
- NOA: forecasting method applied to non-overlapping aggregated data but a direct model for the cumulative target,
- OA: forecasting method applied to overlapping aggregated data but a direct model for the cumulative target,

MSE of non-aggrgate approach

$$
\begin{aligned}
MSE_{NA} = {} & m\gamma_0 + \gamma_1 \left(\sum_{i=1}^{m-1} 2(m-k)\phi^{k-1} \right) + \\
& (m^2) \left(\begin{array}{l} \frac{\alpha\gamma_0\left(1-(1-\alpha)^{2N}\right)}{(2-\alpha)} + 2\alpha(1-\alpha)^{2N-1}\gamma_0 + (1-\alpha)^{2N}\gamma_0 + \\ \frac{2\alpha\gamma_1}{(2-\alpha)} \sum_{i=1}^{N-1}(1-\alpha)^i \phi^{i-1}\left(1-(1-\alpha)^{2(N-i)}\right) + \sum_{i=1}^{N-1} 2\alpha\phi^{N-i-1}(1-\alpha)^{N+i-1}\gamma_1 \end{array} \right) \\
& - \frac{2\alpha\gamma_1(1-\phi^m)}{(1-\phi)(1-\phi+\alpha\phi)} \times \left(1-(\phi-\alpha\phi)^N\right) + (1-\alpha)^N \phi^{N-1}\gamma_1 \times \frac{1-\phi^m}{1-\phi}.
\end{aligned}
$$

MSE of non-overlapping aggregation approach

$$MSE_{NOA} = \left(m\gamma_0 + \gamma_1\left(\sum_{k=1}^{m-1} 2(m-k)\phi^{k-1}\right)\right)\left(\frac{\beta_N\left(1-\left(1-\beta_N\right)^{2[\frac{N}{m}]}\right)}{(2-\beta_N)} + (1-\beta_N)^{2[\frac{N}{m}]} + 2\beta_N(1-\beta_N)^{2[\frac{N}{m}]-1}\right)$$

$$+m\gamma_0 + \gamma_1\left(\sum_{k=1}^{m-1} 2(m-k)\phi^{k-1}\right) + \sum_{i=1}^{([\frac{N}{m}]-1)} \frac{2\beta_N\gamma_1}{(2-\beta_N)}\left(\frac{1-\phi^m}{1-\phi}\right)^2 (1-\beta_N)^i \phi^{m(i-1)}(1-(1-\beta_N)^{2[\frac{N}{m}]-i})$$

$$-2\left(\beta_N\gamma_1 \times \left(\frac{1-\phi^m}{1-\phi}\right)^2\left(\frac{1-(\phi^m-\beta_N\phi^m)^{[\frac{N}{m}]}}{(1-\phi^m+\beta_N\phi^m)}\right) + \phi^{m([\frac{N}{m}]-1)}(1-\beta_N)^{2[\frac{N}{m}]}\left(\frac{1-\phi^m}{1-\phi}\right)^2\gamma_1\right).$$

MSE of overlapping aggregation approach

$$MSE_{OA} = m\gamma_0 + 2\gamma_1 \sum_{i=1}^{m-1} (m-i)\phi^{i-1} +$$

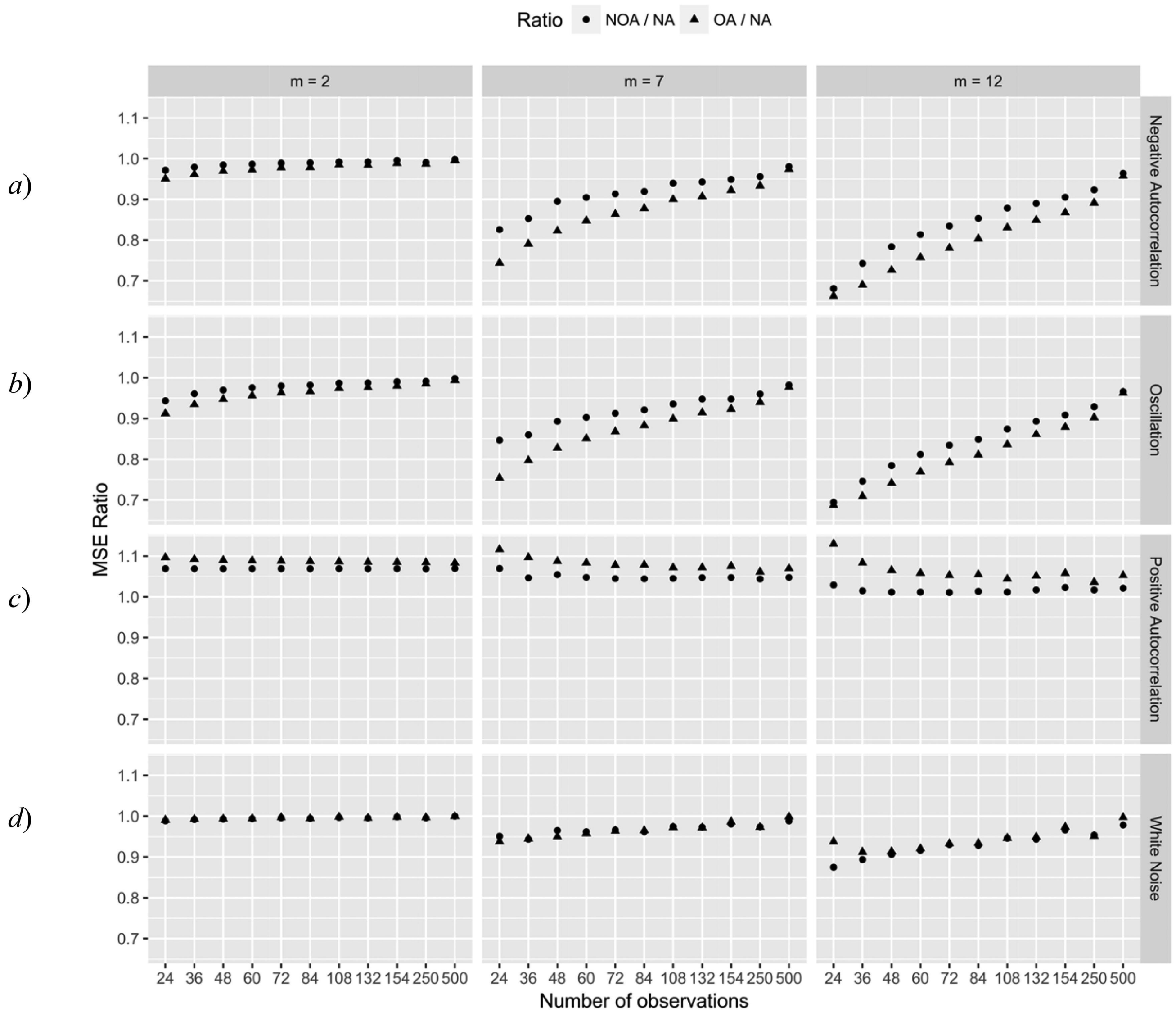
$$\frac{\beta_o \left(m\gamma_0 + 2\gamma_1 \sum_{i=1}^{m-1} (m-i)\phi \right) \left(1 - \left((1-\beta_o)^2 \right)^{N-m+1} \right)}{(2-\beta_o)} - 2\beta_O \gamma_1 \times \left(\frac{1-\phi^m}{1-\phi} \right)^2 \left(\frac{1-(\phi(1-\beta_O))^{N-m+1}}{1-\phi+\phi\beta_O} \right)$$

$$+ \frac{2\beta_o}{(2-\beta_o)} \sum_{k=1}^{N-m} (1-\beta_o)^k \left(\sum_{\substack{i=1 \\ \forall |k+j-i|=0}}^{m} \sum_{j=1}^{m} \gamma_0 + \sum_{\substack{i=1 \\ \forall |k+j-i|=1}}^{m} \sum_{j=1}^{m} \gamma_1 + \sum_{\substack{i=1 \\ \forall |k+j-i|>1}}^{m} \sum_{j=1}^{m} \phi^{\left(|k+j-i|-1 \right)} \gamma_1 \right) \left(1 - \left((1-\beta_o) \right)^{2(N-m+1-k)} \right),$$

Outline

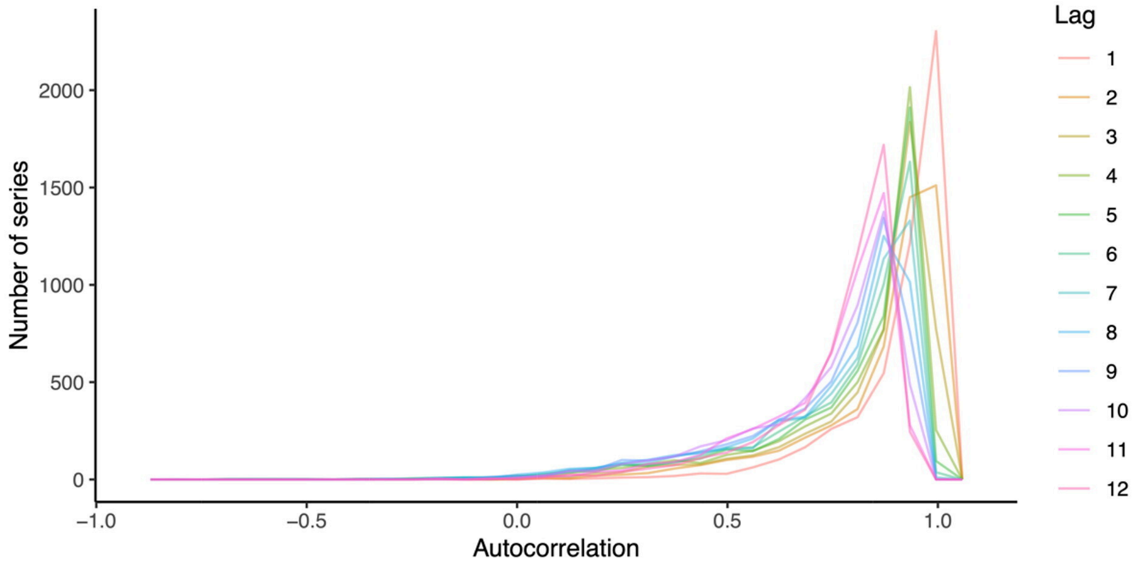
- Introduction & Framing
- Literature Background & Contributions
- Model Setup
- Results: Analytical & Empirical
- Conclusion, Limitations & Future Work

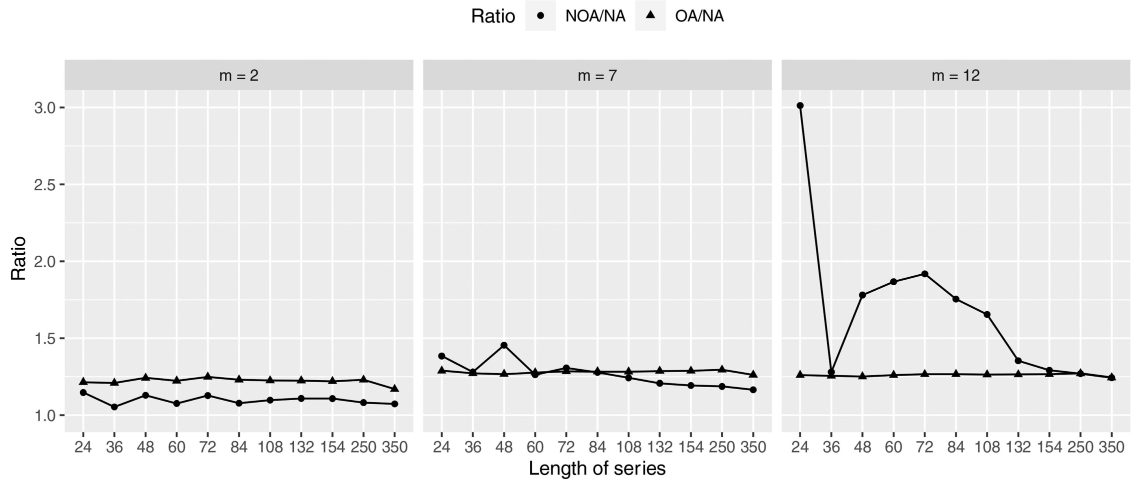
Results



Ratio
NOA / NA
OA / NA
m = 2
m = 7
m = 12
a)
b)
c)
d)
Negative Autocorrelation
Oscillation
Positive Autocorrelation
White Noise
MSE Ratio
1.1
1.0
0.9
0.8
0.7
24 36 48 60 72 84 108 132 154 250 500
Number of observations

Results - empirical data





Outline

- Introduction & Framing
- Literature Background & Contributions
- Model Setup
- Results: Analytical & Empirical
- Conclusion, Limitations & Future Work

Conclusion

The question we address is fundamental— persisting over time and remaining relevant across diverse domains and temporal granularities.

- High positive autocorrelation: Non-aggregated data yields lower MSEs.
- Negative autocorrelation: TA outperforms non-aggregated forecasts.
- Alternating autocorrelation signs: TA performs better than non-aggregated approaches.
- Longer forecast horizons: Both overlapping and non-overlapping TA show improved accuracy.
- Short time series: Overlapping TA is superior; differences diminish as history length increases.
- Diminishing returns: The improvement in forecast accuracy decreases slowly beyond a certain series length for TA approaches.

Limitations

- Data generation process: We assume that the disaggregated (non-aggregated) time series follows a stationary ARMA(1,1) process.
- Forecasting model: We rely exclusively on Simple Exponential Smoothing (SES) as the forecasting method.
- Forecasting horizon: The current framework focuses on generating a cumulative forecast over a fixed horizon M, effectively making it a one-step-ahead forecast in the aggregated setting.
- Empirical data: The empirical evaluation is limited to the M4 competition data, which primarily consists of positive autocorrelated series.

Future Work

Despite the contributions of this work, a key open question remains

> *There is still no general analytical framework that explains when and how temporal aggregation affects forecast accuracy.*
>
> *While specific cases (e.g., ARMA processes with SES) can be studied in isolation, a global understanding—one that applies across models, aggregation schemes, and forecasting horizons—remains elusive.*

Data
Lab for
Social Good

Cardiff University, UK

TO GET LATEST UPDATES

Join Our Mailing List via
http://eepurl.com/iAppLk

OR simply scan the QR code