EUROPEAN JOURNAL OF PURE AND APPLIED MATHEMATICS
Vol. 8, No. 4, 2015, 431-449
ISSN 1307-5543 – www.ejpam.com

On the Variance of Antithetic Time Series

Dennis Ridley[1,*], Pierre Ngnepieba[2]

[1] *SBI, Florida A&M University and Department of Scientific Computing, Florida State University, Tallahassee, FL., USA.*
[2] *Department of Mathematics, Florida A&M University, Tallahassee, FL, 32307, USA.*

Abstract. Combining antithetic time series is used to reduce model fitted and forecast mean square error (MSE). This is accomplished by removing the component of error that represents bias. The potential to reduce error is a function of variance in the time series. The greater the variance the greater is the potential for error reduction. But, the greater the variance the less is the efficacy of reversing correlation and combining antithetic time series. The percentage reduction in MSE increases with variance up to a limit then reduces.

2010 Mathematics Subject Classifications: AMS 62E20, 62F10

Key Words and Phrases: Antithetic time series theory, Time Series analysis, Bias reduction, Inverse correlation, Serial correlation, Forecasting

1. Introduction

Combining antithetic time series to reduce fitted MSE has applications in statistical process control, and the analysis of engineering, scientific, medical and economic time series. The method was demonstrated by Ridley and Ngnepieba [30] with the well-known CompanyX data. They also showed that it satisfies the Diebold and Mariano [12] test for significance in forecast improvement. Before that, the problem of bias in CompanyX data ARIMA models was observed by Chatfield and Prothero [6]. The problem of bias in time series modeling in general was discussed by Copas [10], Griliches [14], Kendall [18], Klein [20], Koyck [21], Marriott and Pope [25], and Nerlov [26].

Prior to time series applications, complementary antithetic random numbers r and $1-r$ were suggested by Hammersley and Morton [16] for Monte Carlo computer simulation (see also Hammersley and Handscomb [15]). The results were mixed (see Kleijnen [19], Hendry [17], Ripley [32] and Davidson and McKinnon [11]). Clements and Hendry [9] discussed the use of antithetic variates to establish various properties concerning bias in forecasting models (see also Calzolari [5], Fisher and Salmon [13] and Mariano and Brown [24]).

*Corresponding author.

Email addresses: dridley@fsu.edu (D. Ridley), Pierre.Ngnepieba@famu.edu (P. Ngnepieba)

Antithetic time series analysis as discussed in this paper is different. And, combining antithetic time series is fundamentally different from traditional combining. Traditional combining applies to independent estimates obtained from two or more different models, which may collectively contain more information than any one model. But, there is no explicit correction of systematic bias that might occur in any one or more of the models. For an extensive review of traditional combining see Armstrong [1], Bates and Granger [2], Bunn [4], Clemen [7], Clemen and Winkler [8], Makridakis and Hibon [23], and Makridakis, *et al.*, [22]. Antithetic combining applies to two estimates. The first estimate is obtained from the single best model, subject only to possible systematic bias. The second estimate is created so as to be inversely correlated with the first, and thereby eliminate systematic bias (Ridley, [27–29]). Since Ridley [29] illustrated antithetic time series combining for an extensive range of autoregressive processes and probability distributions, those results will not be repeated here. Ridley, Ngnepieba and Duke [31] studied the effect of combining parameters on combining. In this paper we extend the simulation studies and derive a closed form formulae for MSE before and after combining. Like in the demonstration with actual data by Ridley and Ngnepieba [30], the variance of a real time series is fixed and cannot be altered. Therefore, this paper focuses on simulated lognormal AR(1) time series where the variance can be altered to illustrate its effect.

The paper is organized as follows. The theory of correlation reversal and bias reduction as a function of variance is reviewed in sections 2 and 3. An analytical function for the percentage reduction in MSE is derived in section 4. The effect of variance on percentage reduction in MSE is illustrated by way of computer simulation in section 5. The conclusions in section 6 contain recommendations for future research.

2. Reverse Correlation and Variance

Definition 1. *Two random variables are antithetic if their correlation is negative. A bivariate collection of random variables is asymptotically antithetic if its limiting correlation approaches minus one asymptotically (see Ridley, [29] for a derivation).*

Definition 2. $\{X(\xi,t)\}$ *is an ensemble of random variables, where* ξ *belongs to a sample space and* t *belongs to an index set representing time, such that* X_t *is a discrete realization of a lognormal stationary stochastic process from the ensemble,* $\ln X_t \sim N(\mu,\sigma)$*, and* $X_t, t = 1,2,3,\ldots$ *are serially correlated.*

Remark 1. *Antithetic time series involves the application of a power transformation that cannot be applied directly to normally distributed numbers that by definition include negative numbers.*

The Ridley [29] antithetic time series theorem states that "if $X_t > 0, t = 1,2,3,\ldots$ is a discrete realization of a lognormal stochastic process, such that $\ln X_t \sim N(\mu,\sigma)$, then if the correlation between X_t and X_t^p is ρ_{XX^p}, then $\lim_{p\to 0^-,\sigma\to 0}\rho_{XX^p} = -1$." The standard deviation σ and variance σ^2 are for the logarithm of the time series. Therefore, in practice, they are naturally small. However, since there is no guarantee that the variance will actually approach zero, there will be some loss of efficacy in reversing correlation and combining antithetic time series. The purpose of this extension of Ridley and Ngnepieba [30] (and Ridley, [29] lognormal theory) is to illustrate the effect on percentage reduction in MSE for various values of σ^2.

3. Antithetic Combining and Bias Reduction

Before antithetic combining can correct bias in a time series model, bias must first occur. Bias occurs when a sample of data is observed, an autoregressive model is fitted to the data, and there is failure to realize all the assumptions of the model, leading to correlation between the error and the lagged variable. The antithetic combining model (Ridley, [29], Ridley and Ngnepieba, [30]), applied to fitted data, is summarized as follows:

$$\widehat{X}_{c,t} = w\widehat{X}_t + (1-w)\widehat{X}'_t, \;\; t = 1, 2, 3, \ldots, n$$

where $\widehat{X}_t$ are fitted values obtained from a time series model $X_t = \Phi X_{t-1} + \epsilon_t$, $t = 1, 2, 3, \ldots n$. The parameter $-\infty < w < \infty$ is a combining weight. The fitted values $\widehat{X}_t$ and $\widehat{X}'_t$ are antithetic in the sense that they contain components of error $\hat{\epsilon}_t$ and $\hat{\epsilon}'_t$, respectively, that are biased and when weighted, $w\hat{\epsilon}_t$ and $(1-w)\hat{\epsilon}'_t$ are perfectly negatively correlated. The antithetic component $\widehat{X}'_t$ is estimated from

$$\widehat{X}'_t = \overline{X} + r_{\hat{X}\hat{X}^p}\left(s_{\hat{X}}/s_{\hat{X}^p}\right)\left(\widehat{X}^p_t - \overline{\widehat{X}^p}\right), \; t = 1, 2, \ldots, n$$

where the exponent of the power transformation is set to the small negative value $p = -0.001$, r denotes sample correlation coefficient and s denotes sample standard deviation.

In data analysis, data are often assumed to be normally distributed in order to use well known normal theory. But, the distribution of many data, including particle physics, engineering, business and economic data is lognormal. For example, it is not possible to manufacture or sell a negative quantity, but the positive amount of the quantity is unlimited. Hence the distribution is positively skewed. In any case, if the data really are normally distributed, they can easily be transformed to a lognormal distribution by simple exponentiation. When the distribution of X_t is lognormal, the distribution of $\lim_{p\to 0^-} X^p_t$ is normal (Ridley, [29]). The exponent p is set to a small negative number to simulate the limit as $p \to 0^-$. Since it is fixed, it is not an added model parameter that must be estimated from data. Also, the role of p is perfect reversal of correlation, not the conversion of a non-linear model to a linear model.

The complete combining function for empirical data is given by

$$\widehat{X}_{c,t} = w\widehat{X}_t + (1-w)\left\{\overline{X} + \left(1 - k\sqrt{n-t+1}\right) r_{\hat{Z}\hat{Z}^p}\left(s_{\hat{Z}}/s_{\hat{Z}^p}\right)\left(\widehat{Z}^p_t - \overline{\widehat{Z}^p}\right)\right\}, \;\; t = 1, 2, \ldots, n,$$

where a shift parameter λ (suggested by Box and Cox, [3]) is used to facilitate the power transformation by adding λ to each X_t to obtain $Z_t = X_t + \lambda$ prior to applying the power transformation and subtracting it after conversion back to their original units, leaving the mean unchanged. Ridley [29] proved the counterintuitive result that $r_{\hat{Z}\hat{Z}^p}$ is independent of λ. The values of k, λ and w are selected so as to minimize the combined fitted MSE in $\widehat{X}_{c,t}$. The expectation is that if $\widehat{X}_t$ are biased, then $\widehat{X}_{c,t}$ will exhibit diminishing bias as p approaches zero from the left (p inside an infinitesimal neighborhood of zero but not zero) and the correlation $r_{\hat{Z}\hat{Z}^p}$ approaches -1. If $\widehat{X}_t$ are unbiased then $w = 1$ and the combined fitted values are just the original fitted values. The factor $k\sqrt{n-t+1}$ is to correct for any apparent heteroscedasticity in the sample data, and k is expected to be small, or zero if there is none. The antithetic combining calculations are performed by FOURCAST[†].

4. Analytic Per Unit Reduction in MSE Function of Variance

In order to gain some insight into how antithetic combining is affected by variance, consider the stationary model

$$X_t = \mu_X + \Phi\left(X_{t-1} - \mu_X\right) + \varepsilon_t,\ t = 2, 3, \dots, n, \tag{1}$$

where X_t are lognormally distributed, $\ln X_t \sim N(\mu, \sigma)$, and whose fitted values are

$$\widehat{X}_t = \widehat{\mu}_X + \widehat{\Phi}\left(X_{t-1} - \widehat{\mu}_X\right),\ t = 2, 3, \dots, n, \tag{2}$$

where $\widehat{\mu}_X$ and $\widehat{\Phi}$ are least-squares estimates of μ_X and Φ, respectively. The least squares estimate $\widehat{\Phi}$ of Φ (see Ridley [29]) is

$$\widehat{\Phi} = \Phi + \mathrm{Cov}\left(\varepsilon_t, X_{t-1}\right)/\mathrm{Var}\left(X_t\right), \tag{3}$$

where fitted values $\widehat{X}_t$ obtained from $\widehat{\mu}_X + \widehat{\Phi}\left(X_{t-1} - \widehat{\mu}_X\right)$ are biased if $\mathrm{Cov}\left(\varepsilon_t, X_{t-1}\right) \neq 0$. From (2) and (3), and given that the time series is stationary so that $\mathrm{Var}\left(X_t\right) = \mathrm{Var}\left(X_{t-1}\right)$, then as $n \to \infty$ and $\widehat{\mu}_X \to \mu_X$,

$$\begin{aligned} \widehat{X}_t =& \mu_X + \Phi\left(X_{t-1} - \mu_X\right) + \left\{\mathrm{Cov}\left(\varepsilon_t, X_{t-1}\right)/\mathrm{Var}\left(X_t\right)\right\}\left(X_{t-1} - \mu_X\right) \\ =& \mathbb{E}\left(X_t\right) + v_t + u_t, \end{aligned} \tag{4}$$

where v_t are purely random errors and $u_t = \left\{\mathrm{Cov}\left(\varepsilon_t, X_{t-1}\right)/\mathrm{Var}\left(X_t\right)\right\}\left(X_{t-1} - \mu_X\right)$ are systematic serially correlated errors.

From Appendix A, the total MSE before combining is given by (A1):

$$\mathrm{MSE}\left(\sigma^2\right) = \left[1 - \left\{\Phi + \frac{\mathrm{Cov}\left(\varepsilon_t, X_{t-1}\right)}{\mathrm{Var}\left(X_t\right)}\right\}^2\right]\mathrm{Var}\left(X_t\right)$$

From Appendix B, the MSE after combining is given by (B6):

$$\begin{aligned} \mathrm{MSEc}\left(\sigma^2\right) = \Bigg[&\left(w^2 - 2w\right)\left\{\Phi + \frac{\mathrm{Cov}\left(\varepsilon_t, X_{t-1}\right)}{\mathrm{Var}\left(X_t\right)}\right\}^2 + 1 - 2(1-w)\exp\left(\overline{\ln\widehat{X}} + s^2/2\right)\frac{\sigma^2/\sqrt{\mathrm{Var}(X_t)}}{\sqrt{\exp(\sigma^2) - 1}} \\ &+ (1-w)^2\exp\left(2\overline{\ln\widehat{X}} + s^2\right)\frac{\mathrm{Var}(\ln\widehat{X}_t)}{\mathrm{Var}(X_t)} \end{aligned}$$

[†]Application program EMC, Inc., `http://www.fourcast.net/fourcast`, Version 2012.1, File: CompanyX.zip

$$+\,2w(1-w)\left\{\Phi+\frac{\mathrm{Cov}\left(\varepsilon_t,X_{t-1}\right)}{\mathrm{Var}(X_t)}\right\}\exp\left(\overline{\ln\widehat{X}}+s^2/2\right)\frac{\mathrm{Cov}\left(X_{t-1},\ln\widehat{X}_t\right)}{\mathrm{Var}\left(X_t\right)}\Bigg]\mathrm{Var}\left(X_t\right)$$

From Appendix C, per unit reduction in MSE due to combining is given by (C1):

$$\mathrm{PMSEc}\left(\sigma^2\right)=\frac{A+B}{1-\left\{\Phi+\frac{\mathrm{Cov}(\varepsilon_t,X_{t-1})}{\mathrm{Var}(X_t)}\right\}^2},$$

where

$$A=-\left[\left[\left\{1-\frac{\mathrm{Cov}\left(\varepsilon_t,X_{t-1}\right)}{\mathrm{Var}\left(X_t\right)}\right\}^{-1}-1\right]^2\left\{\Phi+\frac{\mathrm{Cov}\left(\varepsilon_t,X_{t-1}\right)}{\mathrm{Var}\left(X_t\right)}\right\}^2+\left[1-\left\{1-\frac{\mathrm{Cov}\left(\varepsilon_t,X_{t-1}\right)}{\mathrm{Var}\left(X_t\right)}\right\}^{-2}\right]\right]$$
$$\times\frac{\sigma^2}{\left(\exp\left(\sigma^2\right)-1\right)}$$

and

$$B=-2\left\{1-\frac{\mathrm{Cov}\left(\varepsilon_t,X_{t-1}\right)}{\mathrm{Var}\left(X_t\right)}\right\}^{-1}\left[1-\left\{1-\frac{\mathrm{Cov}\left(\varepsilon_t,X_{t-1}\right)}{\mathrm{Var}\left(X_t\right)}\right\}^{-1}\right]\left\{\Phi+\frac{\mathrm{Cov}\left(\varepsilon_t,X_{t-1}\right)}{\mathrm{Var}\left(X_t\right)}\right\}$$
$$\times\frac{\mathrm{Cov}\left(X_{t-1},\ln\widehat{X}_t\right)}{\exp\left\{\mu+\frac{\sigma^2}{2}\right\}\left(\exp\left(\sigma^2\right)-1\right)}$$

and substituting for $\mathrm{Var}\left(X_t\right)$

$$\mathrm{PMSEc}\left(\sigma^2\right)=\frac{C+D}{1-\left\{\Phi+\frac{\mathrm{Cov}(\varepsilon_t,X_{t-1})}{\exp\{2\mu+\sigma^2\}(\exp(\sigma^2)-1)}\right\}^2},$$

where

$$C=-\left[\left[\left\{1-\frac{\mathrm{Cov}\left(\varepsilon_t,X_{t-1}\right)}{\exp\left\{2\mu+\sigma^2\right\}\left(\exp\left(\sigma^2\right)-1\right)}\right\}^{-1}-1\right]^2\left\{\Phi+\frac{\mathrm{Cov}\left(\varepsilon_t,X_{t-1}\right)}{\exp\left\{2\mu+\sigma^2\right\}\left(\exp\left(\sigma^2\right)-1\right)}\right\}^2\right.$$
$$\left.+\left[1-\left\{1-\frac{\mathrm{Cov}\left(\varepsilon_t,X_{t-1}\right)}{\exp\left\{2\mu+\sigma^2\right\}\left(\exp\left(\sigma^2\right)-1\right)}\right\}^{-2}\right]\right]\frac{\sigma^2}{\left(\exp\left(\sigma^2\right)-1\right)}$$

$$D=-2\left\{1-\frac{\mathrm{Cov}\left(\varepsilon_t,X_{t-1}\right)}{\exp\left\{2\mu+\sigma^2\right\}\left(\exp\left(\sigma^2\right)-1\right)}\right\}^{-1}\left[1-\left\{1-\frac{\mathrm{Cov}\left(\varepsilon_t,X_{t-1}\right)}{\exp\left\{2\mu+\sigma^2\right\}\left(\exp\left(\sigma^2\right)-1\right)}\right\}^{-1}\right]$$

$$\times\left\{\Phi+\frac{\operatorname{Cov}\left(\varepsilon_t,X_{t-1}\right)}{\exp\left\{2\mu+\sigma^2\right\}\left(\exp\left(\sigma^2\right)-1\right)}\right\}\frac{\operatorname{Cov}\left(\varepsilon_t,X_{t-1}\right)}{\exp\left\{\mu+\frac{\sigma^2}{2}\right\}\left(\exp\left(\sigma^2\right)-1\right)}$$

where $\frac{\operatorname{Cov}(\varepsilon_t,X_{t-1})}{\exp\{2\mu+\sigma^2\}(\exp(\sigma^2)-1)}\neq 1$.

We are interested in the per unit reduction in MSE for time series with different variances. Setting $\mu=0$,

$$\operatorname{PMSEc}\left(\sigma^2\right)=\frac{E+F}{1-\left\{\Phi+\frac{\operatorname{Cov}(\varepsilon_t,X_{t-1})}{\exp\{\sigma^2\}(\exp(\sigma^2)-1)}\right\}^2},\tag{5}$$

where

$$E=-\left[\left\{1-\frac{\operatorname{Cov}\left(\varepsilon_t,X_{t-1}\right)}{\exp\left\{\sigma^2\right\}\left(\exp\left(\sigma^2\right)-1\right)}\right\}^{-1}-1\right]^2\left\{\Phi+\frac{\operatorname{Cov}\left(\varepsilon_t,X_{t-1}\right)}{\exp\left\{\sigma^2\right\}\left(\exp\left(\sigma^2\right)-1\right)}\right\}^2$$
$$+\left[1-\left\{1-\frac{\operatorname{Cov}\left(\varepsilon_t,X_{t-1}\right)}{\exp\left\{\sigma^2\right\}\left(\exp\left(\sigma^2\right)-1\right)}\right\}^{-2}\right]\frac{\sigma^2}{\left(\exp\left(\sigma^2\right)-1\right)},$$

$$F=-2\left\{1-\frac{\operatorname{Cov}\left(\varepsilon_t,X_{t-1}\right)}{\exp\left\{\sigma^2\right\}\left(\exp\left(\sigma^2\right)-1\right)}\right\}^{-1}\left[1-\left\{1-\frac{\operatorname{Cov}\left(\varepsilon_t,X_{t-1}\right)}{\exp\left\{\sigma^2\right\}\left(\exp\left(\sigma^2\right)-1\right)}\right\}^{-1}\right]$$
$$\times\left\{\Phi+\frac{\operatorname{Cov}\left(\varepsilon_t,X_{t-1}\right)}{\exp\left\{\sigma^2\right\}\left(\exp\left(\sigma^2\right)-1\right)}\right\}\frac{\operatorname{Cov}\left(X_{t-1},\ln\hat{X}_t\right)}{\exp\left\{\frac{\sigma^2}{2}\right\}\left(\exp\left(\sigma^2\right)-1\right)},$$

and $\frac{\operatorname{Cov}(\varepsilon_t,X_{t-1})}{\exp\{\sigma^2\}(\exp(\sigma^2)-1)}\neq 1$.

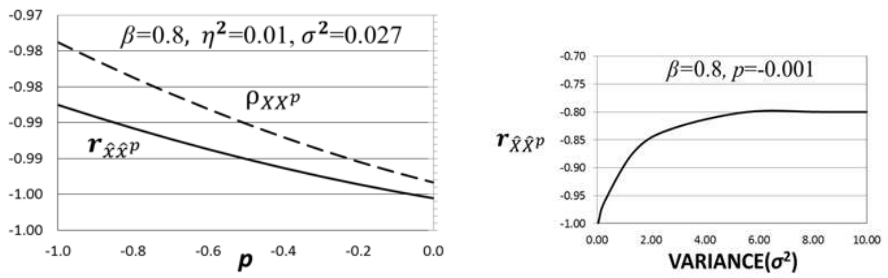
This analytic function (5) for the per unit reduction in MSE due to combining is somewhat complicated. The errors ε_t are unobservable. Furthermore, since the model estimate of parameter Φ is biased, its true value is unknown. The objective of antithetic time series analysis is not to correct the bias in the model parameter. The objective is to correct the bias in the model fitted values. Still, due to these unknowns, it is not possible to evaluate this function. An illustration of how combining reduced MSE, and how it changes with variance, can be accomplished by computer simulations. The simulations and results are explained below in Section 5.

5. Simulated Percentage Reduction in MSE and Variance

The ability to reverse the correlation ($\lim_{p\to 0^-,\sigma\to 0}\rho_{XX^p}=-1$) depends on small variance. Therefore, the ability to reduce bias and combined fitted MSE also depends on small variance.

However, the sample variance will depend on the particular data. To illustrate the effect of variance, let us consider the results of hypothetical simulations for the lognormal time series $X_t = \exp\left(Y_t\right)$, where $Y_t = \beta Y_{t-1} + e_t$, $t = 2, 3, \ldots, 1000$, $e_t \sim N\left(0, \eta\right)$ and $\sigma^2 = \frac{\eta^2}{\left(1-\beta^2\right)}$. Let the model to be fitted to X_t be $X_t = \Phi X_{t-1} + \epsilon_t$, $t = 2, 3, \ldots, 1000$. The least squares estimate $\widehat{\Phi}$ of Φ and the fitted values $\hat{X}_t$ obtained from $\widehat{\Phi} X_{t-1}$ will be biased if there is any deficiency (eg. inter alia, lack of normality, non-stationarity, unavoidably missing variables, sampling bias, serial correlation in the errors) resulting in the covariance $\mathrm{Cov}\left(\varepsilon_t, X_{t-1}\right) \neq 0$. Since X_t are positive and stationary, k and λ (see section 3) are both zero.

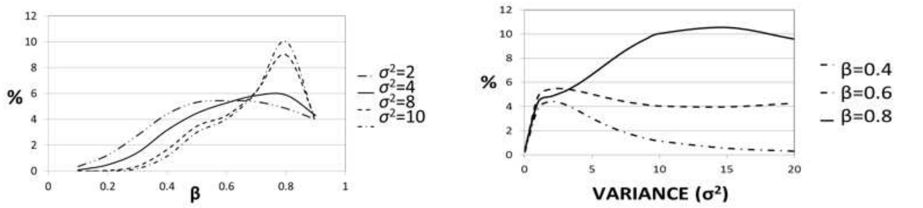
Consider the case of $\eta^2 = 0$, $\sigma^2 = 0$. That is, no variance, no variations in the time series to be explained by the time series model, and no MSE. Then, no improvement in MSE is possible by antithetic combining. Antithetic combining does not include $\sigma^2 = 0$. To simulate small variance, consider $\eta^2 = 0.01$, $\sigma^2 = 0.027$ and $p = -0.001$ as shown in Figure 1a. The correlation $r_{\hat{X}\hat{X}^p}$ starts out at approximately -1 where the combining contribution to MSE reduction is near maximum. As the variance is increased, the amount of variance unexplained by the fitted model increases, contributing to the MSE improvement that is possible. At the same time, the correlation moves away from -1 (see Figure 1b), reducing the ability for antithetic combining to improve MSE. The net effect is illustrated below in Figures 2 and 3.



(a) Correlation coefficient vs p (b) Correlation coefficient vs σ^2

Figure 1: Correlation coefficient of simulated population values ρ_{XX^p} and fitted values $r_{\hat{X}\hat{X}^p}$.

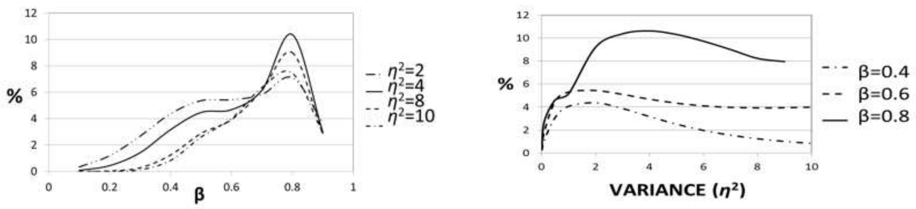
Figure 2a shows the percentage improvement in MSE for different levels of variance ($\sigma^2 = 2, 4, 8, 10$) as the autoregressive coefficient β increases. For $\sigma^2 = 2$, the percentage improvement increases until about $\beta = 0.55$ then falls. For larger σ^2, the percentage improvement increases until about $\beta = 0.8$ then falls. Figure 2b shows the percentage improvement in MSE for different levels of autoregressive coefficient ($\beta = 0.4, 0.6, 0.8$) as σ^2 increases. For all values of β, the improvement increases up to a maximum value then falls. For example, when $\beta = 0.8$, the maximum improvement in MSE is about 10.5% and occurs when σ^2 is about 15. Therefore, in practice, antithetic combining will perform more or less well depending on the variance of an actual time series. Forecasting is outside the scope of this paper, but Ridley and Ngnepieba [30] demonstrated how small idiopathic error bias accumulates to produce large ex ante multiple-period-ahead forecast errors that are corrected dynamically by antithetic combining.



(a) Percentage improvement in MSE vs β (b) Percentage improvement in MSE vs σ^2

Figure 2: Percentage improvement in MSE for $p = -0.001$ and versus β and σ^2.

Figure 3a shows the percentage improvement in MSE for different levels of variance ($\eta^2 = 2, 4, 8, 10$) as the autoregressive coefficient β increases. For all levels of η^2, the percentage improvement increases until about $\beta = 0.8$ then falls. Figure 3b shows the percentage improvement in MSE for different levels of autoregressive coefficient ($\beta = 0.4, 0.6, 0.8$) as η^2 increases. For all values of β, the improvement increases up to a maximum value then falls. For example, when $\beta = 0.8$, the maximum improvement in MSE is about 10.6% and occurs when η^2 is about 4.



(a) Percentage improvement in MSE vs β (b) Percentage improvement in MSE vs η^2

Figure 3: Percentage improvement in MSE for $p = -0.001$ and versus β and η^2.

6. Conclusions

The variance of the logarithm of a time series is naturally small. Therefore, combining antithetic time series to reduce idiopathic bias in the fitted values from a time series model is generally applicable. Furthermore, as the variance increases, the percentage improvement in fitted MSE is actually better. This is due to the higher potential to improve MSE. As the variance continues to increase, thereby departing from the strict requirements of antithetic time series theory, the ability of antithetic combining to reverse correlation and improve MSE diminishes (some similarity in the loss of efficacy of feedforward control was reported by Shi and Kapur [34] [33]). As the two effects interact, the net percentage improvement in fitted MSE increases to a maximum then it declines.

ACKNOWLEDGEMENTS The authors would like to thank Minjing Tao for her helpful comments and suggestions to improve the clarity of the paper, and Dennis Duke for good discussions.

References

[1] J. S. Armstrong. Combining forecasts: The end of the beginning or the beginning of the end. *International Journal of Forecasting*, 5(4):585–588, 1989.

[2] J. M. Bates and C. W. J. Granger. The Combination of Forecasts. *Operational Research Quarterly*, 20(4):451–468, 1969.

[3] G. E. Box and D. R. Cox. An analysis of transformations. *Journal of the Royal Statistical Society B*, 26(2):211–252, 1964.

[4] D. W. Bunn. The synthesis of predictive models in marketing research. *Journal of Marketing Research*, 16(May):280–283, 1979.

[5] G. Calzolari. Antithetic variates to estimate the simulation bias in non-linear models. *Economics Letters*, 4(4):323–328, 1979.

[6] C. Chatfield and D. L. Prothero. Box-Jenkins seasonal forecasting: problems in a case study. *Journal of the Royal Statistical Society: Series A*, 136(3):295–336, 1973.

[7] R. T. Clemen. Combining forecasts: a review and annotated bibliography. *International Journal of Forecasting*, 5(4):559–583, 1989.

[8] R. T. Clemen and R. L. Winkler. Combining economic forecasts. *Journal of Business and Economic Statistics*, 4(1):39–46, 1986.

[9] M. P. Clements and D. F. Hendry. *Forecasting Economic Time Series.* University Press, Cambridge: Cambridge, 1998.

[10] J. B. Copas. Monte Carlo results for estimation in a stable Markov time series. *Journal of the Royal Statistical Society, Series A*, 129(1):110–116, 1966.

[11] R. Davidson and McKinnon. *Estimation and Inference in Econometrics.* Oxford University Press, Oxford, 1993.

[12] F. X. Diebold and R. S. Mariano. Comparing Predictive Accuracy. *Journal of Business and Economic Statistics*, 13(3):134–144, 1995.

[13] P. Fisher and M. Salmon. On evaluating the importance of nonlinearity in large macroeconometric models. *International Economic Review*, 27(3):625–646, 1986.

[14] Z. Griliches. A note on serial correlation bias in estimates of distributed lags. *Econometrica*, 29(1):65–73, 1961.

[15] J. M. Hammersley and D. C. Handscomb. *Monte Carlo Methods*. Chapman and Hall, London, 1964.

[16] J. M. Hammersley and K. W. Morton. A new Monte Carlo technique: antithetic variates. *Proceedings of the Cambridge Philosophical Society*, 52(3):449–475, 1956.

[17] D. F. Hendry. Monte Carlo experimentation in econometrics. In Z. Griliches and M. D. Intriligator, editors, *Handbook of Econometrics, vol. 2*, volume 2, Amsterdam, 1984. North-Holland Publishing Co.

[18] M. G. Kendall. Note on bias in the estimation of autocorrelation. *Biometrika*, 41(3-4):403–404, 1954.

[19] J. P. C. Kleijnen. Antithetic variates, common random numbers and optimal computer time allocation in simulation. *Management Science*, 21(10):1176–1185, 1975.

[20] L. R. Klein. The Estimation of Distributed Lags. *Econometrica*, 26(4):553–618, 1958.

[21] L. M. Koyck. *Distributed Lags and Investment Analysis.* North-Holland Publishing Co., Amsterdam, 1954.

[22] S. Makridakis, A. Anderson, R. Carbone, R. Fildes, M. Hibon, R. Lewandowski, J. Newton, E. Parzen, and R. Winkler. The accuracy of extrapolation (times series) methods: Results of a forecasting competition. *Journal of Forecasting*, 1(2):111–153, 1982.

[23] S. Makridakis and M. Hibon. Accuracy of forecasting:An empirical investigation (with discussion). *Journal of the Royal Statistical Society, Series A*, 142(2):97–145, 1979.

[24] R. S. Mariano and B. W. Brown. Stochastic simulation, prediction and validation of nonlinear models. In L. R. Kline and J. Marquez, editors, *Economics in Theory and Practice: an Eclectic Approach*.

[25] F. H. C. Marriott and J. A. Pope. Bias in the estimation of autocorrelations. *Biometrika*, 41(3/4):390–402, 1954.

[26] M. Nerlove. Distributed Lags and Demand Analysis for Agricultural and Other Commodities. In *U.S. D.A. Agricultural Handbook No. 141*, Washington, 1958. U.S. Department of Agriculture.

[27] A. D. Ridley. Combining Global Antithetic Forecasts. *International Transactions in Operational Research*, 2(4):387–398, 1995.

[28] A. D. Ridley. Optimal Weights for Combining Antithetic Forecasts. *Computers & Industrial Engineering*, 2(2):371–381, 1997.

[29] A. D. Ridley. Optimal antithetic weights for lognormal time series forecasting. *Computers & Operations Research*, 26(3):189–209, 1999.

[30] A. D. Ridley and P. Ngnepieba. Antithetic time series analysis and the CompanyX data. *Journal of the Royal Statistical Society A*, 177(1):83–94, 2014.

[31] A. D. Ridley, P. Ngnepieba, and D. Duke. Parameter Optimization for Combining Lognormal Antithetic Time Series. *European Journal of Mathematical Sciences*, 2(2):235–245, 2013.

[32] B. D. Ripley. *Stochastic Simulation.* Wiley, New York, June, 1987.

[33] L. Shi and K. Kapur. Quasi Feedforward and Feedback Control for Random Step Shift Disturbance Models. *Quality Technology and Quantitative Management*, 12(1):69–82, 2014.

[34] L. Shi and K. Kapur. A Synthesis of Feedback and Feedforward Control for Stationary and Nonstationary Disturbance Models. *Quality, Reliability, and Engineering International*, 31(3):343–354, 2015.

Appendix A: MSE Before Combining

Substituting from (4), the total error in $\hat{X}_t$, including purely random error and systematic error due to bias is $\hat{X}_t - X_t = +\Phi\left(X_{t-1} - \mu_X\right) + \left\{\frac{\text{Cov}(\varepsilon_t, X_{t-1})}{\text{Var}(X_t)}\right\}\left(X_{t-1} - \mu_X\right) - X_t$. Since μ_X, Φ and $\frac{\text{Cov}(\varepsilon_t, X_{t-1})}{\text{Var}(X_t)}$ are constant for a specified time series and model, the variance of the total error before combining is the MSE, given by

$$\begin{aligned}
\text{MSE}\left(\sigma^2\right) =& \text{Var}\left(\hat{X}_t - X_t\right) \\
=& \text{Var}\left[\mu_X + \Phi\left(X_{t-1} - \mu_X\right) + \left\{\frac{\text{Cov}\left(\varepsilon_t, X_{t-1}\right)}{\text{Var}\left(X_t\right)}\left(X_{t-1} - \mu_X\right) - X_t\right\}\right] \\
=& \left[\left\{\Phi + \frac{\text{Cov}\left(\varepsilon_t, X_{t-1}\right)}{\text{Var}\left(X_t\right)}\right\}^2 + 1\right]\text{Var}\left(X_t\right) - 2\left\{\Phi + \frac{\text{Cov}\left(\varepsilon_t, X_{t-1}\right)}{\text{Var}\left(X_t\right)}\right\}\text{Cov}\left(X_t, X_{t-1}\right)
\end{aligned}$$

From (1), $\text{Cov}\left(X_t, X_{t-1}\right) = \Phi\text{Var}\left(X_t\right) + \text{Cov}\left(\varepsilon_t, X_{t-1}\right)$. Therefore,

$$\begin{aligned}
\text{MSE}\left(\sigma^2\right) =& \left[\left\{\Phi + \frac{\text{Cov}\left(\varepsilon_t, X_{t-1}\right)}{\text{Var}\left(X_t\right)}\right\}^2 + 1\right]\text{Var}\left(X_t\right) \\
&- 2\left\{\Phi + \frac{\text{Cov}\left(\varepsilon_t, X_{t-1}\right)}{\text{Var}\left(X_t\right)}\right\}\left\{\Phi\text{Var}\left(X_t\right) + \frac{\text{Cov}\left(\varepsilon_t, X_{t-1}\right)}{\text{Var}\left(X_t\right)}\text{Var}\left(X_t\right)\right\} \\
=& \left[\left\{\Phi + \frac{\text{Cov}\left(\varepsilon_t, X_{t-1}\right)}{\text{Var}\left(X_t\right)}\right\}^2 + 1 - 2\left\{\Phi + \frac{\text{Cov}\left(\varepsilon_t, X_{t-1}\right)}{\text{Var}\left(X_t\right)}\right\}\right.
\end{aligned}$$

$$\times\left\{\Phi+\frac{\operatorname{Cov}\left(\varepsilon_t,X_{t-1}\right)}{\operatorname{Var}\left(X_t\right)}\right\}\Bigg]\operatorname{Var}\left(X_t\right)$$

$$=\left[\left\{\Phi+\frac{\operatorname{Cov}\left(\varepsilon_t,X_{t-1}\right)}{\operatorname{Var}\left(X_t\right)}\right\}^2+1-2\left\{\Phi+\frac{\operatorname{Cov}\left(\varepsilon_t,X_{t-1}\right)}{\operatorname{Var}\left(X_t\right)}\right\}^2\right]\operatorname{Var}\left(X_t\right)$$

$$=\left[1-\left\{\Phi+\frac{\operatorname{Cov}\left(\varepsilon_t,X_{t-1}\right)}{\operatorname{Var}\left(X_t\right)}\right\}^2\right]\operatorname{Var}\left(X_t\right) \tag{A1}$$

Appendix B: MSE After Combining

Consider the combined fitted values $\widehat{X}_{c,t}=w\widehat{X}_t+(1-w)\widehat{X}_t'$, where

$$\widehat{X}_t=\mu_X+\Phi\left(X_{t-1}-\mu_X\right)+\left\{\frac{\operatorname{Cov}\left(\varepsilon_t,X_{t-1}\right)}{\operatorname{Var}\left(X_t\right)}\right\}\left(X_t-\mu_X\right) \tag{B1}$$

and $\widehat{X}_t'=\lim_{s_{\widehat{X}}\to 0,\,p\to 0^-}\left\{\overline{X}+r_{\hat{X}\hat{X}^p}\left(s_{\hat{X}}/s_{\hat{X}^p}\right)\left(\widehat{X}_t^p-\overline{\widehat{X}_t^p}\right)\right\}$. Since we are investigating the effect of changing variance, consider only the limit as $p\to 0^-$, and allow s_x to retreat from 0 as follows

$$\widehat{X}_{c,t}=w\widehat{X}_t+(1-w)\left\{\overline{X}+r_{\hat{X}\hat{X}^p}\left(s_{\hat{X}}/s_{\hat{X}^p}\right)\left(\widehat{X}_t^p-\overline{\widehat{X}_t^p}\right)\right\},\ p\to 0^-. \tag{B2}$$

For the case of lognormal X_t, such that $\ln X_t\sim N(\mu,\sigma)$, the expected values and standard deviations are $\mathbb{E}\left[X_t\right]=\exp\left(\mu+\frac{\sigma^2}{2}\right)$, $\mathbb{E}\left[X_t^p\right]=\exp\left(p\mu++\frac{p^2\sigma^2}{2}\right)$ and $\sigma_X=\exp\left\{\mu+\frac{\sigma^2}{2}\right\}\sqrt{\exp\left(\sigma^2\right)-1}$, $\sigma_{X^p}=\exp\left\{p\mu+\frac{p^2\sigma^2}{2}\right\}\sqrt{\exp\left(p^2\sigma^2\right)-1}$, respectively. Replacing population values with sample values and substituting in (B2),

$$\begin{aligned}\widehat{X}_{c,t}=&w\widehat{X}_t+(1-w)\left[\overline{X}+r_{\hat{X}\hat{X}^p}\left(s_{\hat{X}}/s_{\hat{X}^p}\right)\frac{\exp\left\{\overline{\ln\hat{X}}+\frac{s^2}{2}\right\}\sqrt{\exp\left(s^2\right)-1}}{\exp\left\{p\overline{\ln\hat{X}}+\frac{p^2s^2}{2}\right\}\sqrt{\exp\left(p^2s^2\right)-1}}\right.\\&\left.\times\left\{\widehat{X}_t^p-\exp\left(p\overline{\ln\hat{X}}+\frac{p^2s^2}{2}\right)\right\}\right],\ p\to 0^-\\=&w\widehat{X}_t+(1-w)\left[\overline{X}+\varphi\left(p,\widehat{X}_t,s\right)\right],\ p\to 0^-,\end{aligned} \tag{B3}$$

where

$$\varphi\left(p,\widehat{X}_t,s\right)=r_{\hat{X}\hat{X}^p}\frac{\exp\left\{\overline{\ln\hat{X}}+\frac{s^2}{2}\right\}\sqrt{\exp\left(s^2\right)-1}}{\exp\left\{p\overline{\ln\hat{X}}+\frac{p^2s^2}{2}\right\}\sqrt{\exp\left(p^2s^2\right)-1}}\left[\widehat{X}_t^p-\exp\left(p\overline{\ln\hat{X}}+\frac{p^2s^2}{2}\right)\right],\ p\to 0^-.$$

From Appendix D,

$$\lim_{p\to 0^-}\varphi\left(p,\widehat{X}_t,s\right)=\exp\left(\overline{\ln\hat{X}}+\frac{s^2}{2}\right)\left\{-\overline{\ln\hat{X}}+\ln\hat{X}_t\right\}.$$

Therefore, as $p \to 0^-$, (B3) becomes

$$\widehat{X}_{c,t} = w\widehat{X}_t + (1-w)\left[\overline{X} + \exp\left(\overline{\ln\hat{X}} + \frac{s^2}{2}\right)\left\{-\overline{\ln\hat{X}} + \ln\hat{X}_t\right\}\right]. \tag{B4}$$

The combined mean square error MSEc $= \mathrm{Var}\left(\widehat{X}_{c,t} - X_t\right)$. Substituting for $\widehat{X}_{c,t}$ from (B4)

$$\mathrm{MSEc}\left(\sigma^2\right) = \mathrm{Var}\left[w\widehat{X}_t + (1-w)\left\{\overline{X} + \exp\left(\overline{\ln\hat{X}} + \frac{s^2}{2}\right)\left(-\overline{\ln\hat{X}} + \ln\hat{X}_t\right)\right\} - X_t\right]$$

where s^2 is the variance of the logarithm $\ln\widehat{X}_t$ and $w = \left\{1 - \frac{\mathrm{Cov}(\varepsilon_t, X_{t-1})}{\mathrm{Var}(X_t)}\right\}^{-1}$ (Ridley [29]).

Substituting for $\widehat{X}_t$ from (B1)

$$\begin{aligned}\mathrm{MSEc}\left(\sigma^2\right) =& \mathrm{Var}\left[w\left\{\mu_X + \Phi\left(X_{t-1} - \mu_X\right) + \left\{\frac{\mathrm{Cov}\left(\varepsilon_t, X_{t-1}\right)}{\mathrm{Var}\left(X_t\right)}\right\}\left(X_{t-1} - \mu_X\right)\right\}\right.\\ &\left. + (1-w)\left\{\overline{X} + \exp\left(\overline{\ln\hat{X}} + \frac{s^2}{2}\right)\left(-\overline{\ln\hat{X}} + \ln\hat{X}_t\right)\right\} - X_t\right]\end{aligned}$$

Since $\overline{X} \to \mu_X$ as $n \to \infty$ and μ_X, Φ, and $\frac{\mathrm{Cov}(\varepsilon_t, X_{t-1})}{\mathrm{Var}(X_t)}$ are constant for a specified time series and model, then the post combining MSEc, is given by

$$\begin{aligned}\mathrm{MSEc}\left(\sigma^2\right) =& \mathrm{Var}\left[w\left\{\Phi X_{t-1} + \frac{\mathrm{Cov}\left(\varepsilon_t, X_{t-1}\right)}{\mathrm{Var}\left(X_t\right)} X_{t-1}\right\} + (1-w)\exp\left(\overline{\ln\hat{X}} + \frac{s^2}{2}\right)\ln\hat{X}_t - X_t\right]\\ =& w^2\left\{\Phi + \frac{\mathrm{Cov}\left(\varepsilon_t, X_{t-1}\right)}{\mathrm{Var}\left(X_t\right)}\right\}^2 \mathrm{Var}\left(X_t\right) + (1-w)^2 \exp\left(2\overline{\ln\hat{X}} + s^2\right)\mathrm{Var}\left(\ln\hat{X}_t\right) + \mathrm{Var}\left(X_t\right)\\ &+ 2w(1-w)\left\{\Phi + \frac{\mathrm{Cov}\left(\varepsilon_t, X_{t-1}\right)}{\mathrm{Var}\left(X_t\right)}\right\}\exp\left(\overline{\ln\hat{X}} + \frac{s^2}{2}\right)\mathrm{Cov}\left(X_{t-1}, \ln\hat{X}_t\right) \qquad \text{(B5)}\\ &- 2w\left\{\Phi + \frac{\mathrm{Cov}\left(\varepsilon_t, X_{t-1}\right)}{\mathrm{Var}\left(X_t\right)}\right\}\mathrm{Cov}\left(X_t, X_{t-1}\right) - 2(1-w)\exp\left(\overline{\ln\hat{X}} + \frac{s^2}{2}\right)\mathrm{Cov}\left(X_t, \ln\hat{X}_t\right)\end{aligned}$$

From (1) $\mathrm{Cov}\left(X_t, X_{t-1}\right) = \left\{\Phi\mathrm{Var}\left(X_t\right) + \mathrm{Cov}\left(\varepsilon_t, X_{t-1}\right)\right\} = \left\{\Phi + \frac{\mathrm{Cov}(\varepsilon_t, X_{t-1})}{\mathrm{Var}(X_t)}\right\}\mathrm{Var}\left(X_t\right)$. From the Ridley [29] antithetic time series theorem, $\rho_{\ln\hat{X}_t, X_t} = \frac{\sigma}{\sqrt{\exp(\sigma^2) - 1}} = \frac{\mathrm{Cov}(\ln X_t, X_t)}{\sqrt{\mathrm{Var}(\ln X_t)}\sqrt{\mathrm{Var}(X_t)}}$, and for a stationary time series, as $n \to \infty$, $\ln\hat{X}_t \to \ln X_t$, $\mathrm{Var}\left(\ln\hat{X}_t\right) = s^2$, $\mathrm{Var}\left(\hat{X}_t\right) = \sigma^2$ and $s^2 \to \sigma^2$. So,

$$\mathrm{Cov}\left(\ln\hat{X}_t, X_t\right) = \frac{\sigma}{\sqrt{\exp(\sigma^2) - 1}}\sqrt{\mathrm{Var}\left(\ln X_t\right)}\sqrt{\mathrm{Var}\left(X_t\right)}$$

$$=\frac{\sigma^2/\sqrt{\mathrm{Var}\left(X_t\right)}}{\sqrt{\exp(\sigma^2)-1}}\mathrm{Var}\left(X_t\right).$$

Substituting for $\mathrm{Cov}\left(X_{t-1},X_t\right)$ and $\mathrm{Cov}\left(\ln\hat{X}_t,X_t\right)$ in (B5),

$$\begin{aligned}
\mathrm{MSEc}\left(\sigma^2\right) =& w^2\left\{\Phi+\frac{\mathrm{Cov}\left(\varepsilon_t,X_{t-1}\right)}{\mathrm{Var}\left(X_t\right)}\right\}^2\mathrm{Var}\left(X_t\right)+(1-w)^2\exp\left(2\overline{\ln\hat{X}}+s^2\right)\mathrm{Var}\left(\ln\hat{X}_t\right)+\mathrm{Var}\left(X_t\right)\\
&+2w(1-w)\left\{\Phi+\frac{\mathrm{Cov}\left(\varepsilon_t,X_{t-1}\right)}{\mathrm{Var}\left(X_t\right)}\right\}\exp\left(\overline{\ln\hat{X}}+\frac{s^2}{2}\right)\mathrm{Cov}\left(X_{t-1},\ln\hat{X}_t\right)\\
&-2w\left\{\Phi+\frac{\mathrm{Cov}\left(\varepsilon_t,X_{t-1}\right)}{\mathrm{Var}\left(X_t\right)}\right\}\left\{\Phi+\frac{\mathrm{Cov}\left(\varepsilon_t,X_{t-1}\right)}{\mathrm{Var}\left(X_t\right)}\right\}\mathrm{Var}\left(X_t\right)-2(1-w)\exp\left(\overline{\ln\hat{X}}+\frac{s^2}{2}\right)\\
&\times\frac{\sigma^2/\sqrt{\mathrm{Var}\left(X_t\right)}}{\sqrt{\exp(\sigma^2)-1}}\mathrm{Var}\left(X_t\right)\\
=& w^2\left\{\Phi+\frac{\mathrm{Cov}\left(\varepsilon_t,X_{t-1}\right)}{\mathrm{Var}\left(X_t\right)}\right\}^2\mathrm{Var}\left(X_t\right)+(1-w)^2\exp\left(2\overline{\ln\hat{X}}+s^2\right)\mathrm{Var}\left(\ln\hat{X}_t\right)+\mathrm{Var}\left(X_t\right)\\
&+2w(1-w)\left\{\Phi+\frac{\mathrm{Cov}\left(\varepsilon_t,X_{t-1}\right)}{\mathrm{Var}\left(X_t\right)}\right\}\exp\left(\overline{\ln\hat{X}}+\frac{s^2}{2}\right)\mathrm{Cov}\left(X_{t-1},\ln\hat{X}_t\right) \qquad \text{(B6)}\\
&-2w\left\{\Phi+\frac{\mathrm{Cov}\left(\varepsilon_t,X_{t-1}\right)}{\mathrm{Var}\left(X_t\right)}\right\}^2\mathrm{Var}\left(X_t\right)-2(1-w)\exp\left(\overline{\ln\hat{X}}+\frac{s^2}{2}\right)\frac{\sigma^2/\sqrt{\mathrm{Var}\left(X_t\right)}}{\sqrt{\exp(\sigma^2)-1}}\mathrm{Var}\left(X_t\right)\\
=& \Bigg[\left(w^2-2w\right)\left\{\Phi+\frac{\mathrm{Cov}\left(\varepsilon_t,X_{t-1}\right)}{\mathrm{Var}\left(X_t\right)}\right\}^2+1-2(1-w)\exp\left(\overline{\ln\hat{X}}+\frac{s^2}{2}\right)\frac{\sigma^2/\sqrt{\mathrm{Var}\left(X_t\right)}}{\sqrt{\exp(\sigma^2)-1}}\\
&+(1-w)^2\exp\left(2\overline{\ln\hat{X}}+s^2\right)\frac{\mathrm{Var}\left(\ln\hat{X}_t\right)}{\mathrm{Var}\left(X_t\right)}+2w(1-w)\left\{\Phi+\frac{\mathrm{Cov}\left(\varepsilon_t,X_{t-1}\right)}{\mathrm{Var}\left(X_t\right)}\right\}\\
&\times\exp\left(\overline{\ln\hat{X}}+\frac{s^2}{2}\right)\frac{\mathrm{Cov}\left(\ln\hat{X}_t,X_{t-1}\right)}{\mathrm{Var}\left(X_t\right)}\Bigg]\mathrm{Var}\left(X_t\right)
\end{aligned}$$

Appendix C: Per Unit Reduction in MSE

Denoting the per unit reduction in MSE due to combining as

$$\mathrm{PMSEc}\left(\sigma^2\right)=\frac{\mathrm{MSE}\left(\sigma^2\right)-\mathrm{MSEc}\left(\sigma^2\right)}{\mathrm{MSE}\left(\sigma^2\right)}$$

Substituting for $\mathrm{MSE}\left(\sigma^2\right)$ and $\mathrm{MSEc}\left(\sigma^2\right)$ from Appendix A and Appendix B.

$$\mathrm{MSE}\left(\sigma^2\right)=\left[1-\left\{\Phi+\frac{\mathrm{Cov}\left(\varepsilon_t,X_{t-1}\right)}{\mathrm{Var}\left(X_t\right)}\right\}^2\right]\mathrm{Var}\left(X_t\right)$$

$$\begin{aligned}\mathrm{MSEc}\left(\sigma^2\right)=&\left[\left(w^2-2w\right)\left\{\Phi+\frac{\mathrm{Cov}\left(\varepsilon_t,X_{t-1}\right)}{\mathrm{Var}\left(X_t\right)}\right\}^2+1-2\left(1-w\right)\exp\left(\overline{\ln\hat{X}}+\frac{s^2}{2}\right)\frac{\sigma^2/\sqrt{\mathrm{Var}\left(X_t\right)}}{\sqrt{\exp(\sigma^2)-1}}\right.\\&+\left(1-w\right)^2\exp\left(2\overline{\ln\hat{X}}+s^2\right)\frac{\mathrm{Var}\left(\ln\hat{X}_t\right)}{\mathrm{Var}\left(X_t\right)}+2w\left(1-w\right)\left\{\Phi+\frac{\mathrm{Cov}\left(\varepsilon_t,X_{t-1}\right)}{\mathrm{Var}\left(X_t\right)}\right\}\\&\left.\times\exp\left(\overline{\ln\hat{X}}+\frac{s^2}{2}\right)\frac{\mathrm{Cov}\left(\ln\hat{X}_t,X_{t-1}\right)}{\mathrm{Var}\left(X_t\right)}\right]\mathrm{Var}\left(X_t\right)\end{aligned}$$

From which

$$\begin{aligned}\mathrm{MSE}\left(\sigma^2\right)-\mathrm{MSEc}\left(\sigma^2\right)=&\left[1-\left\{\Phi+\frac{\mathrm{Cov}\left(\varepsilon_t,X_{t-1}\right)}{\mathrm{Var}\left(X_t\right)}\right\}^2\right]\mathrm{Var}\left(X_t\right)-\left[\left(w^2-2w\right)\right.\\&\times\left\{\Phi+\frac{\mathrm{Cov}\left(\varepsilon_t,X_{t-1}\right)}{\mathrm{Var}\left(X_t\right)}\right\}^2+1-2\left(1-w\right)\exp\left(\overline{\ln\hat{X}}+\frac{s^2}{2}\right)\frac{\sigma^2/\sqrt{\mathrm{Var}\left(X_t\right)}}{\sqrt{\exp(\sigma^2)-1}}\\&+\left(1-w\right)^2\exp\left(2\overline{\ln\hat{X}}+s^2\right)\frac{\mathrm{Var}\left(\ln\hat{X}_t\right)}{\mathrm{Var}\left(X_t\right)}+2w\left(1-w\right)\left\{\Phi+\frac{\mathrm{Cov}\left(\varepsilon_t,X_{t-1}\right)}{\mathrm{Var}\left(X_t\right)}\right\}\\&\left.\times\exp\left(\overline{\ln\hat{X}}+\frac{s^2}{2}\right)\frac{\mathrm{Cov}\left(\ln\hat{X}_t,X_{t-1}\right)}{\mathrm{Var}\left(X_t\right)}\right]\mathrm{Var}\left(X_t\right)\\=&\left[-\left(w^2-2w+1\right)\left\{\Phi+\frac{\mathrm{Cov}\left(\varepsilon_t,X_{t-1}\right)}{\mathrm{Var}\left(X_t\right)}\right\}^2+2\left(1-w\right)\exp\left(\overline{\ln\hat{X}}+\frac{s^2}{2}\right)\frac{\sigma^2/\sqrt{\mathrm{Var}\left(X_t\right)}}{\sqrt{\exp(\sigma^2)-1}}\right.\\&-\left(1-w\right)^2\exp\left(2\overline{\ln\hat{X}}+s^2\right)\frac{\mathrm{Var}\left(\ln\hat{X}_t\right)}{\mathrm{Var}\left(X_t\right)}-2w\left(1-w\right)\left\{\Phi+\frac{\mathrm{Cov}\left(\varepsilon_t,X_{t-1}\right)}{\mathrm{Var}\left(X_t\right)}\right\}\\&\left.\times\exp\left(\overline{\ln\hat{X}}+\frac{s^2}{2}\right)\frac{\mathrm{Cov}\left(\ln\hat{X}_t,X_{t-1}\right)}{\mathrm{Var}\left(X_t\right)}\right]\mathrm{Var}\left(X_t\right)\\=&\left[-(w-1)^2\left\{\Phi+\frac{\mathrm{Cov}\left(\varepsilon_t,X_{t-1}\right)}{\mathrm{Var}\left(X_t\right)}\right\}^2+2\left(1-w\right)\exp\left(\overline{\ln\hat{X}}+\frac{s^2}{2}\right)\frac{\sigma^2/\sqrt{\mathrm{Var}\left(X_t\right)}}{\sqrt{\exp(\sigma^2)-1}}\right.\\&-\left(1-w\right)^2\exp\left(2\overline{\ln\hat{X}}+s^2\right)\frac{\mathrm{Var}\left(\ln\hat{X}_t\right)}{\mathrm{Var}\left(X_t\right)}-2w\left(1-w\right)\left\{\Phi+\frac{\mathrm{Cov}\left(\varepsilon_t,X_{t-1}\right)}{\mathrm{Var}\left(X_t\right)}\right\}\end{aligned}$$

$$\times\ \exp\left(\overline{\ln\hat{X}}+\frac{s^2}{2}\right)\frac{\mathrm{Cov}\left(\ln\hat{X}_t,X_{t-1}\right)}{\mathrm{Var}\left(X_t\right)}\Bigg]\mathrm{Var}\left(X_t\right).$$

Substituting and canceling $\mathrm{Var}\left(X_t\right)$,

$$\mathrm{PMSEc}\left(\sigma^2\right)=\frac{A_1}{1-\left\{\Phi+\frac{\mathrm{Cov}(\varepsilon_t,X_{t-1})}{\mathrm{Var}(X_t)}\right\}^2},$$

where

$$\begin{aligned}A_1=&-(w-1)^2\left\{\Phi+\frac{\mathrm{Cov}\left(\varepsilon_t,X_{t-1}\right)}{\mathrm{Var}\left(X_t\right)}\right\}^2+2(1-w)\exp\left(\overline{\ln\hat{X}}+\frac{s^2}{2}\right)\frac{\sigma^2/\sqrt{\mathrm{Var}\left(X_t\right)}}{\sqrt{\exp(\sigma^2)-1}}\\&-(1-w)^2\exp\left(2\overline{\ln\hat{X}}+s^2\right)\frac{\mathrm{Var}\left(\ln\hat{X}_t\right)}{\mathrm{Var}\left(X_t\right)}-2w(1-w)\left\{\Phi+\frac{\mathrm{Cov}\left(\varepsilon_t,X_{t-1}\right)}{\mathrm{Var}\left(X_t\right)}\right\}\\&\times\exp\left(\overline{\ln\hat{X}}+\frac{s^2}{2}\right)\frac{\mathrm{Cov}\left(\ln\hat{X}_t,X_{t-1}\right)}{\mathrm{Var}\left(X_t\right)}.\end{aligned}$$

Since $\overline{\ln\hat{X}}\to\mu$ and $s\to\sigma$, as $n\to\infty$ then writing $\sqrt{\mathrm{Var}\left(X_t\right)}$, $\mathrm{Var}\left(X_t\right)$ and $\mathrm{Var}\left(\ln\widehat{X}_t\right)$ in terms of σ and simplifying,

$$\mathrm{PMSEc}\left(\sigma^2\right)=\frac{A_2}{1-\left\{\Phi+\frac{\mathrm{Cov}(\varepsilon_t,X_{t-1})}{\mathrm{Var}(X_t)}\right\}^2},$$

$$\begin{aligned}A_2=&-(w-1)^2\left\{\Phi+\frac{\mathrm{Cov}\left(\varepsilon_t,X_{t-1}\right)}{\mathrm{Var}\left(X_t\right)}\right\}^2+2(1-w)\exp\left(\mu+\frac{\sigma^2}{2}\right)\\&\times\frac{\sigma^2}{\sqrt{\exp(\sigma^2)-1}\exp\left(\mu+\frac{\sigma^2}{2}\right)\sqrt{\exp(\sigma^2)-1}}\\&-(1-w)^2\exp\left(2\mu+\sigma^2\right)\frac{\sigma^2}{\exp\left(2\mu+\sigma^2\right)\left(\exp(\sigma^2)-1\right)}\\&-2w(1-w)\left\{\Phi+\frac{\mathrm{Cov}\left(\varepsilon_t,X_{t-1}\right)}{\mathrm{Var}\left(X_t\right)}\right\}\exp\left(\mu+\frac{\sigma^2}{2}\right)\\&\times\frac{\mathrm{Cov}\left(\ln\hat{X}_t,X_{t-1}\right)}{\exp\left(2\mu+\sigma^2\right)\left(\exp(\sigma^2)-1\right)}.\end{aligned}$$

$$\text{PMSEc}\left(\sigma^2\right) = \frac{A_3}{1 - \left\{\Phi + \frac{\text{Cov}(\varepsilon_t, X_{t-1})}{\text{Var}(X_t)}\right\}^2},$$

$$A_3 = -(w-1)^2\left\{\Phi + \frac{\text{Cov}\left(\varepsilon_t, X_{t-1}\right)}{\text{Var}\left(X_t\right)}\right\}^2 + \left(1-w^2\right)\frac{\sigma^2}{\left(\exp(\sigma^2)-1\right)}$$
$$-2w(1-w)\left\{\Phi + \frac{\text{Cov}\left(\varepsilon_t, X_{t-1}\right)}{\text{Var}\left(X_t\right)}\right\}\frac{\text{Cov}\left(\ln\hat{X}_t, X_{t-1}\right)}{\exp\left(\mu+\frac{\sigma^2}{2}\right)\left(\exp(\sigma^2)-1\right)}.$$

Substituting for w

$$\text{PMSEc}\left(\sigma^2\right) = \frac{A+B}{1 - \left\{\Phi + \frac{\text{Cov}(\varepsilon_t, X_{t-1})}{\text{Var}(X_t)}\right\}^2}, \tag{C1}$$

where

$$A = -\left[\left\{1 - \frac{\text{Cov}\left(\varepsilon_t, X_{t-1}\right)}{\text{Var}\left(X_t\right)}\right\}^{-1} - 1\right]^2 \left\{\Phi + \frac{\text{Cov}\left(\varepsilon_t, X_{t-1}\right)}{\text{Var}\left(X_t\right)}\right\}^2 + \left[1 - \left\{1 - \frac{\text{Cov}\left(\varepsilon_t, X_{t-1}\right)}{\text{Var}\left(X_t\right)}\right\}^{-2}\right]$$
$$\times\left[\frac{\sigma^2}{\left(\exp(\sigma^2)-1\right)}\right]$$

and

$$B = -2\left\{1 - \frac{\text{Cov}\left(\varepsilon_t, X_{t-1}\right)}{\text{Var}\left(X_t\right)}\right\}^{-1}\left[1 - \left\{1 - \frac{\text{Cov}\left(\varepsilon_t, X_{t-1}\right)}{\text{Var}\left(X_t\right)}\right\}^{-1}\right]\left\{\Phi + \frac{\text{Cov}\left(\varepsilon_t, X_{t-1}\right)}{\text{Var}\left(X_t\right)}\right\}$$
$$\times\frac{\text{Cov}\left(X_{t-1}, \ln\widehat{X}_t\right)}{\exp\left\{\mu+\frac{\sigma^2}{2}\right\}\left(\exp(\sigma^2)-1\right)}$$

and $\frac{\text{Cov}(\varepsilon_t, X_{t-1})}{\exp\{2\mu+\sigma^2\}(\exp(\sigma^2)-1)} \neq 1$.

Substituting for $\text{Var}\left(X_t\right)$

$$\text{PMSEc}\left(\sigma^2\right) = \frac{C+D}{1 - \left\{\Phi + \frac{\text{Cov}(\varepsilon_t, X_{t-1})}{\exp\{2\mu+\sigma^2\}(\exp(\sigma^2)-1)}\right\}^2}, \tag{C2}$$

where

$$C = -\left[\left\{1 - \frac{\text{Cov}\left(\varepsilon_t, X_{t-1}\right)}{\exp\left\{2\mu+\sigma^2\right\}\left(\exp(\sigma^2)-1\right)}\right\}^{-1} - 1\right]^2\left\{\Phi + \frac{\text{Cov}\left(\varepsilon_t, X_{t-1}\right)}{\exp\left\{2\mu+\sigma^2\right\}\left(\exp(\sigma^2)-1\right)}\right\}^2$$

$$+\left[1-\left\{1-\frac{\mathrm{Cov}\left(\varepsilon_t,X_{t-1}\right)}{\exp\left\{2\mu+\sigma^2\right\}\left(\exp(\sigma^2)-1\right)}\right\}^{-2}\right]\frac{\sigma^2}{\left(\exp(\sigma^2)-1\right)}$$

$$D=-2\left\{1-\frac{\mathrm{Cov}\left(\varepsilon_t,X_{t-1}\right)}{\exp\left\{2\mu+\sigma^2\right\}\left(\exp(\sigma^2)-1\right)}\right\}^{-1}\left[1-\left\{1-\frac{\mathrm{Cov}\left(\varepsilon_t,X_{t-1}\right)}{\exp\left\{2\mu+\sigma^2\right\}\left(\exp(\sigma^2)-1\right)}\right\}^{-1}\right]$$
$$\times\left\{\Phi+\frac{\mathrm{Cov}\left(\varepsilon_t,X_{t-1}\right)}{\exp\left\{2\mu+\sigma^2\right\}\left(\exp(\sigma^2)-1\right)}\right\}\frac{\mathrm{Cov}\left(\varepsilon_t,X_{t-1}\right)}{\exp\left\{\mu+\frac{\sigma^2}{2}\right\}\left(\exp(\sigma^2)-1\right)}.$$

Appendix D: Limit as $p \to 0^-$

We wish to find

$$\lim_{p\to 0^-,s\to 0}\varphi\left(p,\widehat{X}_t,s\right)=\lim_{p\to 0^-}r_{\widehat{z}\widehat{z}^p}\frac{\exp\left\{\overline{\ln\widehat{X}}+s^2/2\right\}\sqrt{\exp(s^2)-1}}{\exp\left\{p\overline{\ln\widehat{X}}+p^2s^2/2\right\}\sqrt{\exp\left(p^2s^2\right)-1}}$$
$$\times\left[\widehat{X}_t^p-\exp\left(p\cdot\overline{\ln\widehat{X}}+p^2s^2/2\right)\right].$$

Consider the Taylor expansion at $p=0$ of

$$\exp(p)=1+p+\frac{p^2}{2}+\frac{p^3}{6}+\frac{p^4}{24}+o\left(p^4\right). \tag{D1}$$

Using (D1), we derive the Taylor expansion at $p=0$ of $\exp(p^2s^2)-1=p^2s^2+\frac{p^4s^4}{2}+\frac{p^6s^6}{6}+o\left(p^6\right)$. Hence,

$$\frac{1}{\sqrt{\exp\left(p^2s^2\right)-1}}=\frac{1}{\sqrt{p^2s^2+\frac{p^4s^4}{2}+\frac{p^6s^6}{6}+o\left(p^6\right)}}=\frac{1}{s|p|\sqrt{1+\frac{p^2s^2}{2}+\frac{p^4s^4}{6}+o\left(p^4\right)}} \tag{D2}$$

Using equation (D1), we derive the Taylor expansion at $p=0$ of

$$\begin{aligned}\exp\left(p\cdot\overline{\ln\widehat{X}}+p^2s^2/2\right)&=1+p\cdot\overline{\ln\widehat{X}}+p^2s^2/2+\frac{1}{2}\left\{p\cdot\overline{\ln\widehat{X}}+p^2s^2/2\right\}^2+o\left(p^2\right)\\&=1+p\cdot\overline{\ln\widehat{X}}+p^2s^2/2+\frac{p^2}{2}\cdot\overline{\ln\widehat{X}}^2+o\left(p^2\right)\end{aligned} \tag{D3}$$

The Taylor expansion at $p=0$ of $\left(\widehat{X}_t\right)^p$ is

$$\widehat{X}_t^p=1+p\cdot\ln\widehat{X}+\frac{p^2}{2}\cdot\ln^2\widehat{X}+o\left(p^2\right) \tag{D4}$$

By subtracting (D3) from (D4), we obtain the Taylor expansion at $p = 0$ of

$$\left\{\widehat{X}_t^p - \exp\left(p \cdot \overline{\ln\widehat{X}} + p^2 s^2/2\right)\right\} = p\left(-\overline{\ln\widehat{X}} + \ln\widehat{X}_t\right) + \frac{p^2}{2}\left(-\overline{\ln\widehat{X}}^2 - s^2 + \ln^2\widehat{X}_t\right) + o\left(p^2\right) \tag{D5}$$

The Taylor expansion of $\varphi\left(p, \widehat{X}_t, s\right)$ at $p = 0$ may now be obtained by multiplication of $r_{\widehat{X}\widehat{X}^p}\frac{\exp\left\{\overline{\ln\widehat{X}} + s^2/2\right\}\sqrt{\exp(s^2)-1}}{\exp\left\{p \cdot \overline{\ln\widehat{X}} + p^2 s^2/2\right\}}$ and the expansions (D2) and (D5) as follows.

$$\varphi\left(p, \widehat{X}_t, s\right) = r_{\widehat{X}\widehat{X}^p}\frac{\exp\left(\overline{\ln\widehat{X}} + s^2/2\right)\sqrt{\exp\left(s^2\right)-1}}{\exp\left(p \cdot \overline{\ln\widehat{X}} + p^2 s^2/2\right)} \cdot \frac{1}{s|p|\sqrt{1 + \frac{p^2 s^2}{2} + \frac{p^4 s^4}{6} + o\left(p^4\right)}}$$
$$\times\left[p\left(-\overline{\ln\widehat{X}} + \ln\widehat{X}_t\right) + \frac{p^2}{2}\left(-\overline{\ln\widehat{X}}^2 - s^2 + \ln^2\widehat{X}_t\right) + o\left(p^2\right)\right].$$

From Ridley [29] antithetic time series theorem, $r_{\widehat{X}\widehat{X}^p} \downarrow -\frac{s}{\sqrt{\exp(s^2)-1}}$ as $p \downarrow 0^-$. Also, $\lim_{p\to 0^-}\frac{p}{|p|} = -1$, and, therefore,

$$\lim_{p\to 0^-,}\varphi\left(p, \widehat{X}_t, s\right) = \exp\left(\overline{\ln\widehat{X}} + \frac{s^2}{2}\right)\left(-\overline{\ln\widehat{X}} + \ln\widehat{X}_t\right).$$