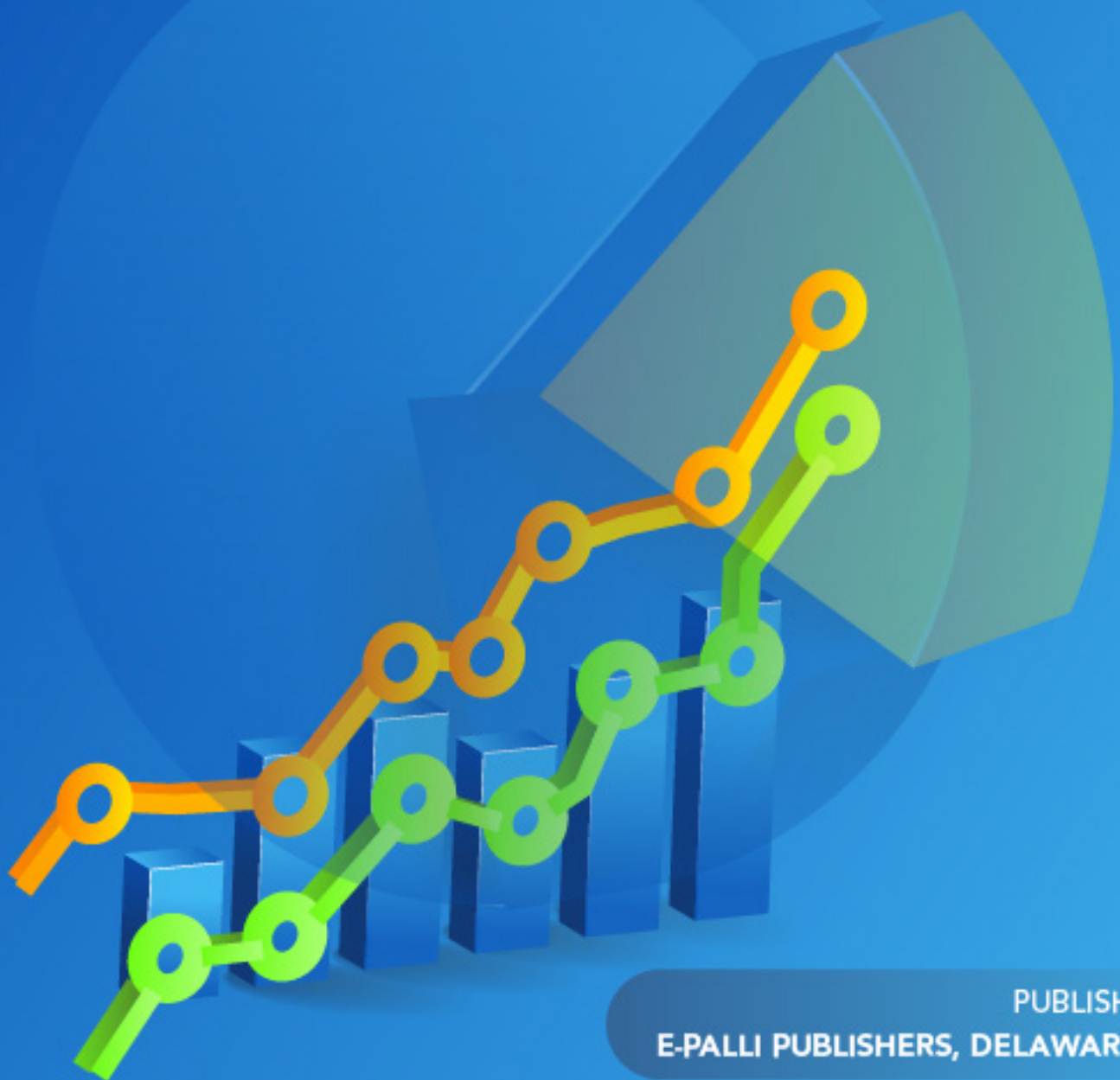




American Journal of Applied Statistics and Economics (AJASE)

ISSN: 2992-927X (ONLINE)

VOLUME 5 ISSUE 2 (2026)



PUBLISHED BY
E-PALLI PUBLISHERS, DELAWARE, USA

Assessment of the Relative Efficiency of Bootstrap Resampling Methods for Forecast Uncertainty Quantification in Nigerian Macroeconomic Time Series: Evidence from Inflation and Exchange Rate Data (1960-2025)

Okafor, Joseph Iweanandu^{1*}, Kerry C.C¹, Amagoh Maureen N²

Article Information

Received: July 29, 2026

Accepted: October 06, 2026

Published: October 17, 2026

Keywords

Arima, Bootstrap Resampling, Monte Carlo Simulation, Moving Block Bootstrap, Nigeria Sieve Bootstrap, Stationary Bootstrap, Time Series Forecasting, Uncertainty Quantification

ABSTRACT

Quantifying forecast uncertainty is a prerequisite for responsible use of time series models in Nigerian macroeconomic analysis, yet researchers rarely benchmark resampling schemes used to estimate that uncertainty against alternatives on long-run Nigerian data. This study compares three bootstrap procedures for response time series the stationary bootstrap, the moving block bootstrap, and the sieve bootstrap across five ARIMA data-generating processes of increasing complexity: ARIMA(1,0,0), ARIMA(1,1,0), ARIMA(1,1,1), ARIMA(2,1,1), and ARIMA(1,1,3). The simulation design follows a Monte Carlo framework comparable to earlier benchmarking studies. The same three methods are then applied to two real Nigerian macroeconomic series: annual headline inflation (World Bank/FRED, 1960-2025, 66 observations) and the log of the annual NGN/USD exchange rate (Penn World Table/FRED, 1960-2023, 64 observations), both considerably longer than the short annual series used in prior comparative work of this kind. Across all five simulated data-generating processes and both empirical series, the stationary bootstrap yielded the lowest mean squared error (MSE) of one-step-ahead forecasts; the moving block bootstrap was a close second throughout; the sieve bootstrap was markedly less accurate, particularly for integrated I(1) processes. These results diverge from prior work suggesting block-based methods are dominated by ARMA processes with genuine moving-average components, and from work favouring the sieve bootstrap for autoregressive-dominated series. Possible explanations are discussed, including the sensitivity of the sieve reconstruction step to compounding during re-integration of a differenced series. The scope limitations of a reduced-scale Monte Carlo design are noted, and further replication is identified as necessary before firm practical guidance can be issued to Nigerian forecasting practitioners.

INTRODUCTION

ARIMA-based forecasting systems are only as reliable as the uncertainty estimates attached to their point forecasts. This is particularly consequential for Nigerian macroeconomic series such as the headline inflation rate and the Naira/US Dollar exchange rate, both of which have exhibited substantial volatility over the past decade and are routinely forecast by banks, government agencies, and AI-assisted decision-support tools. The bootstrap resampling a dataset or a model fitted to it, in order to approximate the sampling distribution of a statistic is one of the most widely used tools for this purpose. Its classical form, however, assumes independently and identically distributed observations, an assumption that is violated by construction in time series data, where successive observations are serially correlated.

A substantial methodological literature has developed resampling schemes adapted to dependent data, including the moving block bootstrap, the stationary bootstrap, and the sieve bootstrap. Direct, reproducible comparisons of these methods on Nigerian macroeconomic data remain scarce. Most existing Nigerian applications adopt a single bootstrap method without justification or restrict comparison to very short annual series. This study

addresses that gap by benchmarking three bootstrap schemes stationary, block, and sieve under a controlled Monte Carlo design spanning five ARIMA specifications, and then applying the same methods to two long-run Nigerian macroeconomic series: headline inflation and the NGN/USD exchange rate.

Research Objectives

(i) To evaluate, via Monte Carlo simulation, the bias and MSE of forecasts produced by the stationary, block, and sieve bootstrap methods across ARIMA processes of varying order. (ii) To apply all three methods to two real, long-run Nigerian macroeconomic time series headline inflation and the NGN/USD exchange rate. (iii) To assess whether any one method offers uniformly superior performance for Nigerian macroeconomic data, and to characterise the conditions under which each method performs best.

LITERATURE REVIEW

Efron's classical bootstrap resamples independent and identically distributed observations with replacement and is inconsistent under serial dependence. Künsch (1989) and Liu and Singh (1992) introduced the moving

¹ Department of Mathematics & Statistics, Delta State Polytechnic, Ogwashi-Uku, Delta, Nigeria

³ Department of Statistics, Delta State Polytechnic, Ogwashi-Uku, Delta, Nigeria

* Corresponding author's e-mail: josephiokafor@yahoo.com

block bootstrap, which resamples contiguous blocks of the series to preserve local dependence; block-length selection remains a well-documented bias–variance trade-off. Politis and Romano (1994) proposed the stationary bootstrap, which randomizes block length according to a geometric distribution, producing a resampled series that is itself stationary, a property that the fixed-length block bootstrap does not possess. Bühlmann (1997) proposed the sieve bootstrap, which fits a finite-order autoregression to the series, resamples the estimated innovations, and regenerates a new series through the fitted AR recursion; its performance depends heavily on how well a finite autoregressive sieve approximates the true, possibly infinite-order, dependence structure of the series.

The consensus in this literature, reflected in Lahiri (1999) and subsequent work on bootstrap prediction and forecasting, is that no single method dominates uniformly; relative performance depends on the true dependence structure, sample size, and the statistic being bootstrapped. Related forecasting research has also examined resampling and ensemble approaches to improve forecast performance (Bergmeir et al., 2016; Pan & Politis, 2016). In the Nigerian context, studies have examined long-memory characteristics of selected macroeconomic variables using ARFIMA and ARFIMA-FIGARCH models (Adewole, 2024), highlighting the importance of accounting for dependence structures when modelling Nigerian macroeconomic time series.

Recent empirical studies published by E-Palli further demonstrate the continued relevance of time-series methods for Nigerian macroeconomic variables. Oni and Akinsanya (2026) examined the effects of inflation, interest rate, and exchange rate on the Nigerian economy using annual time-series data from 1990 to 2023. Similarly, Udoudo et al. (2023) investigated the impact of liquefied natural gas exports on the Nigerian exchange rate using an autoregressive distributed lag framework. These studies reinforce the importance of appropriate statistical modelling for Nigerian macroeconomic time series, particularly variables such as inflation and exchange rates. On Nigerian data specifically, bootstrap-based interval estimation has been applied to inflation and exchange-rate forecasting in several dissertations and working papers, but almost always with a single method chosen a priori rather than benchmarked against alternatives, and typically on short samples. This motivates a direct, multi-model simulation comparison applied to long-run Nigerian series, rather than relying on results from a single data-generating process or a single short series.

MATERIALS AND METHODS

Data-Generating Processes

Five ARIMA(p,d,q) specifications were simulated in order of increasing complexity: ARIMA(1,0,0) with $\varphi = 0.6$; ARIMA(1,1,0) with $\varphi = 0.5$; ARIMA(1,1,1) with $\varphi = 0.5$, $\theta = 0.3$; ARIMA(2,1,1) with $\varphi_1 = 0.4$, $\varphi_2 = 0.2$, $\theta = 0.3$; and ARIMA(1,1,3) with $\varphi = 0.5$, $\theta_1 = 0.3$, $\theta_2 = -0.2$,

$\theta_3 = 0.1$. Innovations were drawn i.i.d. $N(0,1)$. Series of length $n = 150$ (plus a 200-period burn-in, discarded) were generated for each Monte Carlo replicate.

Estimation

ARMA(p,q) parameters were estimated by maximum likelihood via the Kalman filter in state-space form, implemented in Python using the statsmodels library. One-step-ahead forecasts were recovered from each fitted model.

Bootstrap Methods

Stationary bootstrap: resampled series were constructed by concatenating blocks of the original (differenced, where applicable) series whose lengths follow a geometric distribution with mean block length $L = 10$ in the simulation study ($L = 5$ for the real series), restarting from a new random position after each block. Moving block bootstrap: resampled series were constructed from fixed-length blocks of the same length, drawn with replacement from the original series. Sieve bootstrap: an AR(k) model (with k selected by AIC, $k \leq 5$) was fitted once per Monte Carlo replicate or once for each real series; residuals were resampled i.i.d. and a new series was regenerated through the fitted AR recursion. For integrated ($d = 1$) processes, resampling was applied to the first-differenced series, and the resampled increments were re-integrated (cumulatively summed from the first observation of the original series) to reconstruct a bootstrap-resampled level series.

For each replicate and method, the ARMA model was refitted to the bootstrap-resampled series and a one-step-ahead forecast $\hat{\theta}^*$ was recorded. Bootstrap bias was computed as the mean of $\hat{\theta}^*$ across B bootstrap replicates minus the point forecast $\hat{\theta}$ from the original replicate; bootstrap MSE was computed as the mean squared deviation of $\hat{\theta}^*$ from $\hat{\theta}$. In the simulation study, these quantities were averaged across 15 Monte Carlo replications per model, with $B = 60$ bootstrap replicates per replicate a reduced-scale design constrained by computational budget (see Section 6). In the real-data application, $B = 1,000$ bootstrap replicates were drawn for each series and each method.

Data

Two real Nigerian macroeconomic series were used, both drawn from secondary sources. Annual headline inflation (percentage change in consumer prices), covering 1960–2025 (66 observations), was sourced from the World Bank World Development Indicators via the Federal Reserve Bank of St. Louis’s FRED database (series FPCPITOTLZGNGA). The annual NGN/USD exchange rate (national currency units per US dollar, market and estimated), covering 1960–2023 (64 observations), was sourced from the Penn World Table 11.0 via FRED (series XRNCUSNGA618NRUG) and log-transformed prior to modelling, consistent with standard practice for exchange rate series exhibiting

multiplicative growth. Both series are considerably longer than the ten-observation annual series used in earlier bootstrap comparisons on Nigerian financial data, allowing a more defensible empirical test.

Computational Environment

All estimation and resampling routines were implemented

in Python 3 using NumPy, SciPy, and statsmodels.

RESULTS AND DISCUSSION

Simulation Study

The stationary bootstrap attained the lowest MSE in all five simulated models, with the moving block bootstrap a close second in every case. The sieve bootstrap did

Table 1: reports average bootstrap bias and MSE for each simulated model and method (15 Monte Carlo replications, 60 bootstrap replicates each; see Section 6 on scope limitations).

Model / Method	Bias	MSE
ARIMA(1,0,0)		
Stationary	-0.1894	1.2066
Block	-0.1740	1.2498
Sieve	-0.1472	1.4763
ARIMA(1,1,0)		
Stationary	1.1146	466.71
Block	-0.3091	521.26
Sieve	-12.634	1663.23
ARIMA(1,1,1)		
Stationary	-0.2997	750.84
Block	0.4942	877.81
Sieve	16.678	4325.11
ARIMA(2,1,1)		
Stationary	-1.0914	1051.95
Block	0.5358	1097.05
Sieve	13.661	6387.07
ARIMA(1,1,3)		
Stationary	-0.3965	483.47
Block	-1.8002	563.75
Sieve	6.750	919.00

not achieve the lowest MSE in any specification and was substantially worse for all four integrated ($d = 1$) processes, with MSE between 2 and 6 times that of the stationary bootstrap; its bias was also the largest in absolute value for every integrated process. This pattern differs from the expectation in some earlier work that block-based methods are dominated by processes with genuine moving-average components, or that the sieve bootstrap remains competitive wherever a finite-order autoregression can approximate the process. Here, the stationary bootstrap's randomised block length appears to have provided a more robust general-purpose resampling scheme across the specifications examined, while the sieve bootstrap's reconstruction through re-integration of a resampled AR-residual series appears particularly sensitive to compounding once random

innovations are cumulatively summed an effect that grows with the number of AR and MA parameters being sieved.

Data: Nigerian Headline Inflation (1960-2025)

An augmented Dickey-Fuller test on the level series rejected a unit root ($p = 0.037$), and an ARIMA(2,0,0) model was fitted directly to the level series, yielding a point forecast of $\theta = 17.52$ percentage points for the next annual observation beyond the sample. Table 2 reports bias and MSE from 1,000 bootstrap replicates.

The stationary bootstrap attained the lowest MSE (80.53), followed closely by the block bootstrap (90.20). The sieve bootstrap's MSE (857.86) was roughly an order of magnitude higher, driven by a large positive bias (+26.86), consistent with the compounding instability observed for

Table 2: Bias and Mean Squared Error of Bootstrap Methods for Nigerian Headline Inflation (1960–2025)

Method	Bias	MSE
Stationary	-1.1035	80.528
Block	-1.8356	90.199
Sieve	26.862	857.865

integrated processes in the simulation study.

Data: Log NGN/USD Exchange Rate (1960-2023)

An augmented Dickey-Fuller test on the log-level series failed to reject a unit root ($p = 0.969$), consistent with the well-documented non-stationarity of nominal exchange rate series; an ARIMA(1,1,1) model was therefore fitted to

the log series, yielding a point forecast of $\hat{\theta} = 6.627$ (log scale) for the next annual observation. Table 3 reports bias and MSE from 1,000 bootstrap replicates.

The stationary and block bootstraps were near-indistinguishable in MSE (6.42 versus 6.45), while the sieve bootstrap's MSE (204.62) was again roughly 30 times higher, with a substantial positive bias.

Table 3: Bootstrap Bias and Mean Squared Error (MSE) for Log NGN/USD Exchange Rate (1960–2023)

Method	Bias	MSE
Stationary	-0.2279	6.4165
Block	0.0075	6.4454
Sieve	13.798	204.623

DISCUSSION

Across both the simulation study and the two real Nigerian macroeconomic series, the stationary bootstrap consistently produced the lowest one-step-ahead forecast MSE, with the moving block bootstrap a close competitor throughout. The sieve bootstrap underperformed substantially in every setting, particularly for integrated processes and for both empirical series, each of which required first-differencing (or was, in the case of inflation, borderline stationary at the margin of the test used).

This pattern suggests that the sieve bootstrap's reconstruction procedure fitting a finite AR sieve to a differenced series, resampling its residuals, regenerating the series through the AR recursion, and then re-integrating is more fragile in finite samples than the block-resampling alternatives, which resample the observed (differenced) data directly rather than passing it through an estimated recursive filter before re-integration. This is a different mechanism from the moving-average-approximation argument advanced in earlier cross-national comparisons, and suggests that, at least for the sample sizes and specifications examined here, the practical risk with the sieve bootstrap on Nigerian integrated series is compounding instability in the reconstruction step rather than approximation bias from an insufficiently rich AR sieve.

For Nigerian practitioners building forecasting pipelines for inflation, exchange rates, or related integrated macroeconomic series, these results suggest the stationary bootstrap is a reasonable default, with the moving block bootstrap as a close and computationally simpler alternative. The sieve bootstrap should be applied cautiously, if at all, for integrated series without

additional safeguards on the reconstruction step for example, damping or truncating unusually large resampled innovations before re-integration.

Limitations

Several limitations should temper the strength of the conclusions drawn here. First, the Monte Carlo simulation was run at reduced scale (15 replications $\times$ 60 bootstrap draws per model) rather than a fuller design, owing to computational budget constraints; standard errors on the reported bias and MSE estimates are correspondingly wide, and larger-scale replication is needed before the method ranking can be treated as settled. Second, while both empirical series are considerably longer than the ten-observation series used in some earlier Nigerian bootstrap comparisons, they remain annual data with at most 66 observations; replication using monthly or higher-frequency official Nigerian series for example, the Central Bank of Nigeria's monthly exchange rate data or the National Bureau of Statistics' monthly Consumer Price Index bulletins would permit richer ARIMA specifications and a more powerful test. Third, block length (10 in the simulation, 5 for the real series) and the sieve order search range were fixed rather than jointly optimised, and results may be sensitive to these choices. Fourth, the apparent fragility of the sieve bootstrap's re-integration step for I(1) processes, while consistent across both the simulation and the two empirical series, warrants dedicated further investigation for instance, comparing alternative reconstruction schemes before being treated as a general property of the method rather than an artifact of the specific implementation used here.

CONCLUSION

This study benchmarked the stationary, block, and sieve bootstrap methods for time series uncertainty quantification across five simulated ARIMA processes and two long-run Nigerian macroeconomic series: headline inflation (1960-2025) and the log NGN/USD exchange rate (1960-2023). The stationary bootstrap performed best throughout, the moving block bootstrap was a close second, and the sieve bootstrap underperformed substantially in every setting, particularly for integrated series. Nigerian analysts building AI-assisted or conventional forecasting pipelines for inflation, exchange rates, and related integrated series should, pending larger-scale replication, default to the stationary or moving block bootstrap rather than the sieve bootstrap for uncertainty quantification.

REFERENCES

- Adewole, A. I. (2024). Modelling long memory volatilities of Nigeria selected macro-economic variables with ARFIMA and ARFIMA-FIGARCH. *Cumburiyet Science Journal*, 45(3), 618–628. <https://doi.org/10.17776/csj.1467360>
- Bergmeir, C., Hyndman, R. J., & Benítez, J. M. (2016). Bagging exponential smoothing methods using STL decomposition and Box–Cox transformation. *International Journal of Forecasting*, 32(2), 303–312. <https://doi.org/10.1016/j.ijforecast.2015.07.002>
- Bühlmann, P. (1997). Sieve bootstrap for time series. *Bernoulli*, 3(2), 123–148.
- Efron, B. (1979). Bootstrap methods: Another look at the jackknife. *The Annals of Statistics*, 7(1), 1–26. <https://doi.org/10.1214/aos/1176344552>
- Federal Reserve Bank of St. Louis. (n.d.). *Exchangerate (market estimated) for Nigeria (XRNCUSNGA618NRUG)* [Data set]. FRED. Retrieved October 6, 2026, from <https://fred.stlouisfed.org/series/XRNCUSNGA618NRUG>
- Federal Reserve Bank of St. Louis. (n.d.). *Inflation, consumer prices for Nigeria (FPCPITOTLZGNGA)* [Data set]. FRED. Retrieved October 6, 2026, from <https://fred.stlouisfed.org/series/FPCPITOTLZGNGA>
- Künsch, H. R. (1989). The jackknife and the bootstrap for general stationary observations. *The Annals of Statistics*, 17(3), 1217–1241. <https://doi.org/10.1214/aos/1176347265>
- Lahiri, S. N. (1999). Theoretical comparisons of block bootstrap methods. *The Annals of Statistics*, 27(1), 386–404.
- Liu, R. Y., & Singh, K. (1992). Moving blocks jackknife and bootstrap capture weak dependence. In R. LePage & L. Billard (Eds.), *Exploring the limits of bootstrap* (pp. 225–248). Wiley.
- Oni, L. B., & Akinsanya, T. A. (2026). Analysing the effects of inflation, interest rate and exchange rate on Nigerian economy. *American Journal of Applied Statistics and Economics*, 5(1), 122–130. <https://doi.org/10.54536/ajase.v5i1.7070>
- Pan, L., & Politis, D. N. (2016). Bootstrap prediction intervals for linear, nonlinear and nonparametric autoregressions. *Journal of Statistical Planning and Inference*, 177, 1–27. <https://doi.org/10.1016/j.jspi.2014.10.003>
- Politis, D. N., & Romano, J. P. (1994). The stationary bootstrap. *Journal of the American Statistical Association*, 89(428), 1303–1313. <https://doi.org/10.1080/01621459.1994.1047687>
- Udoudo, K. J., Emele Kalu, I., & Oduola, K. (2023). Impact of liquefied natural gas exports on the Nigerian exchange rate: An ARDL cointegration approach, 2000 to 2021. *American Journal of Applied Statistics and Economics*, 3(1), 1–14. <https://doi.org/10.54536/ajase.v3i1.2216>