Article

Bootstrap Prediction Intervals of Temporal Disaggregation

Bu Hyoung Lee

Department of Mathematics and Statistics, Loyola University Maryland, 4501 N. Charles Street, Baltimore, MD 21210, USA; blee4@loyola.edu

Abstract: In this article, we propose an interval estimation method to trace an unknown disaggregate series within certain bandwidths. First, we consider two model-based disaggregation methods called the GLS disaggregation and the ARIMA disaggregation. Then, we develop iterative steps to construct AR-sieve bootstrap prediction intervals for model-based temporal disaggregation. As an illustration, we analyze the quarterly total balances of U.S. international trade in goods and services between the first quarter of 1992 and the fourth quarter of 2020.

Keywords: temporal disaggregation; temporal aggregation; bootstrap prediction intervals; ARIMA models; international trade balances

1. Introduction

Temporally aggregated data with calendrical periods, e.g., weekly, monthly, quarterly, or annually, have been widely used in various fields because the aggregation technique is simple and convenient for summarizing long sequential measurements and reducing their data lengths. However, it is inevitable that temporal aggregation causes a significant loss of information as a time series of relatively high frequency is compressed into periodic totals of relatively low frequency (see [1–3]). Thus, when decompressing the totals, restoring the original information remains a considerable challenge in time series analysis, and numerous solutions for temporal disaggregation have been proposed in modern statistical and econometric literature.

Refs. [4–8] solved disaggregation problems by using additional information obtained from related indicators observed at a desired high frequency. Refs. [9,10] estimated unknown consecutive observations in a disaggregate series based on the model structure of a given aggregate series. In addition, as discussed in [11–13], most of recent publications on disaggregation problems focus on exploring point estimation approaches. In other words, there has been, as yet, no systematic examination of interval estimation for temporal disaggregation. The aim of this study is therefore to propose an interval estimation method and to broaden our understanding of temporal disaggregation.

Refs. [14–16] introduced an AR-sieve bootstrap procedure to construct nonparametric prediction intervals for linear time series models such as stationary and invertible autoregressive integrated moving-average (ARIMA) processes. AR-sieve bootstrap interval estimation provides a theoretical basis for solving the model-based disaggregation problems of [9,10]. This article is organized as follows. Section 2 presents two model-based disaggregation methods—(1) the generalized least squares (GLS) disaggregation and (2) the ARIMA disaggregation, developed by [9,10], respectively. In Section 3, we propose a modified procedure for finding AR-sieve bootstrap prediction intervals of temporal disaggregation. In Section 4, we analyze a real time series, U.S. international trade balances in goods and services from 1992 to 2020, using the proposed approach. Moreover, some further remarks are given in Section 5.
2. Model-Based Temporal Disaggregation

2.1. The GLS Disaggregation

Suppose that \( x_t \) is a finite time series for \( t = 1, 2, \ldots, n \), and the \( d \)th differenced series

\[
u_t = (1 - B)^d x_t = \left( \sum_{i=0}^{d} \delta_i B^{d-i} \right) x_t \tag{1}\]

is covariance stationary with mean zero and autocovariance \( \gamma_u(j) \) for lag \( j \in \mathbb{Z} \), where \( B \) is the backshift operator of \( B^j x_t = x_{t-j} \), \( \delta_i \) is the \( i \)th coefficient of \( -B + 1 \)\(^d \) = \( \sum_{i=0}^{d} \delta_i B^{d-i} \), and \( \mathbb{Z} \) is the set of integers. The \( m \)-periodic temporal aggregate (or, simply, \( m \)-aggregate) series \( X_T \) is defined as the nonoverlapping sum of the \( m \) sequential observations, \( x_{m(T-1)+1}, x_{m(T-1)+2}, \ldots, x_{mT} \), i.e.,

\[
X_T = \sum_{t=m(T-1)+1}^{mT} x_t = \left( 1 + \sum_{j=1}^{m-1} B^j \right) x_{mT} \tag{2}
\]

for \( T = 1, 2, \ldots, N \) and \( n = mN \), and the \( d \)th differenced series of \( X_T \) is given by

\[
V_T = (1 - B)^d X_T = \left( \sum_{i=0}^{d} \delta_i B^{d-i} \right) X_T \tag{3}
\]

with autocovariance \( \gamma_V(j) \) for lag \( j \in \mathbb{Z} \), where \( B = B^m \) is the backshift operator of \( B^j X_T = X_{T-j} \). Let \( x = (x_1, x_2, \ldots, x_n)' \), \( u = (u_{d+1}, u_{d+2}, \ldots, u_n)' \), \( x = (X_1, X_2, \ldots, X_N)' \), and \( V = (V_{d+1}, V_{d+2}, \ldots, V_n)' \). The two expressions (1) and (3) can be rewritten as matrix forms, \( u = \Delta_n x \) and \( V = \Delta_n X \), where

\[
\Delta_n = \begin{pmatrix}
\delta_0 & \delta_1 & \cdots & \delta_d & 0 & \cdots & 0 \\
0 & \delta_0 & \delta_1 & \cdots & \delta_d & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
0 & \cdots & 0 & \delta_0 & \delta_1 & \cdots & \delta_d
\end{pmatrix}_{(d-1) \times d} \tag{4}
\]

In addition, \( u \) and \( V \) are linearly associated, i.e., \( V = \Omega u \) with

\[
\Omega = \begin{pmatrix}
w' & 0_{(N-d-1)m}' & 0_{(N-d-2)m}' & 0_{(N-d-3)m}' & \ldots & 0_{(N-d-1)m}' & w' \\
0_{m} & 0_{(N-d-1)m}' & 0_{(N-d-2)m}' & 0_{(N-d-3)m}' & \ldots & 0_{(N-d-1)m}' & w' \\
0_{m} & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
0_{m} & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
\end{pmatrix}_{(N-d) \times (n-d)} \tag{5}
\]

\( 0_k = (0, \ldots, 0)' \) of size \( k \), and \( w = (w_0, w_1, \ldots, w_{(m-1)(d+1)})' \), where \( w_i \) is the \( i \)th coefficient of \( (1 + B + \cdots + B^{m-1})^d+1 = \sum_{j=0}^{(m-1)(d+1)} w_i B^j \) and \( w_i = W_{(m-1)(d+1)-i} \). Then, the GLS disaggregation estimates of \( x \) and \( u \) are given by

\[
x = \Delta_n^{-1} \left( \Sigma_u \Omega' \Sigma_V^{-1} \Lambda_N \right) x \tag{6}
\]

and

\[
\hat{u} = \Sigma_u \Omega' \Sigma_V^{-1} V \tag{7}
\]

respectively, where \( \Sigma_u \) and \( \Sigma_V \) are the covariance matrices of \( u_t \) and \( V_T \), \( \delta_{i \times j} \) is the zero matrix of dimension \( i \times j \), \( \mathbf{1}_m = (1, \ldots, 1)' \) is the all-ones vector of size \( m \), and \( \mathbf{1}_d \) is the identity matrix of dimension \( d \times d \). The symbol \( \otimes \) denotes the Kronecker product, and the vertical and horizontal lines indicate partitioned matrices in (6). (For a more detailed discussion, see [9].) However, the solutions of either (6) or (7) cannot be directly derived
unless the covariance matrix $\Sigma_u$ is known. In the following subsection, a model-based solution for the unknown covariance matrix $\Sigma_u$ is discussed.

2.2. Disaggregate ARIMA Models

Ref. [17] showed a linear relationship between the autocovariances of $u_t$ and $V_T$ such that

$$\gamma_V(j) = \left(1 + B + \cdots + B^{m-1}\right)^{2(d+1)} \gamma_u(mj + (m-1)(d+1))$$

for $j = 0, 1, \ldots, k$ and $k \in \mathbb{Z}$, or equivalently,

$$\Gamma_V(k) = \mathbf{A} \Gamma_u(mk + (m-1)(d+1)),$$

where $\Gamma_u(mk + (m-1)(d+1)) = (\gamma_u(0), \gamma_u(1), \ldots, \gamma_u(mk + (m-1)(d+1)))'$, $\Gamma_V(k) = (\gamma_V(0), \gamma_V(1), \ldots, \gamma_V(k))'$, and $\mathbf{A}$ is a resultant coefficient matrix of dimension $(k+1) \times (mk + (m-1)(d+1) + 1)$ called the aggregation transformation matrix.

Ref. [10] pointed out that (9) is not necessarily a one-to-one mapping, and they solved the disaggregation problem under some reasonable assumptions on ARIMA models. Suppose that an ARIMA$(p,d,q)$ model of $x_t$,

$$\left(1 - \sum_{i=1}^{p} \phi_i B^i\right)(1 - B)^d x_t = \left(1 - \sum_{i=1}^{q} \theta_i B^i\right) \epsilon_t$$

is a disaggregation solution for an ARIMA$(p,d,r)$ model of $X_T$,

$$\left(1 - \sum_{i=1}^{p} \Phi_i B^i\right)(1 - B)^d X_T = \left(1 - \sum_{i=1}^{r} \Theta_i B^i\right) E_T$$

with $r \leq p + d + 1$, where $\epsilon_t$ and $E_T$ are Gaussian white noise with mean zero and constant variances $\sigma^2_\epsilon$ and $\sigma^2_E$, respectively. It is also assumed that the AR and MA polynomials of (10), $1 - \sum_{i=1}^{p} \phi_i B^i$ and $1 - \sum_{j=1}^{q} \theta_j B^j$, share no common factors. The same condition of no common factors is assumed for the model polynomials in (11), $1 - \sum_{i=1}^{p} \Phi_i B^i$ and $1 - \sum_{j=1}^{r} \Theta_j B^j$.

Under the condition of no hidden periodicity for aggregation order $m$ (see [18], pp. 512–513), the model parameters $\phi_i$ and $\theta_j$ in (10) can be solved as follows [10]:

1. The AR part:
   - When $m$ is odd and the AR polynomial of the aggregate model is factorized as
     $$1 - \sum_{i=1}^{p} \phi_i B^i = \prod_{i=1}^{p} (1 - \Psi_i B)$$

     the disaggregate AR parameters $\phi_i$ are given by
     $$1 - \sum_{i=1}^{p} \phi_i B^i = \prod_{i=1}^{p} \left(1 - \Psi_i^{1/m} B\right).$$

   - Otherwise, a disaggregate model cannot be uniquely determined.

2. The MA part:
   Ref. [19] showed that the $m$-aggregate model for ARIMA$(p,d,q)$ with $q \leq p + d + 1$ is ARIMA$(p,d,r)$ with $r \leq p + d + 1$. Thus, it is reasonable to assume that the maximum MA order of the disaggregate model is $q = p + d + 1$. 


Consider \( k = p + d + 1 \) in (9). Because \( \gamma_V(j) \) for \( j > d + 1 \) depends only on \( \gamma_u(i) \) for \( i > d + 1 \), the aggregation transformation matrix \( A \) in (9) can be partitioned as

\[
A = \begin{pmatrix}
    P_{11} & P_{12} \\
    0_{p \times (d+2)} & P_{22}
\end{pmatrix}
\]

(14)

where \( P_{11} \) is a square matrix of size \( d + 2 \), such as \( \Gamma_V(d+1) = P_{11} \Gamma_u(d+1) \), and \( P_{12} \) and \( P_{22} \) are block matrices of dimensions \( (d + 2) \times ((l+1) - (d+2)) \) and \( p \times ((l+1) - (d+2)) \), respectively, with \( l = m(p + d + 1) + (m-1)(d+1) \).

Using the approaches given in Theorem 1 of [10], Equation (9) is rewritten as

\[
\Gamma_V(p + d + 1) = \left( \begin{array}{cc}
P_{11} & P_{12} \\
0_{p \times (d+2)} & P_{22}
\end{array} \right) \Gamma_u(p + d + 1),
\]

(15)

where \( H \) is a coefficient matrix of dimension \( ((l+1) - (d+2)) \times p \), relating \( \gamma_u(i) \) to itself for \( d + 2 \leq i \leq p + d + 1 \) and \( \gamma_u(i) \) to \( \gamma_u(d+2), \gamma_u(d+3), \ldots, \gamma_u(p + d + 1) \) for \( i > p + d + 1 \). In other words, if the matrix in (15) is not singular, we have

\[
\Gamma_u(p + d + 1) = \left( \begin{array}{cc}
P_{11} & P_{12} \\
0_{p \times (d+2)} & P_{22}
\end{array} \right)^{-1} \Gamma_V(p + d + 1).
\]

(16)

As discussed in [10], the inverse matrix in (16) exists under the condition of no hidden periodicity for order \( m \). Then, we can derive the disaggregate MA parameters \( \theta_i \) from the autocovariances \( \gamma_u(0), \gamma_u(1), \ldots, \gamma_u(p + d + 1) \) included in \( \Gamma_u(p + d + 1) \). Two special cases of the ARIMA disaggregation are illustrated below.

**Example 1.** Consider an \( m \)-aggregate series \( X_T \) written as the ARIMA(1,1,0) model

\[
(1 - \Phi B)(1 - B)X_T = E_T
\]

(17)

with aggregation order \( m = 3 \). We remark that the ARIMA(1,1,3) model

\[
(1 - \phi B)(1 - B)x_t = \left(1 - \theta_1 B - \theta_2 B^2 - \theta_3 B^3\right) e_t
\]

(18)

is a possible solution for the unknown disaggregate series \( x_t \), under the conditions of \( p = 1 \) and \( q \leq p + d + 1 = 3 \). We find the model parameters of (18) as follows.

1. The disaggregate AR parameter:
   
   From Equations (12) and (13), it is clear that \( \phi = \Phi^{1/3} \).

2. The disaggregate MA parameters:
   
   Let \( V_T = (1 - B)X_T \) and \( u_t = (1 - B)x_t \). From (8) and (9), we find the linear relationship

   \[
   \gamma_V(j) = (1 + B + B^2)^j \gamma_u(3j + 4)
   \]

   for \( j = 0, 1, 2, 3 \), and so

   \[
   \Gamma_V(3) = \begin{pmatrix}
P_{11} & P_{12} \\
0_{1 \times 3} & P_{22}
\end{pmatrix} \Gamma_u(13),
\]

(19)

where

\[
P_{11} = \begin{pmatrix}19 & 32 & 20 \\
4 & 11 & 16
\end{pmatrix}, \quad P_{12} = \begin{pmatrix}8 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
19 & 16 & 10 & 4 & 1 & 0 & 0 & 0 & 0 & 0
\end{pmatrix},
\]

\[
P_{22} = \begin{pmatrix}0 & 0 & 1 & 4 & 10 & 16 & 19 & 16 & 10 & 4 & 1 & 0 & 0 & 0
\end{pmatrix},
\]

and \( P_{22} \). Moreover, since the AR(1) component implies \( \gamma_u(k) = \phi \gamma_u(k - 1) = \Phi^{1/3} \gamma_u(k - 1) \) for \( k > 3 \), \( \Gamma_u(13) \) can be rewritten as

\[
\Gamma_u(13) = \begin{pmatrix}I_5 & 0_{3 \times 1} \\
0_{11 \times 3} & H
\end{pmatrix} \Gamma_u(3),
\]

(20)
where \( H = (1, \Phi^{1/3}, \Phi^{2/3}, \ldots, \Phi^{10/3})' \). We therefore obtain

\[
\Gamma_V(3) = \begin{pmatrix}
19 & 32 & 20 \\
19 & 32 & 20 \\
0 & 0 & 1 \\
19 & 32 & 20 \\
0 & 0 & 1 \\
0 & 0 & 0 \\
0 & 0 & 0 \\
0 & 0 & 0 \\
\end{pmatrix}
\begin{pmatrix}
19 & 32 & 20 \\
19 & 32 & 20 \\
0 & 0 & 1 \\
19 & 32 & 20 \\
0 & 0 & 1 \\
0 & 0 & 0 \\
0 & 0 & 0 \\
0 & 0 & 0 \\
\end{pmatrix} \Gamma_u(3).
\] (21)

When the matrix in (21) is not a singular matrix, the autocovariances \( \gamma_u(0), \gamma_u(1), \gamma_u(2), \) and \( \gamma_u(3) \) can be solved. Consider \( \psi_t = (1 - \Phi^{1/3} B) u_t \), which is MA(3) from (18). The autocovariances of \( \psi_t \) are expressed as

\[
\gamma_w(0) = (1 + \Phi^{2/3}) \gamma_u(0) - 2\Phi^{1/3} \gamma_u(1) = (1 + \theta_1^2 + \theta_2^2 + \theta_3^2) \sigma_e^2,
\]
\[
\gamma_w(1) = -\Phi^{1/3} \gamma_u(0) + (1 + \Phi^{2/3}) \gamma_u(1) - \Phi^{1/3} \gamma_u(2) = (-\theta_1 + \theta_1 \theta_2 + \theta_2 \theta_3) \sigma_e^2,
\]
\[
\gamma_w(2) = -\Phi^{1/3} \gamma_u(1) + (1 + \Phi^{2/3}) \gamma_u(2) - \Phi^{1/3} \gamma_u(3) = (-\theta_2 + \theta_1 \theta_3) \sigma_e^2,
\]
\[
\gamma_w(3) = -\Phi^{1/3} \gamma_u(2) + (1 + \Phi^{2/3}) \gamma_u(3) - \Phi^{1/3} \gamma_u(4) = -\theta_3 \sigma_e^2.
\] (22)

Here, we can derive the MA parameters \( \theta_1, \theta_2, \) and \( \theta_3 \) from the above autocovariance equations, including the fixed values of \( \Phi^{1/3}, \gamma_u(0), \gamma_u(1), \gamma_u(2), \) and \( \gamma_u(3) \).

**Example 2.** Consider an \( m \)-aggregate series \( X_T \) expressed as the ARIMA(0,1,1) model

\[
(1 - B)X_T = (1 - \Theta B)E_T,
\] (23)

with \( m = 3 \). Based on the model conditions of \( p = 0 \) and \( q \leq p + d + 1 = 2 \), we suppose that the ARIMA(0,1,2) model

\[
(1 - B)x_t = (1 - \theta_1 B - \theta_2 B^2) e_t
\] (24)

is a possible solution for the unknown disaggregate series \( x_t \). The MA parameters of (24) are calculated as follows.

Let \( V_T = (1 - B)X_T \) and \( u_t = (1 - B)x_t \). From (8) and (9), we find the linear relationship

\[
\gamma_V(j) = (1 + B + B^2)^j \gamma_u(3j + 4) \text{ for } j = 0, 1, 2, \text{ and so}
\]

\[
\Gamma_V(2) = (P_{11} P_{12}) \Gamma_u(10),
\] (25)

where

\[
P_{11} = \begin{pmatrix}
19 & 32 & 20 \\
4 & 11 & 16 \\
0 & 0 & 1 \\
\end{pmatrix}
\text{ and } P_{12} = \begin{pmatrix}
8 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
19 & 16 & 10 & 4 & 1 & 0 & 0 & 0 & 0 & 0 \\
4 & 10 & 16 & 19 & 16 & 10 & 4 & 1 & 0 & 0 \\
\end{pmatrix}.
\]

Furthermore, since \( \gamma_u(3) = \cdots = \gamma_u(10) = 0 \), we can rewrite (25) as \( \Gamma_V(2) = P_{11} \Gamma_u(2) \). Thus, we obtain

\[
\Gamma_u(2) = P_{11}^{-1} \Gamma_V(2) = \begin{pmatrix}
11/81 & -32/81 & 292/81 \\
-4/81 & 19/81 & -224/81 \\
0 & 0 & 1 \\
\end{pmatrix} \Gamma_V(2).
\] (26)

That is, \( \gamma_u(0) = \frac{11}{81} \gamma_V(0) - \frac{32}{81} \gamma_V(1), \gamma_u(1) = -\frac{4}{81} \gamma_V(0) + \frac{19}{81} \gamma_V(1), \) and \( \gamma_u(2) = \gamma_V(2) = 0 \). These autocovariances imply that \( x_t \) reduces to the ARIMA(0,1,1) of \( (1 - B)x_t = (1 - \theta B)e_t \). We now solve the following equation and find the MA parameter \( \theta \),

\[
\frac{\gamma_u(0)}{\gamma_u(1)} = \frac{11 \gamma_V(0) - 32 \gamma_V(1)}{-4 \gamma_V(0) + 19 \gamma_V(1)} = \frac{1 + \theta^2}{-\theta}.
\] (27)

**3. AR-Sieve Bootstrap Prediction Intervals of Temporal Disaggregation**

In this section, we extend the framework of temporal disaggregation to interval estimation. The traditional parametric approaches of finding prediction intervals are not applicable to this disaggregation problem because the true distributions of realizations are unknown at disaggregate time points. Thus, we alternatively consider a nonparametric
method for \((1 - \alpha)\%\) AR-sieve bootstrap prediction intervals, developed by [14–16]. Based on the bootstrap approach, we present a modified procedure for constructing bootstrap prediction intervals of temporal disaggregation.

Step 1: Using the disaggregation method introduced in Section 2.2, we identify an ARIMA\((p, d, q)\) expression for the unknown disaggregate series \(x_t\) and find the autocovariance estimates, \(\hat{\gamma}_u(0), \hat{\gamma}_u(1), \ldots, \hat{\gamma}_u(p + d + 1)\), and the coefficient estimates, \(\hat{\phi}_1, \ldots, \hat{\phi}_p\) and \(\hat{\theta}_1, \ldots, \hat{\theta}_q\), of the disaggregate model. Then, we derive an estimated time series \(\hat{x}_t\) from the GLS disaggregation shown in Section 2.1.

Step 2: The \(d\)th differenced series \(u_t = (1 - B)^d x_t\) is assumed to be stationary and invertible in (10). The invertibility admits an AR\((\infty)\) representation for \(u_t\) such that

\[
(1 - \sum_{i=1}^{\infty} \pi_i B^i) u_t = e_t, \tag{28}
\]

and the \(i\)th coefficient estimate \(\hat{\pi}_i\) is derived from the numerical association

\[
1 - \sum_{i=1}^{\infty} \hat{\pi}_i B^i = \frac{1 - \sum_{i=1}^{p} \hat{\phi}_i B^i}{1 - \sum_{j=1}^{q} \hat{\theta}_j B^j}. \tag{29}
\]

Furthermore, we decide an appropriate AR approximation of order \(s\) under the condition of \(|\hat{\pi}_{s+h}| < \tau\) for \(h = 1, 2, \ldots\), where \(\tau\) is a predetermined positive value close to zero.

Step 3: We compute the centered residuals of the AR\((s)\) approximation, defined as

\[
\tilde{e}_t = e_t - \frac{1}{n - s} \sum_{t=s+1}^{n} \tilde{e}_t, \tag{30}
\]

for \(t = s + 1, \ldots, n\), where

\[
\tilde{e}_t = \left(1 - \sum_{i=1}^{s} \hat{\pi}_i B^i\right) \hat{u}_t. \tag{31}
\]

In addition, we obtain the empirical distribution function of the centered residuals,

\[
\hat{F}_n(\tilde{e}_t) = \frac{1}{n - s} \sum_{t=s+1}^{n} \mathbf{1}_{\tilde{e}_t \leq z}, \tag{32}
\]

where the indicator \(\mathbf{1}_{\tilde{e}_t \leq z} = 1\) if \(\tilde{e}_t \leq z\) or 0 elsewhere.

Step 4: We generate a bootstrap resample of i.i.d. observations \(e_t^*, t = 1, \ldots, n\) from \(\hat{F}_n(\tilde{e}_t)\) in (32).

Step 5: We derive a pseudo time series \(u_t^*\) from the forward model

\[
u_t^* = \hat{\pi}_1 u_{t-1}^* + \cdots + \hat{\pi}_s u_{t-s}^* + e_t^* \tag{33}
\]

for \(t = s + 1, \ldots, n\), assuming \(u_1^* = \cdots = u_s^* = 0\). Then, we replace the first \(s\) observations with the backward values

\[
u_t^* = \hat{\pi}_1 u_{t+1}^* + \cdots + \hat{\pi}_s u_{t+s}^* + e_t^* \tag{34}
\]

for \(t = s, s - 1, \ldots, 1\) (see [20], pp. 205–206).

Step 6: For the pseudo series \(u_t^*, t = 1, \ldots, n\), we obtain the AR coefficient estimates, \(\hat{\pi}_1^*, \ldots, \hat{\pi}_s^*\), through Yule–Walker estimation (see [21], pp. 239–240).
Step 7: We calculate the predicted bootstrap observations, \( \hat{x}_t^* = \hat{u}_t^* + \hat{x}_{t-1}^* \), using \( \hat{u}_t^* \) defined as
\[
\hat{u}_t^* = \hat{x}_1^* u_{t-1}^* + \cdots + \hat{x}_s^* u_{t-s}^* + e_t^*
\] (35)
for \( t = s + 1, \ldots, n \), and
\[
\hat{u}_t^* = \hat{x}_1^* u_{t+1}^* + \cdots + \hat{x}_s^* u_{t+s}^* + e_t^*
\] (36)
for \( t = s, s-1, \ldots, 1 \), and assuming \( \hat{x}_0^* = \frac{1}{n} \sum_{t=1}^{n} x_t \).

Step 8: We repeat Steps 4–7 many times and find the bootstrap distribution function of \( \hat{x}_t^* \), denoted by \( F_n(\hat{x}_t^*) \). Finally, we construct the \( (1 - \alpha) \)% prediction interval for the unknown disaggregate observation \( x_t \), given by
\[
\left[ Q^*(\frac{\alpha}{2}), Q^*(1 - \frac{\alpha}{2}) \right],
\] (37)
where \( Q^*(k) \) is the \( k \)th percentile of \( F_n(\hat{x}_t^*) \).

When the original procedure proposed by [14–16] and our modified version are compared, two main differences exist. First, the AIC (Akaike information criterion) or BIC (Bayesian information criterion) model identification is replaced with the GLS and ARIMA disaggregations in Step 1. From this perspective, the unknown disaggregate model can be directly selected from the structure of the known aggregate series without extra information about the model criteria of the disaggregate series. Next, while the \( h \)-step ahead forecasting is used in the original procedure for \( h > 0 \), we sequentially estimate the unknown disaggregate series with the one-step forward predictions for \( t = s + 1, \ldots, n \) and the one-step backward predictions for \( t = s, s-1, \ldots, 1 \) in Step 7. Since the one-step ahead forecast limit is narrower than the \( h \)-step ahead forecast limit in general, the modification can prevent the prediction intervals from being unreasonably wide or unbounded.

Furthermore, we conduct a simulation study to analyze the properties of the AR-sieve bootstrap resamples of Step 4. Suppose that a finite disaggregate series \( x_t, t = 1, 2, \ldots, 5000 \), follows an ARIMA\((1,1,1)\) model of \( (1 - \phi B)(1 - B)x_t = (1 - \theta)\epsilon_t \), where \( \epsilon_t \) is Gaussian white noise with mean 0 and standard deviation \( \sigma = \sqrt{3} \). We generate 12 different time series for all the coefficient pairs of \( \phi \) and \( \theta \in \{-0.75, -0.35, 0.35, 0.75\} \) and \( \phi \neq \theta \). Through Steps 2–4, we obtain the centered residuals \( \tilde{e}_t \) and an AR-sieve bootstrap resample of size 5000 for each case. For all 12 pairs, the probability density plots of the bootstrapped \( \tilde{e}_t \) are drawn in Figure 1, and their mean values, standard deviations, and the Shapiro–Wilk test results including the W-statistics and \( p \)-values are reported in Table 1. It is clear from these results that each bootstrap resample has a bell-shaped distribution analogous to the assumption of mean 0 and \( \sigma = \sqrt{3} \approx 1.73205 \), and its normality is not rejected at significance level \( \alpha = 0.05 \).
Figure 1. Density plots of bootstrap resamples of size 5000 for the centered residuals $\tilde{e}_t$.

Table 1. Mean values, standard deviations, and the Shapiro–Wilk test results for the bootstrap resamples drawn in Figure 1.

| $\phi$ | $\theta$ | Mean  | sd    | W-Statistic | $p$-Value |
|--------|----------|-------|-------|-------------|-----------|
| $-0.75$ | $-0.35$  | $-0.03279$ | $1.73068$ | $0.99247$   | $0.61783$ |
| $-0.75$ | $0.35$    | $0.00814$  | $1.78736$ | $0.99271$   | $0.64584$ |
| $-0.75$ | $0.75$    | $-0.00993$ | $1.67396$ | $0.99321$   | $0.70362$ |
| $-0.35$ | $-0.75$  | $-0.00597$ | $1.80319$ | $0.99254$   | $0.62554$ |
| $-0.35$ | $0.35$    | $0.00768$  | $1.73869$ | $0.99416$   | $0.80999$ |
| $-0.35$ | $0.75$    | $0.00223$  | $1.75250$ | $0.99419$   | $0.81240$ |
| $0.35$  | $-0.75$  | $0.03641$  | $1.68577$ | $0.99342$   | $0.72751$ |
| $0.35$  | $-0.35$  | $-0.01750$ | $1.74479$ | $0.99294$   | $0.67239$ |
| $0.35$  | $0.75$    | $-0.05127$ | $1.69289$ | $0.99169$   | $0.53010$ |
| $0.75$  | $-0.75$  | $0.02573$  | $1.76438$ | $0.99207$   | $0.57279$ |
| $0.75$  | $-0.35$  | $-0.01051$ | $1.79086$ | $0.99271$   | $0.64575$ |
| $0.75$  | $0.35$    | $-0.05146$ | $1.71417$ | $0.99310$   | $0.69043$ |

4. Real Data Analysis: The U.S. International Trade in Goods and Services

The U.S. Census Bureau and the U.S. Bureau of Economic Analysis periodically publish the total balance, defined as the difference between exports and imports, of U.S. international trade in goods and services. The trade balance data, as one of the main economic indicators, play an important role in explaining the U.S. economy (see [22]). To demonstrate the applicability of the bootstrap interval estimation proposed in Section 3, we investigate the quarterly balances of U.S. international trade in goods and services from the first quarter of 1992 to the fourth quarter of 2020, collected from [23,24].
Let $X_T$ be the quarterly total balance of U.S. international trade in goods and services at time $T = 1, 2, \ldots, 116$. Since the augmented Dickey–Fuller test for $X_T$ provides strong evidence (i.e., test statistic: $-2.1213$ and $p$-value: 0.5262) of a unit-root, the lag one differenced series $V_T = X_T - X_{T-1}$ is now considered. In addition, the BIC values of the fitted ARMA($p, q$) models of $V_T$ for $p$ and $q = 0, 1, \ldots, 10$, are computed with maximum likelihood (ML) estimation. Since the smallest BIC value $2439.75$ is found at $p = 0$ and $q = 1$ as listed in Table 2, it is believed that the MA(1) model of $V_T$ or, equivalently, the IMA(1,1) model of $X_t$ given below

$$V_T = (1 - B)X_T = (1 - \hat{\theta} B)E_T$$

is the most appropriate fit for the data. Here, $\hat{\gamma}_V(0) = 112841784$, $\hat{\gamma}_V(1) = -46810651$, and $\hat{\sigma}^2_T = 87918193$.

Table 2. The BIC values of the fitted ARMA($p,q$) models of $V_T$ for $p$ and $q = 0, 1, \ldots, 10$.

| $p = 0$ | 1  | 2  | 3  | 4  | 5  | 6  | 7  | 8  | 9  | 10 |
|--------|----|----|----|----|----|----|----|----|----|----|
| $q = 0$ | 2458.10 | 2447.40 | 2447.39 | 2451.29 | 2449.77 | 2454.26 | 2458.19 | 2461.61 | 2466.23 | 2470.66 | 2475.38 |
| 1      | 2439.75 | 2443.08 | 2447.79 | 2452.52 | 2454.36 | 2457.65 | 2462.10 | 2466.14 | 2470.88 | 2475.40 | 2479.71 |
| 2      | 2443.21 | 2447.80 | 2452.57 | 2456.98 | 2458.48 | 2462.05 | 2467.14 | 2467.64 | 2475.63 | 2480.12 | 2480.09 |
| 3      | 2447.81 | 2452.51 | 2457.18 | 2457.58 | 2462.46 | 2466.70 | 2471.40 | 2471.63 | 2476.32 | 2480.10 | 2480.64 |
| 4      | 2452.25 | 2456.36 | 2458.65 | 2462.00 | 2466.64 | 2471.14 | 2475.81 | 2472.15 | 2476.98 | 2484.85 | 2484.58 |
| 5      | 2454.33 | 2459.07 | 2462.15 | 2466.71 | 2471.48 | 2473.01 | 2479.69 | 2476.58 | 2481.51 | 2484.40 | 2489.01 |
| 6      | 2459.07 | 2462.63 | 2466.64 | 2471.28 | 2475.10 | 2479.20 | 2482.06 | 2486.42 | 2489.60 | 2489.12 | 2493.61 |
| 7      | 2462.52 | 2466.53 | 2471.24 | 2475.55 | 2476.00 | 2482.14 | 2486.52 | 2485.05 | 2494.68 | 2500.22 | 2498.04 |
| 8      | 2466.27 | 2470.60 | 2473.06 | 2476.34 | 2480.72 | 2486.30 | 2487.91 | 2490.70 | 2497.10 | 2498.50 | 2503.03 |
| 9      | 2470.83 | 2475.26 | 2480.07 | 2480.91 | 2485.65 | 2490.12 | 2494.56 | 2495.01 | 2496.75 | 2505.48 | 2510.27 |
| 10     | 2474.49 | 2479.21 | 2483.43 | 2484.44 | 2488.40 | 2492.39 | 2497.26 | 2497.14 | 2501.86 | 2505.34 | 2509.34 |

Assume that $X_T$ is an $m$-aggregate series of $m = 3$ and the monthly trade balance series $x_t$, $t = 1, 2, \ldots, 348$, is unknown and needs to be estimated. From (26) in Example 2, we obtain $\hat{\gamma}_u(0) = \frac{11}{12} \hat{\gamma}_V(0) - \frac{32}{39} \hat{\gamma}_V(1) = 33817290$ and $\hat{\gamma}_u(1) = -\frac{4}{12} \hat{\gamma}_V(0) + \frac{8}{39} \hat{\gamma}_V(1) = -16552710$. As described in (27), we solve the quadratic equation

$$\frac{\hat{\gamma}_u(0)}{\hat{\gamma}_u(1)} = -2.043006 = \frac{1 + \hat{\theta}^2}{\hat{\theta}}$$

and so find $\hat{\theta} = 0.8130118$ or $1.2299944$. However, due to the invertibility condition $|\theta| < 1$, we only retain the IMA(1,1) model of $x_t$,

$$(1 - B)x_t = (1 - 0.8130118B)v_t$$

with $\hat{\sigma}^2_t = \hat{\gamma}_u(0)/(1 + \hat{\theta}^2) = -\hat{\gamma}_u(1)/\hat{\theta}^2 = 20359741$.

Furthermore, we calculate $\hat{x}_t$, $t = 1, 2, \ldots, 348$, using the GLS disaggregate formula of (6). The monthly disaggregate estimates $\hat{x}_t$ and the monthly averages of the observed aggregate series, i.e., $X_T/3$, are compared in Figure 2. As illustrated in the lined graph of Figure 2, it seems obvious that $\hat{x}_t$ is well interpolated into the gaps between the monthly averages $X_T/3$ and sufficiently explains the unknown monthly pattern.
Next, we construct 95% AR-sieve bootstrap prediction intervals for the unknown monthly trade balances $x_t$. From Steps 1 and 2 in Section 3, assuming $\tau = 0.001$, we select the AR(33) approximation for $\hat{u}_t = \hat{x}_t - \hat{u}_{t-1}$, including the coefficient estimates

$$
\hat{\alpha}_1 = -0.81301, \quad \hat{\alpha}_2 = 0.66099, \quad \hat{\alpha}_3 = -0.53739, \quad \hat{\alpha}_4 = 0.43691, \quad \hat{\alpha}_5 = -0.35521, \\
\hat{\alpha}_6 = 0.28879, \quad \hat{\alpha}_7 = -0.23479, \quad \hat{\alpha}_8 = 0.19089, \quad \hat{\alpha}_9 = -0.15519, \quad \hat{\alpha}_{10} = 0.12617, \\
\hat{\alpha}_{11} = -0.10258, \quad \hat{\alpha}_{12} = 0.08340, \quad \hat{\alpha}_{13} = -0.06780, \quad \hat{\alpha}_{14} = 0.05513, \quad \hat{\alpha}_{15} = -0.04482, \\
\hat{\alpha}_{16} = 0.03644, \quad \hat{\alpha}_{17} = -0.02962, \quad \hat{\alpha}_{18} = 0.02408, \quad \hat{\alpha}_{19} = -0.01958, \quad \hat{\alpha}_{20} = 0.01592, \\
\hat{\alpha}_{21} = -0.01294, \quad \hat{\alpha}_{22} = 0.01052, \quad \hat{\alpha}_{23} = -0.00856, \quad \hat{\alpha}_{24} = 0.00696, \quad \hat{\alpha}_{25} = -0.00565, \\
\hat{\alpha}_{26} = 0.00460, \quad \hat{\alpha}_{27} = -0.00374, \quad \hat{\alpha}_{28} = 0.00304, \quad \hat{\alpha}_{29} = -0.00247, \quad \hat{\alpha}_{30} = 0.00201, \\
\hat{\alpha}_{31} = -0.00163, \quad \hat{\alpha}_{32} = 0.00133, \quad \hat{\alpha}_{33} = -0.00108.
$$

Using the AR-sieve technique presented in Step 3, we compute the empirical distribution function $\hat{F}_n(\hat{e}_t)$ of the centered residuals $\hat{e}_t$. The S-like shape of $\hat{F}_n(\hat{e}_t)$ in Figure 3 is the typical pattern of zero-mean Gaussian white noise. We remark that other AR approximation conditions, $\tau = 0.1, 0.01$, and 0.0001, are also investigated. The centered residuals for $\tau = 0.0001$ depict a similar pattern to $\tau = 0.001$, while the two other conditions do not. Thus, we only consider $\tau = 0.001$ as a predetermined positive value close to zero.

Through a large number (e.g., 25000 in our study) of consecutive iterations of Steps 4–7, we obtain the bootstrap distribution function $F_n(\hat{x}_t^*)$ in Figure 3 for each bootstrap prediction $\hat{x}_t^*$, and we construct 95% AR-sieve bootstrap prediction intervals for all unknown $x_t$, $t = 1, \ldots, 348$, in the form of (37). The 95% bootstrap prediction intervals of $x_t$ and the maximum and minimum $\hat{x}_t^*$ values of $F_n(\hat{x}_t^*)$ for $t = 1, \ldots, 348$ are drawn in Figure 4. We notice that the prediction intervals completely surround the sequence of the monthly averages $X_T/3$ over the indexed time and capture the exact pattern of the monthly averages. Therefore, we can conclude from Figure 4 that the AR-sieve bootstrap prediction intervals can provide reasonable bandwidths for estimation of the unknown monthly trade balances $x_t$. 

![Figure 2. Monthly averages of the observed aggregate series vs. estimated disaggregate series.](image-url)
5. Concluding Remarks

In this research, we have developed a modified procedure for constructing AR-sieve bootstrap prediction intervals of an unknown disaggregate series \( x_t \). As demonstrated in Section 4, when the upper and lower limits of the prediction intervals are linked over the indexed time, we can visually outline the sequential pattern of unknown disaggregate observations within certain bandwidths. Therefore, the proposed prediction intervals can provide a measure of reliability for temporal disaggregation and account for the uncertainty in estimating \( x_t \).

In addition, the proposed method is composed of the model-based point estimation introduced in Section 2 and the nonparametric bootstrap interval estimation described in Section 3. That is, the iterative procedure of bootstrap resampling generates empirical probability distributions for unknown disaggregate observations, while a disaggregate ARIMA model is established under the assumption of Gaussian innovations. From this
perspective, the proposed method can be interpreted as a semiparametric approach to estimating and predicting unknown \( x_t \) values.

Our underlying assumption is that the white noise distribution is Gaussian. However, non-Gaussian time series are also frequently found in real life. For example, if a set of measurements is a time series that contains whole numbered counts, they can be modeled with a Poisson count process. Refs. [25–28] developed the generalized AR models and related estimation methods for the Poisson scenarios. In future work, as an extension of the AR-sieve method to non-Gaussian models, we will explore a wide range of possibilities of the Poisson AR-sieve bootstrap prediction intervals of temporal disaggregation. Moreover, since the current GLS and ARIMA disaggregations do not meet the non-Gaussian conditions, we will examine other model-based disaggregation methods, for example, [8,29].

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The quarterly balances of U.S. international trade in goods and services from the first quarter of 1992 to the fourth quarter of 2020. See [25,24].

Acknowledgments: We wish to thank the Special Issue editors and reviewers for their helpful comments, which have improved the presentation of this article.

Conflicts of Interest: The author declares no conflict of interest.

References

1. Lee, B.H.; Wei, W.W.S. The use of temporally aggregated data on detecting a mean change of a time series process. Commun. Stat. Theory Methods 2017, 46, 5851–5871. [CrossRef]
2. Lee, B.H.; Park, J. A spectral measure for the information loss of temporal aggregation. J. Stat. Theory Pract. 2020, 14, 1–23. [CrossRef]
3. Lee, B.H.; Wei, W.W.S. The use of temporally aggregated data in modeling and testing a variance change in a time series. Commun. Stat. Simul. Comput. 2021, Forthcoming. [CrossRef]
4. Chow, G.C.; Lin, A. Best linear unbiased interpolation, distribution and extrapolation of time series by related series. Rev. Econ. Stat. 1971, 53, 372–375. [CrossRef]
5. Denton, F.T. Adjustment of monthly or quarterly series to annuals totals: An approach based on quadratic minimization. J. Am. Stat. Assoc. 1971, 66, 99–102. [CrossRef]
6. Fernández, R.B. A methodological note on the estimation of time series. Rev. Econ. Stat. 1981, 63, 471–476. [CrossRef]
7. Litterman, R.B. A random walk, Markov model for the distribution of time series. J. Bus. Econ. Stat. 1983, 1, 169–173.
8. Guerrero, V.M. Temporal disaggregation of time series: An ARIMA based approach. Int. Stat. Rev. 1990, 58, 29–46. [CrossRef]
9. Stram, D.O.; Wei, W.W.S. A methodological note on the disaggregation of time series totals. J. Time Ser. Anal. 1986, 7, 293–302. [CrossRef]
10. Wei, W.W.S.; Stram, D.O. Disaggregation of time series models. J. R. Stat. Soc. Ser. Stat. Methodol. 1990, 52, 453–467. [CrossRef]
11. Feijóo, S.R.; Caro, A.R.; Quintana, D.D. Methods for quarterly disaggregation without indicators: A comparative study using simulation. Comput. Stat. Data Anal. 2003, 43, 63–78. [CrossRef]
12. Moauro, F.; Savio, G. Temporal disaggregation using multivariate structural time series models. Econom. J. 2005, 8, 214–234. [CrossRef]
13. Sax, C.; Steiner, P. Temporal disaggregation of time series. R J. 2013, 5, 80–87. [CrossRef]
14. Alonso, A.M.; Péna, D.; Romo, J. Forecasting time series with sieve bootstrap. J. Stat. Plan. Inference 2002, 100, 1–11. [CrossRef]
15. Alonso, A.M.; Péna, D.; Romo, J. On sieve bootstrap prediction intervals. Stat. Probab. Lett. 2003, 65, 13–20. [CrossRef]
16. Alonso, A.M.; Sipols, A.E. A time series bootstrap procedure for interpolation intervals. Comput. Stat. Data Anal. 2008, 52, 1792–1805. [CrossRef]
17. Stram, D.O.; Wei, W.W.S. Temporal aggregation in the ARIMA process. J. Time Ser. Anal. 1986, 7, 279–292. [CrossRef]
18. Wei, W.W.S. Time Series Analysis: Univariate and Multivariate Methods, 2nd ed.; Addison Wesley: Boston, MA, USA, 2006.
19. Wei, W.W.S. Seasonal Analysis of Economic Time Series; CHAPTER Some Consequences of Temporal Aggregation in Seasonal Time Series Models; U.S. Department of Commerce, Bureau of the Census: Washington, DC, USA, 1978; pp. 433–444.
20. Box, G.E.P.; Jenkins, G.M.; Reinsel, G.C.; Ljung, G.M. Time Series Analysis: Forecasting and Control, 5th ed.; John Wiley & Sons: Hoboken, NJ, USA, 2016.
21. Brockwell, P.J.; Davis, R.A. Time Series: Theory and Methods, 2nd ed.; Springer: New York, NY, USA, 2006.
22. U.S. Census Bureau Economic Indicators. Available online: https://www.census.gov/economic-indicators (accessed on 28 February 2021).
23. U.S. International Trade Data. Available online: https://www.census.gov/foreign-trade/data/index.html (accessed on 13 February 2021).
24. International Trade in Goods and Services. Available online: https://www.bea.gov/data/intl-trade-investment/international-trade-goods-and-services (accessed on 13 February 2021).
25. Freeland, R.K.; McCabe, B.P.M. Analysis of low count time series data by Poisson autoregression. J. Time Ser. Anal. 2004, 25, 701–722. [CrossRef]
26. Fokianos, A.; Rahbek, A.; Tjøstheim, D. Poisson autoregression. J. Am. Stat. Assoc. 2009, 104, 1430–1439. [CrossRef]
27. Zhu, F.; Wang, D. Estimation and testing for a Poisson autoregressive model. Metrika 2011, 73, 211–230. [CrossRef]
28. Chen, C.W.S.; Lee, S. Generalized Poisson autoregressive models for time series of counts. Comput. Stat. Data Anal. 2016, 99, 51–67. [CrossRef]
29. Hotta, L.K.; Vasconcellos, K.L. Aggregation and disaggregation of structural time series models. J. Time Ser. Anal. 1999, 20, 155–171. [CrossRef]