A review of temporal aggregation and systematic sampling on time-series analysis

Rajaguru, Gulasekaran; Lim, Sheryl; O'Neill, Michael

Published in:
Journal of Accounting Literature

DOI:
10.1108/JAL-09-2024-0237

Received 7 September 2024
Revised 29 November 2024
Accepted 29 November 2024

A review of temporal aggregation and systematic sampling on time-series analysis

Gulasekaran Rajaguru
Centre for Data Analytics, Bond University Business School, Gold Coast, Australia
Sheryl Lim
Bond University Business School, Gold Coast, Australia, and
Michael O'Neill
Centre for Data Analytics, Bond University Business School, Gold Coast, Australia

Abstract

Purpose – This review investigates the effects of temporal aggregation and systematic sampling on time-series analysis, focusing on their influence on data accuracy, interpretability and statistical properties. The purpose of the study is to synthesise existing literature on the topic and offer insights into the trade-offs between these data reduction techniques.

Design/methodology/approach – The research methodology is based on an extensive review of theoretical and empirical studies covering univariate and multivariate time series models, focusing on unit roots, ARIMA, GARCH, cointegration properties and Granger Causality.

Findings – The key findings reveal that while temporal aggregation simplifies data by emphasising long-term trends, it can obscure short-term fluctuations, potentially leading to biases in analysis. Similarly, systematic sampling enhances computational efficiency but risks information loss, especially in non-stationary data, and may result in biased samples if sampling intervals coincide with data periodicity. The review highlights the complexities and trade-offs involved in applying these methods, particularly in fields like economic forecasting, climate modelling and financial analysis.

Originality/value – The originality and value of this study lie in its comprehensive synthesis of the impacts of these techniques across various time series properties. It underscores the importance of context-specific applications to preserve data integrity, offering recommendations for best practices in the use of temporal aggregation and systematic sampling in time-series analysis.

Keywords Temporal aggregation, Systematic sampling, Granger causality

Paper type Literature review

1. Introduction

A common challenge for researchers and analysts is the lack of time series data at the desired frequency. The use of highly temporally aggregated or systematically sampled data for inferences is quite common in the applied econometric literature.

In the field of time-series analysis, the techniques of temporal aggregation and systematic sampling are integral to managing and interpreting large volumes of data. Time series data, which records observations over specific intervals, is often complex and voluminous, making it challenging to analyse without the application of data reduction techniques. Temporal aggregation and systematic sampling serve as two primary methods to simplify this process, yet each has distinct implications for the accuracy, interpretability and reliability of the resulting analysis.

Funding: This research was funded by the Australian Government through the Australian Research Council (DP230101472).

Journal of Accounting Literature
Vol. 47 No. 5, 2025
pp. 110-128
Emerald Publishing Limited
e-ISSN: 2452-1469
p-ISSN: 0737-4607
DOI 10.1108/JAL-09-2024-0237

Temporal aggregation refers to the process of consolidating data points over a defined time period. For instance, hourly data may be aggregated into daily, weekly or monthly averages or sums. This technique is commonly employed to reduce data noise, emphasise longer-term trends and facilitate easier analysis and interpretation. Aggregation can be particularly useful in identifying broader patterns that might be obscured by short-term fluctuations. However, this simplification comes with trade-offs. By averaging or summing data, temporal aggregation may obscure important short-term variations and cyclical patterns, potentially leading to misleading conclusions. Additionally, the choice of aggregation interval can significantly influence the interpretation of data, as different intervals may highlight or conceal different aspects of the underlying time series.

Systematic sampling, on the other hand, involves selecting data points from a time series at regular intervals. This method is often used to reduce the size of the dataset, making it more manageable for analysis without the need to examine every data point. Systematic sampling can be particularly advantageous in scenarios where the original data is dense, as it allows for the analysis of a representative subset of the data, thereby saving time and computational resources. However, this method also introduces potential risks. If the sampling interval coincides with the periodicity or cyclic nature of the data, systematic sampling can result in biased samples, leading to inaccurate conclusions. Moreover, systematic sampling may inadvertently miss critical data points, especially in time series with irregular or complex patterns.

The interplay between temporal aggregation and systematic sampling is complex and can significantly impact the outcomes of time-series analysis. For researchers, analysts and decision-makers, it is crucial to understand how these techniques affect data characteristics and the potential biases they may introduce. This is particularly important in fields such as economics, finance, climate science and engineering, where accurate time-series analysis is essential for forecasting, trend analysis and policy modelling.

This study aims to explore the effects of temporal aggregation and systematic sampling on time-series data, examining how they affect the model dynamics, accuracy, reliability and interpretation of analysis results. By understanding the strengths and limitations of these techniques, we can better navigate the trade-offs involved and make more informed choices when analysing time series data. Ultimately, this knowledge contributes to more robust and reliable conclusions, helping to ensure that time-series analyses reflect the true dynamics of the underlying processes being studied.

2. Literature review

In this section, we conduct an extensive literature review to explore the impacts of temporal aggregation and systematic sampling on various time series econometric properties. Sections 2.1 and 2.2 analyse the effect of aggregation on univariate time series models while section 2.3 conducts extensive review on multivariate time series models with a special focus on causality in section 2.4.

2.1 Unit roots

The effect of temporal aggregation and systematic sampling on time series characteristics, such as unit roots, has been explored extensively over time. Granger (1990) derived an important result that if the disaggregated process x possesses an integration order of $I(d)$, the integration order of the aggregated process x remains $I(d)$. In a thorough overview of the consequences of temporal aggregation in empirical analysis, Marcellino (1999) further reiterated that unit roots are not affected by temporal aggregation. In particular, if the autoregressive component of the basic disaggregated process has the roots $\{\lambda_j, j = 1, \ldots, n\}$, then the corresponding roots for the aggregated process are given by $\{\lambda_j^m, j = 1, \ldots, n\}$ with m representing the order of the aggregation. Marcellino (1999) explained that the decrease in the absolute values of non-unit roots offsets the reduction in sample size. This reinforces that keeping the data span constant, the results of unit root tests are asymptotically consistent across various temporal aggregation methods and sampling frequencies.

Pierse and Snell (1995) demonstrated the asymptotic local power of a one-sided unit root test does not depend on systematic sampling or temporal aggregation, as long as the same data span is used when comparing models. They further pointed out that this result is intuitive because the power loss from reducing the data points by $\frac{m-1}{m}$ is made up by the increased separation of the unit root hypotheses, H_0 (being that there is one unit root) from H_1 (being the one-sided local alternative), induced by temporal aggregation.

In this context, seasonal unit roots are unsurprisingly an important point of study. Granger and Siklos (1995) demonstrated that seasonal unit roots present in high-frequency data might not appear in the aggregated data at lower frequencies. This issue is exclusive to systematic sampling, as temporal aggregation does not produce such results (Hylleberg *et al.*, 1993). Nonetheless, Silvestrini and Veredas (2005) reports that depending on the aggregation frequency chosen, temporal aggregation can cause a seasonal unit root to become a non-seasonal unit root. Franses and Bowijk (1996) explained that observing a periodically integrated autoregressive process S times per year suggests there are S-1 cointegration relations within the annual series that include these seasonal observations, and these relations change with the seasons. This implies a single unit root in the vector autoregression of these annual series. They also determined that temporal aggregation does not affect the existence of this unit root. Therefore, applying a temporal aggregation filter to a periodically integrated series will yield another periodically integrated series.

As we will see later, the invariant property of the unit root tests (asymptotically) plays an important role in various aspects of the time-series analysis, such as Granger causality and cointegration (Mamingi, 1996; Rajaguru, 2004a; Rajaguru and Abeysinghe, 2008; Rajaguru *et al.*, 2018).

The literature referenced above presupposes that the time series is either stationary or non-stationary throughout the entire data span. However, they fail to analyse the effect of temporal aggregation or systematic sampling in cases where the time series exhibits piecewise stationarity. In such cases, the resultant process after applying the aggregation filter could be either stationary or non-stationary. We hypothesise that the resultant process is likely to be non-stationary, as the long run (non-stationarity) is likely to be preserved. If the aggregation scheme condenses the non-stationary component to a single observation, then it is likely to be treated as a stationary process where the condensed observation will appear as an outlier in a stationary process, and vice versa.

To analyse the nature of the unit root properties under various aggregation schemes on the piecewise stationary process, the future research could derive the condition at which the process is (1) stationary, (2) non-stationary and (3) piecewise stationary.

2.2 Univariate models

Having established that unit roots, particularly integer unit roots, are unaffected by temporal aggregation and systematic sampling, this section now provides a summary of the literature on additional properties of univariate models. Silvestrini and Veredas (2005, 2008) and Mamingi (2017) are a selection of the few detailed surveys on the literature in this area. From this section onwards, this literature review draws upon their comprehensive reviews of the work that has been done within the temporal aggregation and systematic sampling framework.

2.2.1 ARIMA models. As unit roots determine the stationarity of a series, the results presented establishes that aggregation also does not alter the stationary or non-stationary characteristics of the series. However, time series structures are generally affected by aggregation.

One of the earliest studies to initiate the study of temporal aggregation on stationary model structures is the seminal work of Amemiya and Wu (1972). They theoretically demonstrated a crucial finding: if the original time series follows a pure autoregressive $AR(p)$ process, then the resultant temporally aggregated non-overlapping series will follow $ARMA(p, q^*)$. In this model, the roots of the polynomial correspond to the m^{th} powers of the roots from the original disaggregated series, where m represents the frequency of aggregation. This occurs regardless of the timespan of aggregation. In particular, they showed that the MA component which arises from aggregation is $q^* = \left[p + 1 - \frac{(p+1)}{m}\right]$, and proved that it is invertible as the roots do not lie on the unit circle.

In an insightful paper challenging researchers' decisions on the "natural" interval of observations, Brewer (1973) builds upon Amemiya and Wu's results and extends them to $ARMA(p, q)$ models. He shows that temporal aggregation transforms this into $ARMA(p, q^*)$. Specifically, Brewer demonstrates that while the AR component retains its form from the pure autoregressive case, the MA component undergoes a transformation, becoming $q^* = \left[p + 1 + \frac{(q-p-1)}{m}\right]$.

A substantial amount of work has also been undertaken in the non-stationary framework. As with the stationary cases, temporal aggregation can change the structure of an integrated moving average process. Working (1960), Wei (1978a, 1978b), Tiao (1972) and Stram and Wei (1986) were some of the pioneers in this area. Collectively, the authors have theoretically established some limiting results. Working (1960) demonstrated substantial bias in developing a spurious correlation where such correlation is not present when the correlation is computed on the first differences of averages in a random chain. Further, under temporal aggregation, an $IMA(d, q)$ process is reduced to an $IMA(d, d)$ process (Stram and Wei, 1986), whereas systematic sampling turns it into an $IMA(d, d-1)$ process (Wei, 1978a). Rossana and Seater (1995) further elaborated that the limiting process could also be an $IMA(d, d-1)$ if the rise in standard errors, resulting from fewer observations, surpasses the increase in estimated autocorrelation coefficients—a scenario more likely in small samples. There is a notable implication from these results, which is that temporally aggregating data has the possibility of converting an MA process into white noise, specifically if the process is stationary ($d = 0$) (Wei, 1981).

The evolution of research in this area typically advances towards ARIMA type models. Tiao's 1972 research demonstrated that an $IMA(d, q)$ model evolving into an $IMA(d, d)$ remains applicable even when the original model includes an AR component. Thus, an $ARIMA(p, d, q)$ model also effectively transforms into an $IMA(d, d)$. Wei's (1978a) derivations also lead to the same conclusion. This strikes as remarkable, because it indicates that in the limit, neither the AR and MA lags (p and q) nor the ARIMA coefficients are relevant in determining the behaviour of a temporally aggregated time series. The order of integration d solely determines everything.

Another extension in this area is to introduce seasonality into the model. The general multiplicative seasonal ARIMA model can be represented by $ARIMA(p, d, q) \times (P, D, Q)_s$. Wei (1978a) and Weiss (1984) achieved the same result by constructing the m-th order non-overlapping aggregate series, which follows an $ARIMA(p, d, q^*) \times (P, D, Q)_S$ process, where $q^* = \left[p + d + 1 + \frac{(q-p-d-1)}{m}\right]$ and $S = \frac{s}{m}$. They have shown that when the aggregation frequency aligns with the seasonal frequency (for example, intra-annual seasonality with annual aggregation), the aggregate model simplifies to a model without seasonality. Thus, the limiting model becomes an $IMA(D + d, D + d)$ process.

For both non-seasonal and seasonal ARIMA models, Wei (1978b) reiterates the possibility that temporal aggregation can change the model structure to the extent of reducing it to a white noise process. This is especially true for the cases where m is large.

Across all the models reviewed, a consistent finding is that the order of integration (d) remains unchanged by temporal aggregation or systematic sampling. This outcome is intuitive, aligning with the findings related to unit roots and stationarity.

Stram and Wei (1986) question these previous conclusions by pointing out that the derived orders for the AR and MA polynomials represent only the maximum possible values and proceed to determine the exact order for the aggregated model. They explored the relationship between the autocovariance functions of the basic disaggregated series and its aggregated counterpart. Their findings reveal that under certain conditions, the AR order of mixed-ARIMA models may shrink due to temporal aggregation. This is related to the “hidden periodicity” of the order m. Rajaguru (2004a, b) further extended it to multivariate case in the context of cross-correlations. Stram and Wei’s (1986) result is the special case of Rajaguru (2004a, b) when the cross-correlations are zero.

Across two detailed theoretical papers, Teles and Sousa (2017) derived the relationships between the parameters of the aggregate and basic disaggregated ARMA models. This extended the results of Ahsanullah and Wei (1984), which derived these relationships for the $AR(1)$ series. Teles and Sousa used these derivations to measure the estimation accuracy due to temporal aggregation. They found that the estimation errors for the temporally aggregated series is very large (in absolute values), and can be exceedingly high, highlighting a substantial negative effect of temporal aggregation. This is evident even at the lowest order of aggregation, m = 2, and the impact intensifies with the order of the ARMA model and the corresponding increase in the number of parameters. Teles and Sousa (2017) cautioned that this low estimation accuracy can lead to severe adverse effects on the use of ARMA models, such as for inference and forecasting.

Interestingly, more recent developments by Wei have yielded further results in more specific cases that conflict with his previous works, underscoring the importance of handling data aggregation with careful consideration. In Lee and Wei (2017), the effect of temporal aggregation on testing for a mean change of time series through a likelihood ratio test was investigated. As opposed to the general belief that aggregation causes information loss, they found that the information of the likelihood ratio test for a mean change is strengthened (in terms of its empirical power) with the m th order temporal aggregation. The authors have also investigated the effects of temporal aggregation on modelling and testing a variance change in Lee and Wei (2023). They showed that the aggregation effect on the variance-change ARMA model can only be explained with non-aggregate model parameters.

2.2.2 GARCH models. While ARIMA models are used for modelling and forecasting the first moment, the generalised autoregressive conditional heteroskedasticity (GARCH) model is designed for modelling the second moment. This is known to be important particularly for financial time series which often exhibit conditional heteroskedasticity and changing volatilities. A crucial difference between the aggregation of ARIMA-type models and GARCH models is highlighted in Hafner (2008). When it comes to aggregation, it is always the first-order moments (i.e. returns) that is aggregated, and not the second-order process (i.e. squared returns). However, a GARCH process models the latter. Hafner notes that this results in the generation of cross-products, which then induces additional noise in the aggregated process. This area has been studied by Nelson (1990), Meddahi and Renault (2004), Hafner and Rombouts (2007) and Hafner (2009), amongst others.

Drost and Nijman (1993) thoroughly investigated the effects of time aggregation on GARCH processes, and they show that the classical GARCH assumptions are not robust to the sampling interval specification. They present variations of ARCH, GARCH and ARMA-GARCH models and the ramifications of aggregation on them when they are stock or flow variables. One of the results they find is that the class of weak (defined as models which only projections of the conditional variance are considered) GARCH models are also closed under aggregation for both stock and flow variables. Precisely, an aggregated symmetric weak $GARCH(p,q)$ model becomes $GARCH\left(r+\left[\frac{p-r}{m}\right],r\right)$ for stock variables, and $GARCH(r,r)$ for flow variables, with $r=\max(p,q)$. This extends to their main result, which states that the class of ARMA models with weak GARCH errors remains closed under temporal aggregation. They proved that temporal aggregation of an $ARMA(P,Q)-GARCH(p,q)$ model turns into an $ARMA(P,Q^*)-GARCH(\widetilde{r},\widetilde{r})$ where $Q^*=P+\left[\frac{Q-P+w}{m}\right]$ and $\widetilde{r}=r+\frac{1}{2}Q^*(Q^*+1)$.

Hafner (2008) expanded these findings to include the multivariate context, showing that the class of weak multivariate GARCH processes remains stable under temporal aggregation. Further, Drost and Nijman (1993) show that the disaggregated model yields estimate of the aggregated variance parameters which are more precise than direct estimates from the aggregated model.

2.2.3 Fractionally integrated models. The autoregressive fractionally integrated moving average (ARFIMA) model, or long memory processes, is said to be more flexible than ARIMA models as it allows orders of integration to be fractional, i.e. $0 \leq d \leq 1$. Chambers (1998) demonstrated that temporal aggregation preserves the order of integration, including fractional orders, of the underlying series. He further showed that the decay rate of autocorrelation functions in discrete-time long memory processes at large lags remains constant, regardless of temporal aggregation or systematic sampling. Therefore, the true long memory parameter, d, can be accurately estimated regardless of the sampling interval.

Hwang (2000) also investigates the impact of time aggregation and systematic sampling on discrete-time long memory processes, as well as their finite sample properties. This study confirmed Chambers' (1998) second result but extends it to show that temporal aggregation and systematic sampling significantly affect autocorrelation functions in short lags. Hwang identified the directional change in the autocorrelation functions due to aggregation filters, which indicates a bias. In particular, the level of autocorrelation functions shifts upwards and downwards when temporal aggregation and systematic sampling filters, respectively, are applied. More specifically, for the $ARFIMA(0, d, 0)$ process, the absolute value of the long memory parameter ($|d|$) is determined to be larger for a temporally aggregated process and smaller for a systematically sampled process, whereas the true parameter lies between them.

Beran and Ocker (2000) explored the effects of temporal aggregation on time series characterised by long memory, short memory, and anti-persistence within the context of fractional autoregressive processes. They extend Tiao's (1972) results to ARFIMA models as follows. Their results show that the long memory and anti-persistence are asymptotically invariant, but the short memory components of the model gradually disappear with an increasing *m*.

Additionally, Ohanissian (2001) builds upon the already established results that the true long memory process retains the same degree of long memory upon temporal aggregation. He shows that spurious long memory processes have long memory parameters that indeed diverge to either 0 or 1 upon temporal aggregation. Thus, he proposed a methodology to examine the nature of the long memory parameter estimate across different levels of aggregation, enabling the distinction between true and spurious long memory.

2.3 Multivariate models

Given their impact on univariate models, it is expected that temporal aggregation and systematic sampling also significantly affects multivariate models.

2.3.1 Dynamic relationships. The effect of temporal aggregation and systematic sampling on the dynamic relationship between variables has been examined by Telser (1967), Sims (1971) and Wei and Metha (1980), amongst others.

In an early paper studying the relationship between a basic linear dynamic difference equation model and a corresponding aggregate model, Tiao and Wei (1976) showed that if the basic model has a one-sided dynamic relationship, aggregation causes it to become two-sided, leading to a feedback mechanism. They find that identifying the fundamental dynamic structure becomes increasingly difficult as data aggregation increases, causing the variables to become coincidental. Therefore, obtaining disaggregated data is desirable, as the information loss due to aggregation for estimation is substantial.

However, in its original disaggregated form, the series may appear in a random frequency whereby modelling becomes merely impossible. This dissertation aims to address this issue in detail.

Wei (1978b) studied the effect of temporal aggregation on parameter estimation specifically in a general finite autoregressive distributed lag model. Despite deriving unbiased and consistent estimators, he finds that there is a loss of information in estimation due to aggregation. Wei explains that temporal aggregation generally induces correlations. Thus, the information loss due to aggregation is through the worsening multicollinearity among the input variables, on top of the inflation of error variance. He also shows that the recovering the parameters of the basic disaggregated model from the aggregated process becomes more difficult as the aggregation level increases. Wei and Mehta (1980) expanded on these findings, demonstrating that aggregation leads to a significant information loss particularly when the input series exhibits negative autocorrelation. These results were further reinforced by Wei (1990), who reiterated that higher levels of aggregation and an increased number of model parameters result in more substantial information loss during estimation.

2.3.2 Vector ARIMA models. A central theme from the results of the studies in Section 2.2 is that the structures and orders of the class of univariate ARIMA models do not remain unaffected. This issue extends into multivariate ARIMA models as well.

Lütkepohl (1987) thoroughly studies the effects of temporal aggregation on Vector Autoregressive moving Average (VARM) models. He examined the effect of temporal aggregation on covariance stationary VARMA models and found that the aggregating the n-variate $VARMA(p,q)$ model transforms it into a $VARMA(pn, pn+q)$ model. This study notes that the lag lengths pn and $pn+q$ are the upper bounds of the lag length for both AR and MA components, respectively.

Expanding these results into a more parsimonious aggregated representation, Marcellino (1999) obtains a proper VARIMA representation which is not final-form and is more useful for evaluating the effects of temporal aggregation on specific properties. From $VARMA(p,q)$ model, the author shows that aggregation yields a $VARMA(p,q^*)$ model with

$$q^* = \begin{cases} p-1-k & \text{for } km < p-q \leq (k+1)m, \text{ where } k = 0,1,\ldots,p-1 \\ q & \text{for } p = q \\ p+k & \text{for } km \leq q-p < (k+1)m, \text{ where } k = 0,1,\ldots \end{cases}.$$

He also derives the results for a systematically sampled $VARMA(p,q)$. In this case, the model turns into $VARMA(p,q^*)$, with

$$q^* = \begin{cases} p-k & \text{for } km < p-q+1 \leq (k+1)m, \text{ where } k = 0,1,\ldots,p \\ q & \text{for } p = q \\ p+1+k & \text{for } km \leq q-1-p < (k+1)m, \text{ where } k = 0,1,\ldots \end{cases}.$$

Note that k is the lowest value such that the inequalities are satisfied. In line with the univariate results, Marcellino (1999) points out that the AR component is generally the same whether the VARMA process is temporally aggregated and systematically sampled. Nonetheless, the order and coefficients of the MA component for both are different. More specifically, the MA order after temporal aggregation is typically higher than that of systematic sampling.

2.3.3 Time series properties. The increased complexity of the structure of VARMA models upon temporal aggregation reverberates across several time series properties. This includes the trend-cycle decompositions of a series. Beveridge and Nelson (1981), Lippi and Reichlin (1991) and Pesaran *et al.* (1991), are some of the studies which have focused on this. Essentially, trend-cycle decompositions are affected by temporal aggregation. The aggregated model leads to distinct and sometimes conflicting trend and cycle results from that of the basic model.

Another property that has been explored is measures of the persistence of shocks, such as impulse response functions and variance decomposition, which have been studied by Swanson and Granger (2012). They show that there is a change in the instantaneous correlation of residuals which can be explained by temporal aggregation. Rosanna and Seater (1995) elaborate that the numerical values of the impulse response function and the variance ratio, as measures of persistence, converge to fixed values which are only dependent on the order of integration of each variable – irrespective of the AR and MA parameters of the underlying data generating process (DGP). They also warn that as the level of temporal aggregation increases, both measures become less reliable indicators of long-run persistence in the original data.

Besides this, the quality of forecasting is crucial, particularly in economic policymaking. A pioneering work in this specific area is Lütkepohl's (1987) comprehensive analysis of the effect of temporal aggregation on the efficiency of forecasts. There are two ways of forecasting aggregated values: forecasting directly from the aggregated model or obtaining the disaggregated forecasts and then aggregating them. Lütkepohl (1987) compares these two approaches for a multivariate ARMA model and concludes that the latter results in a smaller mean squared forecast error. Zellner and Montmarquette (1971), in their examination of the temporal aggregation problem within a general regression model, alluded to similar results prior to Lütkepohl's deep analysis; the predictors based on disaggregated data are more precise than those based on aggregated data. The authors also highlighted the challenge of making meaningful short-run forecasts with temporally aggregated data, as the smoothing effect of temporal aggregation emphasises the long-run trend.

Wei (1990) concurs that forecasting using an aggregated model tends to be less efficient compared to using a disaggregated model. The efficiency loss from aggregation can be substantial, especially when the non-seasonal component of the model is non-stationary. If the non-seasonal component is stationary, the impact on efficiency is less severe. Furthermore, he shows that there is no loss in forecasting efficiency when the underlying model is purely seasonal.

Marcellino (1999) cautions that using temporally aggregated data can diminish the finite-sample power of testing procedures, a consequence of having fewer observations available. This often leads to a higher mean squared forecast error as well. Supporting these findings, Jea *et al*. (2014) in their study on multi-period regression models, observed that the power of tests declines with increasing aggregation levels. They also found that higher levels of aggregation complicate the detection of true and significant relationships between variables.

In multivariate time series modelling, Granger causality tests have become popular among the practitioners to identify the causal directions and its policy implications. The next section will focus on the effect of aggregation on causal inferences.

2.4 Causality

2.4.1 Granger causality. A substantial body of theoretical literature examines the effects of temporal aggregation and systematic sampling on Granger causality, as well as on weak and strong exogeneity. In a seminal paper, Sims (1971) found that a spurious causal relationship can be attributed to time aggregation. Geweke (1982) constructed measures of linear dependence and feedback between two time series, x_t and y_t. He showed that $F_{x,y} = F_{x\to y} + F_{y\to x} + F_{x\cdot y}$, whereby $F_{x,y}$ is denoted as the linear relationship between x_t and y_t, $F_{x\to y}$ as the linear causality from x_t to y_t, $F_{y\to x}$ as the linear causality from y_t to x_t, and $F_{x\cdot y}$ as the instantaneous linear causality between x_t and y_t. Focusing more specifically on causal relationships, Wei (1982) extends the framework to investigate the effect of aggregation on causal dynamics. This approach builds on the formulation by Tiao and Wei (1976) and utilises Geweke's canonical form for analysing linear relationships. Wei (1982) demonstrated that temporal aggregation transforms a unidirectional causality from Y_t to X_t into a spurious two-sided feedback system. As the level of aggregation increases, these causal relationships weaken, and the information increasingly concentrates on false instantaneous causality between the variables, with all lag dependencies gradually disappearing. On the other hand, systematic sampling preserves the causal relationship direction, but this relationship is nonetheless weakened. This suggests that identifying the true causal relationship becomes more challenging even with systematic sampling, although it does not introduce spurious instantaneous causality.

With a Monte Carlo experiment on series which follow a stable AR process, Cunningham and Vilasuso (1995) have shown that compared to systematic sampling, temporal aggregation leads to approximately a tenfold increase in the chance of failing to detect a true causal relationship. Their results support the finding that longer sampling intervals under systematic sampling and temporal aggregation increases the possibility of incorrect analyses.

Silvestrini and Veredas (2008) explain the issue of spurious instantaneous causality by stating that it occurs when x_t and y_t do not have any Granger causal relationship at their natural frequency, yet instantaneous causality emerges in the aggregated variables. In an empirical study focused on the relationship between inflation and output growth from cross-country regressions, Ericsson *et al*. (2001) point out the impact temporal aggregation has on time series analyses such as theirs. They derived the limiting values of the estimated coefficients from a non-contemporaneous regression of a bivariate $VAR(1)$ system which has a true bidirectional causal relationship. Based on the temporal aggregation of two-period non-overlapping averages, they obtained the result that the instantaneous causality between the two variables could be positive, negative, or zero, even if the original disaggregated series do not have a contemporaneous correlation. Furthermore, the probability limit of the coefficient could be zero even when the actual coefficients are non-zero. Rajaguru (2004a, b) and Rajaguru and Abeysinghe (2012) also confirmed these results.

Although the analysis in Ericsson *et al*. (2001) is specifically set up for the authors' empirical application, they conjectured that these results are applicable to the general multiple lag multivariate VAR. This was formally proven by Rajaguru (2004a) and Rajaguru *et al.* (2018) for the non-cointegrated case, and then subsequently extended by Rajaguru and Abeysinghe (2008) for the cointegrated system. These authors took the results a step further and shows that the cross-covariances of the aggregated process can be written as a function of the cross-covariances of the disaggregated process for the general integrated process of order *d*.

Breitung and Swanson (2002) made another significant contribution to the literature on spurious instantaneous causality. A key issue in this area is determining whether the observed instantaneous causality between two variables reflects a genuine underlying causal relationship between the original variables x_t and y_t, or if it is merely an artifact of temporal aggregation. Using asymptotic theory for large aggregation intervals, Breitung and Swanson (2002) derived sufficient conditions to rule out spurious causality in VAR models caused by temporal aggregation. A Monte Carlo experiment utilising a $VAR(1)$ model confirms their findings that as the level of aggregation (m) increases, short-run dynamic structures disappear, and asymptotic results become applicable. Nonetheless, for persistent processes, the dynamic structures remain important at the lower level of aggregation (m). Breitung and Swanson (2002) derived the conditions under which the non-stationary properties are invariant to temporal aggregation (Granger (1990) and Marcellino (1999)). However, if the non-stationary properties are not valid then the validity of the Breitung and Swanson (2002) is questionable. The results by Granger (1990) are applicable when the order of integration 'd' (the source of non-stationarity) is integer. But the results are not generalisable when the variables are fractionally integrated.

Granger (1990) also found that when a pair of series exhibits a one-way causal relationship at the disaggregate level, temporal aggregation tends to transform it into a feedback or two-way causation. However, as the aggregation frequency increases, a stationary series becomes contemporaneously related, leading to a total loss of causality.

Breitung and Swanson (2002) showed that for a moderate sampling interval, however, temporally aggregated (stationary) variables are in fact well approximated by a VMA(1) process. Additionally, as the sampling interval increases, the short-run dynamics also disappear. On the other hand, Cunningham and Vilasuso (1995) concluded that at short aggregation intervals, temporal aggregation highly likely leads to a failure of detecting the true causal relationship.

Fan *et al*. (2024) explore if temporal aggregation preserves the consistency of functional causal models and conditional independence under general (nonlinear) conditions. They discovered that achieving consistency is challenging, and the model becomes unidentifiable after temporal aggregation. They highlighted the detrimental effects of temporal aggregation on instantaneous causal discovery, though a solution remains elusive.

It is evident that the study of aggregation especially in multivariate models does not, and indeed, cannot overlook the significance of Granger causality (which is crucial for understanding spurious instantaneous causality in the first place). Multiple papers previously referenced are included in the significant body of literature which examines the ramifications of systematic sampling and temporal aggregation on Granger causality. Marcellino (1999) cautioned that Granger causality is generally not preserved under aggregation. He noted that while it is often lost after temporal aggregation, it can also be spuriously generated.

Rajaguru (2004a) and Rajaguru and Abeysinghe (2012) investigated the impact of temporal aggregation on causal relationships among variables by establishing a connection between the autocovariances of the original disaggregated series and those of the aggregated series in a multivariate context. This extends the analysis of Stram and Wei (1986), who derived this result for the univariate case. Rajaguru and Abeysinghe (2012) thoroughly investigated this relationship under several cases. First, they find that the causally unrelated series remains unrelated due to temporal aggregation. Second, temporal aggregation creates a spurious feedback effect which is evidenced by the non-zero values of the probability limit of the estimated coefficient. They also found that this spurious feedback is highly likely in practice. Third, if the Granger causality between the disaggregated series is bi-directional, it is possible to conclude that causality is unidirectional for the aggregated series. At the higher level of aggregation, the last two cases tend to show causal relationships between the variables. This paper also included a Monte Carlo study demonstrating that aggregation distortions primarily occur at lower levels of aggregation, specifically when the aggregation order slightly surpasses the true causal lag. At higher levels of aggregation, only the contemporaneous correlation remains evident. Consequently, Rajaguru and Abeysinghe cautioned that standard Granger causality tests, which overlook contemporaneous correlations, should be applied with extreme caution. The absence of causal findings in temporally aggregated data does not inherently imply a lack of causality between variables.

The problems in this regard are further exacerbated as a lack of consensus persists, even within the stationary framework. Rajaguru (2004a) demonstrated that low levels of temporal aggregation are more prone to misleading inferences, suggesting that using more disaggregated data may not always be beneficial. This finding is significant for the current study. Rajaguru and Abeysinghe (2008) conducted a Monte Carlo study and demonstrate that analyses might be better off with highly temporally aggregated data especially in terms of causality testing. There does not seem to be a promising outcome on Granger causality inference when disaggregated data is employed.

In summary, earlier studies using stationary VAR models have found that temporal aggregation distorts causal linkages and leads to spurious causal inferences, whereas systematic sampling does not result in any spurious causation. They highly recommend using the systematic sampling filter instead of the temporal aggregation filter to establish non-spurious causal inferences. Rajaguru (2004b) affirms that systematic sampling maintains unidirectional causality in stationary processes. However, he established through the cross-covariance analysis that in the presence of unit roots, systematic sampling transforms a one-sided causality into a bidirectional causal relationship. This contrasts with stationary VAR models, where systematic sampling preserves causal direction. Rajaguru (2004b) further demonstrated that when non-stationary variables are differenced and used in a VAR framework, systematic sampling can lead to spurious causal inferences. Consequently, it becomes impossible to recover the true causal linkages using aggregated data, particularly when the unidirectional causality originates from a non-stationary series and is directed towards either a stationary or another non-stationary series. As in temporal aggregation filter, systematic sampling also creates contemporaneous causal link when the series possess unit roots.

On a positive note, Rajaguru *et al.* (2018) derived the condition at which the standard Granger Causality tests do not produce spurious results. They showed that if the unidirectional Granger causality originates from a stationary stock variable (where the systematic sampling filter is used) to any other variable (stock, flow, stationary or non-stationary), then one can attain a non-spurious result. All other cases generate spurious causal inferences. This latter study indicates that systematic sampling may have a similar distortionary effect as temporal aggregation when the order of integration is $d > 1$. They also demonstrated that systematic sampling causes all causal information to focus within the contemporaneous relationships among variables. These spurious contemporaneous relationships persist, even when the level of aggregation increases.

The studies above deal with the short-run causal relationship between the variables within a stationary VAR model. If the series are non-stationary, then the VAR model is constructed in a differenced form to examine the underlying causal directions between the variables. But if the integrated variables are cointegrated then one could analyse the nature of the long-run causal relationship in the context of aggregation, which is outlined in the next section.

2.4.2 Cointegration. After establishing that the structure of VARMA models becomes more complicated upon temporal aggregation, Marcellino (1999) also demonstrated that temporal aggregation alters most properties presented at the basic disaggregated level, except long-run properties such as cointegrating vectors (see Rajaguru and Abeysinghe, 2008). As was hinted earlier in this literature review, the theory that cointegration is invariant to temporal aggregation was originally established by Granger (1990). He showed that the $I(d)$ linear combinations remain as $I(d)$ after temporal aggregation, implying that cointegration would similarly remain invariant to aggregation. However, a more recent finding from Tserkezos (2021) has warned that the power of the Augmented Dickey Fuller (ADF) stationarity test is negatively impacted by temporal aggregation. This indicates that the ADF test has the potential to produce misleading results about the stationarity of a temporally aggregated series, particularly as the aggregation interval increases.

Similarly, with cointegration tests, Otero *et al.* (2022) found that the ability to detect the presence of long-term cointegration relationships can decrease if the sample length is short. Through Monte Carlo simulations, the data span of a sample is shown to take precedence over the frequency of observations for the power of a Johansen cointegration test. Marcellino (1999) extended these results, demonstrating that if the vector z is cointegrated of order $CI(1, 0)$, both the number and the composition of the cointegrating vectors remain unchanged upon temporal aggregation. The $I(2)$ case was proven by Marcellino (1996), which indicates that these results can be used for the general $I(d)$ case.

An extension of these cointegration results is found in Rajaguru and Abeysinghe (2008) which demonstrated that with increasing levels of temporal aggregation, a stationary VAR(p) process might approach a $VAR(0)$ model, as all causal information becomes concentrated in contemporaneous links. However, due to the presence of unit roots, a cointegrated $VAR(p)$ process can reduce to $VAR(1)$ and it will not further converge to VAR(0). This persistence of cointegrating relationships despite temporal aggregation offers a promising avenue for Granger causality testing (Rajaguru and Abeysinghe, 2008).

In light of the role of cointegration in this context, Kirchgässner and Wolters (1992) studied the causal relationship between both systematically sampled and temporally aggregated variables. They showed that if the variables are cointegrated, temporal aggregation can lead to spurious instantaneous relations and feedback causal relations. Crucially, they found that if the explanatory variable is temporally aggregated while the dependent variable systematically sampled, spurious feedback relationships which cannot be neglected can occur. This result generally holds whether the variables are cointegrated or not.

In a more detailed setting, Mamingi (1996) used a Monte Carlo experiment to examine how temporal aggregation distorts causality in error correction models (ECMs). As cointegration has been established to be invariant to aggregation, it is not surprising that Mamingi's results are in line with the finding that temporal aggregation leads to a greater Granger causality distortion than systematic sampling does. The latter typically maintains the structure of the ECM and the true causal relationship—causal distortions appear to be absent, especially for variables that exhibit a higher degree of cointegration. Temporal aggregation, however, significantly alters the true causal relationship and the feedback dominates. He also demonstrates that a large data span, rather than a large sample size, is more detrimental to accurately detecting true Granger causality between variables, particularly affecting the power of cointegration tests.

Although these results from Mamingi (1996) are based on ECMs, it is conflicting to identify that systematic sampling does not produce spurious causal relationship in the presence of unit roots when they are cointegrated. Thus, Rajaguru and Abeysinghe (2008) provided analytical results for these findings. Upon analysing the relationship between parameter estimates and t-statistics of aggregated processes compared to parameters of non-aggregated processes, Rajaguru and Abeysinghe (2008) revealed that any empirically observed cointegrating relationship displaying the incorrect sign is likely due to temporal aggregation. To correct this distortion in the sign of the adjustment coefficient in an ECM, they introduced a sign rule based on the adjustment coefficients from the estimated cointegrating vector. This adjustment aids in causal inference and contemporaneous conditioning in regression models. However, the nature of causal relationships under various aggregation filters remains unclear when multiple cointegrating vectors are involved.

McCrorie and Chambers (2006) explored solutions to mitigate the distortional effects of temporal aggregation bias and suggested developing econometric models in continuous time. They argued that continuous time models offer advantages for causality testing because they allow the integration of *a priori* information into the data (via a causal chain model) independently of the sampling rate, without impacting Granger causality relationships. They theoretically demonstrated that models formulated in continuous time provide a framework to correct for temporal aggregation effects in observed discrete data, via a discrete time analogue, in a manner that does not depend on assuming a specific time unit in the data generation process.

The above-mentioned literature analyses the effect of various aggregation filters on causal inferences using synchronous time series data. Even for financial time series, which arrive at random frequencies, they are constructed into equally spaced time series. For instance, the McCrorie and Chambers (2006) technique requires converting a non-synchronous discrete time series into a synchronous time series and subsequently choosing an appropriate functional form. The choice of functional form can alter the nature of the causal relationship. Additionally, a solution to navigate the causal distortion of aggregation was proposed by Rajaguru and Abeysinghe (2008) through the sign rule, and this is applicable to synchronous time series data. However, the effect of aggregation on causal inference with non-synchronous data remains largely unexplored and unknown.

2.5 Aggregate or disaggregate

From the existing literature, an important question for time series analyses arises: should data be aggregated or disaggregated for the Granger causality tests? The literature review above has established the case for aggregating the data for the cointegrated case with one cointegrating vector to establish non-spurious causal inference. However, the nature of the causal relations due to various aggregation filters when there are more than one cointegrating vectors remains unknown. It is also identified from the previous section that inference based on non-cointegrated system produces spurious causal inference. Aggregation leads to many problems, and even these problems are conflicting in the research, especially when the data set becomes more complicated, i.e. non-synchronous data. So, the idea of working with the aggregated data is ruled out.

Some techniques have been explored in terms of disaggregating the data, beginning with Chow and Lin (1971) [1]. In this paper, they studied the problems of interpolation, extrapolation, and distribution, which are essentially the disaggregation counterparts of systematic sampling, temporal aggregation and mixed sampling respectively. In all these cases, they assume that the series to be disaggregated is available in a lower frequency (e.g. quarterly observations), while a related series is available in a higher frequency (e.g. monthly observations). Chow and Lin derived an unbiased estimator which is applicable to all three problems. However, there are two main deterrents of using this estimator, and the authors have addressed one of them. They acknowledged that the usefulness of their method is dependent on the availability of a related series, which is not always feasible.

Yet, the more profound challenge in the disaggregation process lies in determining the extent to which the data should be disaggregated. Even if it is assumed that the related series is available at any frequency, the optimal level of disaggregation remains unclear. Conversely, if there is a clear answer regarding the appropriate frequency for disaggregation, the initial issue arises yet again, in that the related series may not always be available at the required frequency.

The impact of aggregation on causal analyses have been established by the existing literature. As for disaggregation, although some methods exist, they have not been studied as extensively. One possible extension to the literature could be identify the optimal level of disaggregation (if any) to identify the robust causal relationship between the variables. If the optimal level of disaggregation doesn't exist then the proposed solution is to analyse the data as it is, regardless of its synchronicity.

2.6 Causal inference with non-synchronous time series data

One of the most important pioneers of research exploring the framework of non-synchronous time series data is Hayashi and Yoshida (2005). They conceived a new cross-correlation estimator which does not require any synchronisation of the data, whether they are recorded at regularly spaced intervals or not. The Hayashi-Yoshida estimator thus utilises all available tick-by-tick data, which in many financial contexts, are inherently asynchronous. This has been foundational to the seminal study of lead/lag relationships conducted by Hoffmann *et al.* (2013) and Huth and Abergel (2014), particularly when dealing with high frequency data. The Hoffmann *et al.* (2013) study constructs an estimator of the lead/lag parameter, which represents the time shift for which the leader anticipates the lagger. On the other hand, Huth and Abergel (2014) measures the lead/lag relationships between two assets using the lead/lag ratio (LLR), which is based on the asymmetry of the cross-correlation function. This involves calculating the ratio of the sum of squared correlations across all positive lags to the sum of squared correlations across all negative lags. For example, if the sum of squared correlations for all lags of variable X against Y exceeds that of the reverse, then the LLR (Lag Lead Ratio) exceeds 1, indicating that X leads Y and X Granger causes Y, as noted by Huth and Abergel (2014) and O'Neill and Rajaguru (2020, 2023).

The LLR is a common method used to capture causality between two variables, as it shows which variable leads or lags the other. Especially when trading occurs at a high frequency (i.e. intraday), the LLR is becoming more favourable in the analysis of the flow of information and understanding financial market dynamics because of its ability to bypass any preprocessing of the data (Bollen *et al.*, 2017). However, there are several limitations to this method. The first is that it can only directly analyse a bivariate relationship. Secondly, O'Neill and Rajaguru (2020) have pointed out that the LLR method only captures unidirectional causality – the causal relationship either runs from X to Y or from Y to X. If the true underlying relationship either possesses a bidirectional causality or does not have a causal relationship at all, the arbitrary LLR benchmark of 1 can be misleading as it is not capable of accurately determining these cases.

Although the theories established in Bibinger and Mykland (2013) resulted in confidence intervals for the Hayashi-Yoshida covariance matrix, Bollen *et al.* (2017) identified a gap in the literature concerning the significance of lead/lag correlations – the third problem with the LLR. They also noted the lack of clarity in translating the covariance matrix confidence intervals into confidence intervals for cross-correlation functions and the LLR. In their paper which aimed to uncover the lead/lag relation between different VIX-related time series, Bollen, O'Neill, and Whaley attempted to circumvent this issue by bootstrapping the Hayashi–Yoshida covariance matrix under the null using the empirical distribution. To preserve the correlation structure, they performed the resampling in blocks.

Unfortunately, the analysis based on bootstrap confidence intervals assuming a two-tailed alternative can still be misleading, because it identifies the presence of causality rather than establishing the direction of causality. Thus, O'Neill and Rajaguru (2020) proposed a response surface regression-based technique through extensive simulations to improve the statistical power of the LLR tests. Using random simulations, they generated the upper-tail critical values for the LLR. The results of the response surface technique are consistent with the Bollen *et al.* (2017) bootstrap results. O'Neill and Rajaguru also proved that the simulated values are more efficient and convenient. This simulation method offers a reliable solution to the statistical significance problem of the LLR measure, but the results remain applicable only to a bivariate system. Thus, the core issue with the LLR continues to persist. Using the LLR to analyse lead/lag relationships can be misleading and produce biased inferences, if the underlying DGP involves more than two variables.

2.7 Practical implications and further research

It is well established that aggregation preserves causal inferences in the absence of cointegration, provided Granger causality originates from a stationary stock variable. However, when variables are cointegrated, the sign-rule proposed by Rajaguru and Abeysinghe (2008) can be applied to determine the underlying causal inferences. According to their findings, the sign of the error correction coefficients (speed of adjustment coefficients) must be opposite to the sign of the coefficient in the cointegrating vector. If the signs are the same and the coefficients are significant, this indicates that the identified causality is spurious. In such cases, it can be concluded that there is no causality at the disaggregate level. Moreover, it is recommended to conduct the Granger causality using a sign rule in the presence of cointegration with systematically sampled data instead of temporally aggregated data regardless of the nature of the data (flow or stock).

The future research could analyse the nature of the causal relationship between variables when they appear as a mixed sample (i.e. one of them is flow and the other is a stock variable). The nature of causal distortions within a mixed-frequency model remains unclear, especially when one variable is measured quarterly and another monthly. Despite these complications, the sign-rule may still be applicable in this context.

3. Conclusion

This comprehensive review on temporal aggregation and systematic sampling has explained their significant influence on time series analysis, particularly in fields such as economics, and finance. Temporal aggregation, while simplifying data analysis by emphasising longer-term trends, may obscure crucial short-term fluctuations and introduce biases that potentially mislead conclusions. On the other hand, systematic sampling enhances computational efficiency but risks losing information crucial for accurate data representation, especially in non-stationary series.

The review highlights that while both methods offer practical benefits for handling large datasets, they come with trade-offs that must be carefully managed. For instance, higher aggregation levels can significantly alter the dynamics and statistical properties of the data, potentially leading to incorrect inferences about the underlying processes.

Moreover, this review stresses the necessity of context-specific approaches when employing these techniques. The choice of method and its parameters should be tailored to the specific characteristics of the data and the objectives of the analysis. By adhering to the best practices and recommendations discussed, researchers and analysts can better navigate the complexities introduced by these data reduction techniques, thereby enhancing the reliability and accuracy of their time-series analyses.

In conclusion, while temporal aggregation and systematic sampling are invaluable tools in time-series analysis, their application requires an understanding of their effects on data integrity and analysis outcomes. Future research should continue to explore advanced methodologies that mitigate the inherent trade-offs of these techniques, aiming to optimise both data reduction and analytical precision. This ongoing refinement will be crucial as data continues to grow in volume and complexity across various scientific and applied fields.

The review carefully analysed the cases where the causal inferences are preserved after applying various aggregation filters. Temporal aggregation always led to misleading causal inference while the causal inferences are preserved under systematic sampling the true causality runs from a stationary time series to either a stationary or a non-stationary time series. However, in many instances the causal inferences are made on the cases where one of the variables is flow the other variable stock. In such cases, one should apply a mixed sampling filter where temporal aggregation is applied to a flow variable and the systematic sampling is applied to a stock variable. The effect of mixed sampling on causal inference is still unknown and the future studies can theoretically analyse them.

It is also established from the review that neither aggregation nor a disaggregation is a solution to establish the causal inferences. The solution is to model the series as it is where the observation arrives at random frequency. There is clear potential for future studies to develop tools that can establish a causal inference based on the non-synchronous time series data.

Notes

1. Also see Abeysinghe and Rajaguru (2004), Banbura and Modugno (2014), Denton (1971), Fernandez (1981), and Litterman (1983).

References

Abeysinghe, T. and Rajaguru, G. (2004), “Quarterly real GDP estimates for China and ASEAN4 with a forecast evaluation”, *Journal of Forecasting*, Vol. 23 No. 6, pp. 431-447, doi: 10.1002/for.922.

Ahsanullah, M. and Wei, W.W.S. (1984), “The effects of time aggregation on the AR(1) process”, *Computational Statistics*, Vol. 4, pp. 343-352.

Amemiya, T. and Wu, R.Y. (1972), “The effect of aggregation on prediction in the autoregressive model”, *Journal of the American Statistical Association*, Vol. 67 No. 339, pp. 628-632, doi: 10.1080/01621459.1972.10481264.

Bańbura, M. and Modugno, M. (2014), “Maximum likelihood estimation of factor models on datasets with arbitrary pattern of missing data”, *Journal of Applied Econometrics (Chichester, England)*, Vol. 29 No. 1, pp. 133-160, doi: 10.1002/jae.2306.

Beran, J. and Ocker, D. (2000), “Temporal aggregation of stationary and nonstationary FARIMA (p, d, 0) models”, IDEAS Working Paper Series from RePEc.

Beveridge, S. and Nelson, C.R. (1981), “A new approach to decomposition of economic time series into permanent and transitory components with particular attention to measurement of the ‘business cycle’”, *Journal of Monetary Economics*, Vol. 7 No. 2, pp. 151-174, doi: 10.1016/0304-3932(81)90040-4.

Bibinger, M. and Mykland, P.A. (2013), “Inference for multi-dimensional high-frequency data: equivalence of methods, central limit theorems, and an application to conditional independence testing”, arXiv, available at: https://arxiv.org/pdf/1301.2074

Bollen, N.P.B., O’Neill, M.J. and Whaley, R.E. (2017), “Tail wags dog: intraday price discovery in VIX markets”, *Journal of Futures Markets*, Vol. 37 No. 5, pp. 431-451, doi: 10.1002/fut.21805.

Breitung, J. and Swanson, N.R. (2002), "Temporal aggregation and spurious instantaneous causality in multiple time series models", *Journal of Time Series Analysis*, Vol. 23 No. 6, pp. 651-665, doi: 10.1111/1467-9892.00284.

Brewer, K.R.W. (1973), "Some consequences of temporal aggregation and systematic sampling for ARMA and ARMAX models", *Journal of Econometrics*, Vol. 1 No. 2, pp. 133-154, doi: 10.1016/0304-4076(73)90015-8.

Chambers, M.J. (1998), "Long memory and aggregation in macroeconomic time series", *International Economic Review (Philadelphia)*, Vol. 39 No. 4, pp. 1053-1072, doi: 10.2307/2527352.

Chow, G.C. and Lin, A. (1971), "Best linear unbiased interpolation, distribution, and extrapolation of time series by related series", *The Review of Economics and Statistics*, Vol. 53 No. 4, pp. 372-375, doi: 10.2307/1928739.

Cunningham, S. and Vilasuso, J. (1995), "Time aggregation and causality tests: results from a Monte Carlo experiment", *Applied Economics Letters*, Vol. 2 No. 10, pp. 403-405, doi: 10.1080/758519001.

Denton, F.T. (1971), "Adjustment of monthly or quarterly series to annual totals: an approach based on quadratic minimization", *Journal of the American Statistical Association*, Vol. 66 No. 333, pp. 99-102, doi: 10.1080/01621459.1971.10482227.

Drost, F.C. and Nijman, T.E. (1993), "Temporal aggregation of GARCH processes", *Econometrica*, Vol. 61 No. 4, pp. 909-927, doi: 10.2307/2951767.

Ericsson, N.R., Irons, J.S. and Tryon, R.W. (2001), "Output and inflation in the long run", *Journal of Applied Econometrics*, Vol. 16 No. 3, pp. 241-253, doi: 10.1002/jae.614.

Fan, S., Gong, M. and Zhang, K. (2024), "On the recoverability of causal relations from temporally aggregated I.I.D. Data", arXiv:2406.02191 [stat.ML], doi: 10.48550/arxiv.2406.02191.

Fernandez, R. (1981), "A methodological note on the estimation of time series", *The Review of Economics and Statistics*, Vol. 63 No. 3, pp. 471-476, doi: 10.2307/1924371.

Franses, P.H. and Boswijk, H.P. (1996), "Temporal aggregation in a periodically integrated autoregressive process", *Statistics and Probability Letters*, Vol. 30 No. 3, pp. 235-240, doi: 10.1016/0167-7152(95)00225-1.

Geweke, J. (1982), "Measurement of linear dependence and feedback between multiple time series", *Journal of the American Statistical Association*, Vol. 77 No. 378, pp. 304-313, doi: 10.1080/01621459.1982.10477803.

Granger, C.W.J. (1990), "Aggregation of time-series variables: a survey", in Barker, T. and Pesaran, M.H. (Eds), *Disaggregation in Econometric Modelling*, Routledge, London, pp. 17-34.

Granger, C.W.J. and Siklos, P.L. (1995), "Systematic sampling, temporal aggregation, seasonal adjustment, and cointegration theory and evidence", *Journal of Econometrics*, Vol. 66 No. 1, pp. 357-369, doi: 10.1016/0304-4076(94)01622-7.

Hafner, C.M. (2008), "Temporal aggregation of multivariate GARCH processes", *Journal of Econometrics*, Vol. 142 No. 1, pp. 467-483, doi: 10.1016/j.jeconom.2007.08.001.

Hafner, C.M. (2009), "Causality and forecasting in temporally aggregated multivariate GARCH processes", *The Econometrics Journal*, Vol. 12 No. 1, pp. 127-146, doi: 10.1111/j.1368-423X.2008.00276.x.

Hafner, C.M. and Rombouts, J.V.K. (2007), "Estimation of temporally aggregated multivariate GARCH models", *Journal of Statistical Computation and Simulation*, Vol. 77 No. 8, pp. 629-650, doi: 10.1080/10629360600616252.

Hayashi, T. and Yoshida, N. (2005), "On covariance estimation of non-synchronously observed diffusion processes", *Bernoulli*, Vol. 11 No. 2, pp. 359-379, doi: 10.3150/bj/1116340299.

Hoffmann, M., Rosenbaum, M. and Yoshida, N. (2013), "Estimation of the lead-lag parameter from non-synchronous data", *Bernoulli*, Vol. 19 No. 2, pp. 426-461, doi: 10.3150/11-bej407.

Huth, N. and Abergel, F. (2014), "High frequency lead/lag relationships – empirical facts", *Journal of Empirical Finance*, Vol. 26, pp. 41-58, doi: 10.1016/j.jempfin.2014.01.003.

Hwang, S. (2000), "The effects of systematic sampling and temporal aggregation on discrete time long memory processes and their finite sample properties", *Econometric Theory*, Vol. 16 No. 3, pp. 347-372, doi: 10.1017/S0266466600163030.

Hylleberg, S., Jorgansen, C. and Sorensen, N.K. (1993), "Seasonality in macroeconomic time series", *Empirical Economics*, Vol. 18 No. 2, pp. 321-335, doi: 10.1007/bf01205406.

Jea, R., Lin, L., Shih, T.S. and Lee, H.M. (2014), "Time aggregation effect on multi-period regression models", *ICIC Express Letters*, Vol. 8 No. 5, pp. 1505-1510.

Kirchgässner, G. and Wolters, J. (1992), "Implications of temporal aggregation on the relation between two time series", *Statistical Papers*, Vol. 33, pp. 1-19, doi: 10.1007/bf02925307.

Lee, B.H. and Wei, W.W.S. (2017), "The use of temporally aggregated data on detecting a mean change of a time series process", *Communications in Statistics - Theory and Methods*, Vol. 46 No. 12, pp. 5851-5871, doi: 10.1080/03610926.2015.1091082.

Lee, B.H. and Wei, W.W.S. (2023), "The use of temporally aggregated data in modeling and testing a variance change in a time series", *Communications in Statistics - Simulation and Computation*, Vol. 52 No. 7, pp. 3183-3200, doi: 10.1080/03610918.2021.1928197.

Lippi, M. and Reichlin, L. (1991), "Trend-cycle decompositions and measures of persistence: does time aggregation matter?", *The Economic Journal (London)*, Vol. 101 No. 405, pp. 314-323, doi: 10.2307/2233821.

Litterman, R.B. (1983), "A random walk, markov model for the distribution of time series", *Journal of Business and Economic Statistics*, Vol. 1 No. 2, pp. 169-173, doi: 10.1080/07350015.1983.10509336.

Lütkepohl, H. (1987), *Forecasting Aggregated Vector ARMA Processes*, Springer-Verlag, Berlin, Heidelberg, Vol. 284.

Mamingi, N. (1996), "Aggregation over time, error correction models and Granger causality: a Monte Carlo investigation", *Economics Letters*, Vol. 52 No. 1, pp. 7-14, doi: 10.1016/0165-1765(96)00841-5.

Mamingi, N. (2017), "Beauty and ugliness of aggregation over time: a survey", *Jahrbuch für Wirtschaftswissenschaften*, Vol. 68 No. 3, pp. 205-227, doi: 10.1515/roe-2017-0027.

Marcellino, M. (1996), "Aggregation of an I(2) process (in Italian with English summary)", *Proceedings of the XXXVIII Meeting of the Italian Statistical Society*, Vol. 2, pp. 549-556.

Marcellino, M. (1999), "Some consequences of temporal aggregation in empirical analysis", *Journal of Business and Economic Statistics*, Vol. 17 No. 1, pp. 129-136, doi: 10.2307/1392244.

McCrorie, J.R. and Chambers, M.J. (2006), "Granger causality and the sampling of economic processes", *Journal of Econometrics*, Vol. 132 No. 2, pp. 311-336, doi: 10.1016/j.jeconom.2005.02.002.

Meddahi, N. and Renault, E. (2004), "Temporal aggregation of volatility models", *Journal of Econometrics*, Vol. 119 No. 2, pp. 355-379, doi: 10.1016/S0304-4076(03)00200-8.

Nelson, D.B. (1990), "ARCH models as diffusion approximations", *Journal of Econometrics*, Vol. 45 No. 1, pp. 7-38, doi: 10.1016/0304-4076(90)90092-8.

Ohanissian, A. (2001), "Using temporal aggregation to distinguish between true and spurious long memory", Working paper.

Otero, J., Panagiotidis, T. and Papapanagiotou, G. (2022), "Multivariate cointegration and temporal aggregation: some further simulation results", *Computational Economics*, Vol. 59 No. 1, pp. 59-70, doi: 10.1007/s10614-020-10062-w.

O'Neill, M. and Rajaguru, G. (2020), "A response surface analysis of critical values for the lead-lag ratio with application to high frequency and non-synchronous financial data", *Accounting and Finance (Parkville)*, Vol. 60 No. 4, pp. 3979-3990, doi: 10.1111/acfi.12546.

O'Neill, M. and Rajaguru, G. (2023), "Causality of price movements in VIX exchange-traded products and VIX futures contracts", *Journal of Accounting Literature*, Vol. 46 No. 2, pp. 153-169, doi: 10.1108/JAL-12-2022-0126.

Pesaran, M.H., Pierce, R.G. and Lee, K.C. (1991), “Persistence shocks and their sources in a multivariate model of U.K. output growth”, DAE Working Paper 9115, University of Cambridge, Department of Applied Economics.

Pierse, R.G. and Snell, A.J. (1995), “Temporal aggregation and the power of tests for a unit root”, *Journal of Econometrics*, Vol. 65 No. 2, pp. 333-345, doi: 10.1016/0304-4076(93)01589-E.

Rajaguru, G. (2004a), “Effects of temporal aggregation and systematic sampling model dynamics and causal inference”, Unpublished PhD dissertation, National University of Singapore.

Rajaguru, G. (2004b), “Impact of systematic sampling on causality in the presence of unit roots”, *Economics Letters*, Vol. 84 No. 1, pp. 127-132, doi: 10.1016/s0165-1765(04)00042-4.

Rajaguru, G. and Abeysinghe, T. (2008), “Temporal aggregation, cointegration and causality inference”, *Economics Letters*, Vol. 101 No. 3, pp. 223-226, doi: 10.1016/j.econlet.2008.08.012.

Rajaguru, G. and Abeysinghe, T. (2012), “The distortionary effects of temporal aggregation on Granger causality”, in Kuldeep, K. and Chaturvedi, A., *Some recent developments in statistical theory and applications: Selected Proceedings of the International Conference on Recent Developments in Statistics, Econometrics and Forecasting*, University of Allahabad, Brown Walker Press, India, December 27-28, 2010, pp. 38-56.

Rajaguru, G., O’Neill, M. and Abeysinghe, T. (2018), “Does systematic sampling preserve Granger causality with an application to high frequency financial data?”, *Econometrics*, Vol. 6 No. 2, p. 31, doi: 10.3390/econometrics6020031.

Rossana, R.J. and Seater, J.J. (1995), “Temporal aggregation and economic time series”, *Journal of Business and Economic Statistics*, Vol. 13 No. 4, pp. 441-451, doi: 10.1080/07350015.1995.10524618.

Silvestrini, A. and Veredas, D. (2005), “Temporal aggregation of univariate time series models”, CORE Discussion Paper 2005/59, Université catholique de Louvain, Louvain-la-Neuve, Belgium.

Silvestrini, A. and Veredas, D. (2008), “Temporal aggregation of univariate and multivariate time series models: a survey”, *Journal of Economic Surveys*, Vol. 22 No. 3, pp. 458-497, doi: 10.1111/j.1467-6419.2007.00538.x.

Sims, C.A. (1971), “Discrete approximations to continuous time distributed lags in econometrics”, *Econometrica*, Vol. 39 No. 3, pp. 545-563, doi: 10.2307/1913265.

Stram, D.O. and Wei, W.W.S. (1986), “Temporal aggregation in the ARIMA process”, *Journal of Time Series Analysis*, Vol. 7 No. 4, pp. 279-292, doi: 10.1111/j.1467-9892.1986.tb00495.x.

Swanson, N.R. and Granger, C.W.J. (2012), “Impulse response functions based on a causal approach to residual orthogonalization in vector autoregressions”, *Journal of the American Statistical Association*, Vol. 92 No. 437, pp. 357-367, doi: 10.1080/01621459.1997.10473634.

Teles, P. and Sousa, P.S.A. (2017), “The effect of temporal aggregation on the estimation accuracy of time series models”, *Communications in Statistics - Simulation and Computation*, Vol. 46 No. 9, pp. 6738-6759, doi: 10.1080/03610918.2016.1210169.

Telser, L.G. (1967), “Discrete samples and moving sums in stationary stochastic processes”, *Journal of the American Statistical Association*, Vol. 62 No. 318, pp. 484-499, doi: 10.1080/01621459.1967.10482922.

Tiao, G.C. (1972), “Asymptotic behaviour of temporal aggregates of time series”, *Biometrika*, Vol. 59 No. 3, pp. 525-531, doi: 10.1093/biomet/59.3.525.

Tiao, G.C. and Wei, W.S. (1976), “Effect of temporal aggregation on the dynamic relationship of two time series variables”, *Biometrika*, Vol. 63 No. 3, pp. 513-523, doi: 10.1093/biomet/63.3.513.

Tserkezos, D. (2021), “The effects of temporal aggregation and random sampling on the power of the augmented Dickey Fuller stationarity test: a Monte Carlo study”, in *Money, Trade and Finance: Recent Trends and Methodological Issues*, pp. 223-233, doi: 10.1007/978-3-030-73219-6_11.

Wei, W.W.S. (1978a), “Some consequences of temporal aggregation in seasonal time series models”, in Zellner, A. (Ed.), *Seasonal Analysis of Economic Time Series*, pp. 433-444.

Wei, W.W.S. (1978b), “The effect of temporal aggregation on parameter estimation in distributed lag model”, *Journal of Econometrics*, Vol. 8 No. 2, pp. 237-246, doi: 10.1016/0304-4076(78)90032-5.

Wei, W.W.S. (1981), "Effect of systematic sampling on ARIMA models", *Communications in Statistics - Theory and Methods*, Vol. 10 No. 23, pp. 2389-2398, doi: 10.1080/03610928108828197.

Wei, W.W.S. (1982), "Comment: the effects of systematic sampling and temporal aggregation on causality – a cautionary note", *Journal of the American Statistical Association*, Vol. 77 No. 378, pp. 316-319, doi: 10.1080/01621459.1982.10477806.

Wei, W.W.S. (1990), *Time Series Analysis: Univariate and Multivariate Methods*, Addison-Wesley Pub, Reading, MA.

Wei, W.W.S. and Mehta, J. (1980), "Temporal aggregation and information loss in a distributed lag model", in Anderson, O.D. (Ed.), *Analyzing Time Series*, North-Holland, pp. 613-617.

Weiss, A.A. (1984), "Systematic sampling and temporal aggregation in time series models", *Journal of Econometrics*, Vol. 26 No. 3, pp. 271-281, doi: 10.1016/0304-4076(84)90022-8.

Working, H. (1960), "Note on the correlation of first differences of averages in A random chain", *Econometrica: Journal of the Econometric Society*, Vol. 28 No. 4, pp. 916-918, doi: 10.2307/1907574.

Zellner, A. and Montmarquette, C. (1971), "A study of some aspects of temporal aggregation problems in econometric analyses", *The Review of Economics and Statistics*, Vol. 53 No. 4, pp. 335-342, doi: 10.2307/1928734.

Further reading

Shu, J. and Zhang, J.E. (2012), "Causality in the VIX futures market", *The Journal of Futures Markets*, Vol. 32 No. 1, pp. 24-46, doi: 10.1002/fut.20506.

Wei, W.W.S. and Tiao, G.C. (1975), "The effect of temporal aggregation on forecasting in the dynamic model", *ASA Proceedings of Business and Economic Section*, pp. 613-617.

Whaley, R.E. (1993), "Derivatives on market volatility: hedging tools long overdue", *The Journal of Derivatives*, Vol. 1 No. 1, pp. 71-84, doi: 10.3905/jod.1993.407868.

Whaley, R.E. (2000), "The investor fear gauge: explication of the CBOE VIX", *Journal of Portfolio Management*, Vol. 26 No. 3, pp. 12-17, doi: 10.3905/jpm.2000.319728.

Whaley, R.E. (2009), "Understanding the VIX", *Journal of Portfolio Management*, Vol. 35 No. 3, pp. 10-105, doi: 10.3905/jpm.2009.35.3.098.

Whaley, R.E. (2013), "Trading volatility: at what cost?", *Journal of Portfolio Management*, Vol. 40 No. 1, pp. 95-108, doi: 10.3905/jpm.2013.40.1.095.

Yang, Y.H. and Shao, Y.H. (2020), "Time-dependent lead-lag relationships between the VIX and VIX futures markets", *The North American Journal of Economics and Finance*, Vol. 53, 101196, doi: 10.1016/j.najef.2020.101196.

Zhang, J., Ruan, X. and Zhang, J.E. (2023), "Do short-term market swings improve realized volatility forecasts?", *Finance Research Letters*, Vol. 58, 104629, doi: 10.1016/j.frl.2023.10462.

Corresponding author
Gulasekaran Rajaguru can be contacted at: rgulasek@bond.edu.au