Journal of Natural Sciences and Mathematics Research
J. Nat. Scien. & Math. Res. Vol. 10 No. 1 (2024) 1-11
ISSN: 2460-4453 (Online); 2614-6487 (Printed)
Available online at http://journal.walisongo.ac.id/index.php/jnsmr

Bootstrap-after-bootstrap for autoregressive models: an application to Indonesian value of export oil and gas

Umi Mahmudah[1]

[1]Department of Mathematics Education, Faculty of Education and Teacher Training, Universitas Islam Negeri K.H. Abdurrahman Wahid Pekalongan, Indonesia

*Corresponding author's e-mail: umi.mahmudah@uingusdur.ac.id

ABSTRACT

This research focuses on predicting the value of oil and gas exports in Indonesia, employing a hybrid methodology that combines autoregressive models and a bootstrap approach. Specifically, this research applies the bootstrap-after-bootstrap approach to showcase its effectiveness in improving the accuracy of parameter estimates. Analysis results indicate that the autoregressive model with an order of p=2 minimizes the AIC, BIC, and HQ values, yielding AIC=9.833775, BIC=10.03125, and HQ=9.883440, respectively. Consequently, the AR(2) model emerges as the optimal choice for predicting Indonesia's export value of oil and gas. This research utilizes varying numbers of bootstrap replications (B=100, 250, 500, 1000, and 10000) to assess the impact on prediction intervals. Prediction intervals exhibit less smoothness for B=100 and B=250, whereas B=500 and B=1000 result in a considerably smoother pattern. The highest level of smoothness is achieved for B=10000. The findings underscore that bootstrap-after-bootstrap prediction intervals provide the most accurate and conservative assessment of future uncertainty. Moreover, predictive analysis for the upcoming five periods indicates a projected decline in the export value of oil and gas in Indonesia. Overall, this research demonstrates the efficacy of the bootstrap-after-bootstrap approach in enhancing the precision of predictions and providing robust insights into future uncertainties surrounding Indonesia's oil and gas export market.

Keywords: autoregressive; bootstrap; forecasting; oil and gas

Introduction

The Indonesian economy, characterized by its reliance on the export of oil and gas, presents a unique set of challenges for time series analysis. The volatility of the Indonesian value of export in the oil and gas sector is intricately linked to various dynamic factors, demanding sophisticated modeling approaches to capture the complexity of these influences.

Global oil prices exhibit considerable volatility due to factors such as geopolitical tensions (Alqahtani & Klein, 2021; Cunado et al., 2020), supply-demand imbalances (Chai et al., 2021), and changes in production levels by major oil-producing nations (Demirbas et al., 2017; Su et al., 2020). For instance, a sudden geopolitical conflict in a major oil-producing region, like the Middle East, can lead to a spike in global oil prices (El-Gamal & Jaffe, 2018; Noguera-Santaella, 2016). This, in turn, directly impacts the export value of Indonesian oil and gas products, as the country is a significant player in the global energy market. Consider the period between 2014 and 2016 when global oil prices experienced a significant downturn (Razmi et al., 2016). During this time, geopolitical tensions in the Middle East, coupled with a surge in global oil production, led to a substantial drop in oil prices. Indonesia, as a major oil exporter, faced the dual challenge of reduced export volumes and diminished export values due to the global price slump.

The significance of predictive analysis on the value of oil and gas exports in Indonesia lies in its pivotal role for strategic decision-making and economic planning. Given Indonesia's substantial reliance on the oil and gas sector, accurate predictive models offer crucial insights for the government, businesses, and investors. However, in response to uncertainties influences in predictive analysis, employing advanced approach like bootstrap-after-bootstrap becomes essential. This approach, by accounting for serial dependencies and providing robust parameter estimates, offers a more accurate representation of the intricate relationships within the time series data (Errouissi et al., 2015; Kim, 2001; Kim & Shamsuddin, 2020), allowing for a nuanced understanding of the Indonesian oil and gas export sector's dynamics.

Traditional methodologies often encounter challenges, particularly in the context of emerging economies like Indonesia, where non-Gaussian distribution and limited sample sizes prevail (Errouissi et al., 2015; Hyndman & Athanasopoulos, 2018; Kim, 2001; Staszewska-Bystrova et al., 2011). This research uses a hybrid method, in which combined autoregressive models and bootstrap-after-bootstrap approach. Motivated by the limitations of conventional techniques, the bootstrap-after-bootstrap approach is a robust solution for the estimation of autoregressive models in the context of the Indonesian oil and gas export sector. This approach aims to address the intricacies of serial correlation in time series data, providing more accurate and reliable parameter estimates (Bozorg et al., 2021; Clements & Kim, 2007; Kim, 2001).
The primary objective of this research is to predict the value of export oil and gas in Indonesia, in which focuses on the prediction intervals. Through the application of bootstrap-after-bootstrap approach, this research aims to illustrate how this approach enhances the accuracy of parameter estimates and provides robust insights into the serial dependencies within the data (Kim, 2001; Mahmudah, 2023). A case study delves into the dynamic nature of Indonesia's oil and gas export values, considering factors such as global market trends, geopolitical influences, and domestic policy changes.

The conventional (nonparametric) bootstrap method utilized in previous research produces bootstrap replicates that are inherently biased in small samples, owing to biases inherent in AR parameter estimators (Friedrich et al., 2020; Kilian, 1998; Kim, 2001). A study demonstrated that intervals derived through the bootstrap-after-bootstrap approach exhibit significantly superior performance compared to those relying on conventional methods, especially in small sample sizes (Kilian, 1998). Furthermore, a study reported that the bootstrap-after-bootstrap method emerges as a superior alternative to asymptotic and standard bootstrap prediction intervals. Bootstrap-after-bootstrap prediction intervals consistently offer the most precise and cautious evaluation of future uncertainty, particularly in situations with small sample sizes, across a wide range of circumstances, including AR models with roots near or equal to unity (Kim, 2001; Shang, 2018).

The key novelty is the development and application of the bootstrap-after-bootstrap technique to autoregressive models. Unlike traditional bootstrap, it iteratively resamples within each bootstrap iteration, improving precision in parameter estimates and prediction intervals. Applied to Indonesian oil and gas exports, this approach captures intricate patterns, offering a more accurate and stable model for economic trend forecasting. This research not only contributes to methodological advancements in time series analysis but also provides a practical and valuable tool for researchers and policymakers navigating the complexities of the Indonesian oil and gas export sector. The real-world application serves as a concrete example of the methodology's effectiveness in addressing the specific challenges posed by economic time series data in a dynamic and critical sector.

Methods

Autoregressive model

The traditional of an autoregressive (AR) model is a type of model in time series analysis where the values of a variable at a particular time are determined by the values of the variable at previous times. In the case of a K-dimensional stationary AR(p) model, represented as follows (Kim, 2001):

$$Y_t = v + A_1 Y_{t-1} + \cdots + A_p Y_{t-p} + u_t \tag{1}$$

Where Y_t is the observed variable at time t; v is a constant; $A_1, A_2, \dots A_p$ are autoregressive coefficient parameters; $Y_{t-1}, Y_{t-2}, \dots, Y_{t-p}$are the values of the variable at previous times up to p

time steps; u_t is the random disturbance at time t, assumed to have a normal distribution with mean 0 and variance σ^2.

An AR(p) model like this is used to model the relationship between the variable Y at time t and its past values up to p time steps. The autoregressive coefficients $A_1, A_2, \dots A_p$ determine the extent to which previous values influence the current value of the variable.
Stationarity means that the statistical properties of the model do not change over time. This implies that the mean, variance, and covariance of the variable do not depend on time. It is an essential assumption in time series analysis because it allows for consistent estimation of model parameters. Therefore, the AR(p) model enables us to understand and model temporal dependence patterns in time series data.

The backward AR(p) model linked to the forward model (1) can be expressed as

$$Y_t = \mu + H_1 Y_{t+1} + \cdots + H_p Y_{t-p} + v_t \tag{2}$$

Where $t = 0,1,2, \dots$ $E(v_t) = 0$ and $E(v_t v_t') = \sum_v$ is symmetric positive definite matrix with finite elements.

Bootstrap-after-bootstrap Prediction Intervals

The backward model is employed in bootstrapping to generate bootstrap forecasts based on the most recent p observations of the original series. Suppose $\hat{A} = (\hat{v}, \hat{A}_1, \dots, \hat{A}_p)$ and $\widehat{H} = (\hat{\mu}, \widehat{H}_1, \dots, \widehat{H}_p)$ are the least square estimator for A and H. Then, predictions are produced in a conventional manner by utilizing the estimated coefficients (Kim, 2001):

$$\hat{Y}_n(h) = \hat{v} + \hat{A}_1 \hat{Y}_{n-1}(h) + \cdots + \hat{A}_p \hat{Y}_{n-p}(h) \tag{3}$$

Where $\hat{Y}_n(j) = Y_{n+j}$ for $j \leq 0$. The asymptotic prediction interval (API) for the kth AR component, maintaining a nominal coverage rate of $100\left(1 - {}^{\alpha}/_{K}\right)\%$, can be characterized as:

$$API_k = \hat{Y}_{k,n}(h)\, z_\tau \hat{\sigma}_k(h) \tag{4}$$

Where $k = 1,2, \dots, K$. $\hat{Y}_{k,n}(h)$ represents the k-th element of $\hat{Y}_n(h)$, and z_τ is the upper percentile corresponding to the 100 τth percentile of the standard normal distribution with $\tau = .5\left({}^{\alpha}/_{K}\right)$. Additionally, $\hat{\sigma}_k(h)$ is the square root of the k-th diagonal element of $\sum_Y(h)$. The procedure for obtaining bootstrap-after-bootstrap prediction intervals can be derived as follows (Kim, 2001):

Step 1:

Calculate $\widehat{A}, \widehat{H}$, and residuals $\hat{u}_t$ and $\hat{v}_t$ using equation (1) when given n realizations $(Y_1, \dots, Y_n)$.

Step 2:

Calculate the bootstrap estimator $\hat{A}^*$ and $\widehat{H}^*$ for A and H. The calculations for the biases of $\hat{A}$ and $\widehat{H}$ are determined through the estimation of $bias(\hat{A}) = \hat{A}^* - \hat{A}$ and $bias(\widehat{H}) = \widehat{H}^* - \widehat{H}$
bias XY and bias ZY. The pseudo datasets are generated as $Y_t^* = \hat{v} + \hat{A}_1 Y_{t-1}^* + \cdots + \hat{A}_p Y_{t-p}^* + u_t^*$ where u_t^* is a random draw with replacement from $\hat{u}_t$. Further, $Y_t^* = \hat{\mu} + \widehat{H}_1 Y_{t+1}^* + \cdots + \widehat{H}_p Y_{t+p}^* + v_t^*$ where v_t^* is a random draw with replacement from $\hat{v}_t$.

Step 3:

Calculate the bias-corrected estimators $\hat{A}^c$ and $\widehat{H}^c$ using $bias(\hat{A})$ and $bias(\widehat{H})$.

Step 4:

Generate a pseudo dataset using equation (2) as $Y_t^* = \hat{\mu}^c + \widehat{H}_1^c Y_{t+1}^* + \cdots + \widehat{H}_p^c Y_{t+p}^* + v_t^*$. The initial p values are initialized to match the last p values of the original series.

Step 5:

Employing these pseudo datasets, the coefficient matrices of the forward model (1) are estimated using the Least Squares (LS) method, and the resulting estimators are denoted as $\tilde{A}^*$.

Step 6:

Calculate the bias corrected estimator $\tilde{A}^c$ using the biases in $\tilde{A}^*$.

Then, the iterative generation of bootstrap replicates for AR forecasts is carried out as follows:

$$Y_n^*(h) = \tilde{v}^c + \tilde{A}_1^c Y_n^*(h-1) + \cdots + \tilde{A}_p^c Y_n^*(h-p) + u_{n+h}^* \tag{5}$$

Where $Y_n^*(j) = Y_{n+j}^* = Y_{n+j}$ for $j \leq 0$ and u_{n+h}^* is a random term draw from $\hat{u}_t$ with replacement. Generating pseudo datasets repeatedly as outlined in (5), B times, will result in the bootstrap forecast distribution $Y_n^*(h;i)_{i=1}^B$.

Data

This research utilized secondary data in the form of the export values of oil and gas in Indonesia, which were published by Statistics Indonesia in the year 2023. The data covered the period from January 2022 to November 2023, totaling 23 series. Figure 1 illustrates the data series employed in the predictive analysis of the export values of oil and gas in Indonesia.

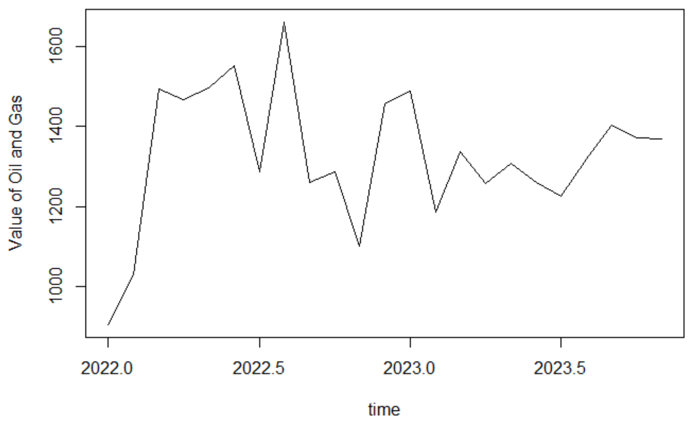


Figure 1. Export values of oil and gas (in million US$)

Figure 1 provides a chronological overview of the monthly variations in the export values of oil and gas, depicting fluctuations and trends over the specified period. The most significant increase in the time series data occurred from January to February 2022, with a notable rise of 463.7. Conversely, the most substantial decrease, indicating a decline in the export value of oil and gas, was observed from July 2022 to August 2022, with a reduction of -403.9.

The COVID-19 pandemic has brought about unprecedented uncertainty across all facets of life, extending its effects to the export values of oil and gas in Indonesia. Predicting these export values entails grappling with inherent uncertainty stemming from various factors. Fluctuations in global oil prices, heavily influenced by complex supply-demand dynamics, geopolitical tensions, and economic conditions, pose a significant challenge to accurate forecasting. Moreover, changes in government policies, such as taxation, production quotas, and environmental regulations, further contribute to the uncertainty surrounding future export values. Additionally, geopolitical events, such as conflicts in oil-producing regions or shifts in international trade agreements, introduce unpredictable elements into the equation. In the face of such uncertainty, traditional forecasting methods may fall short of capturing the full spectrum of potential outcomes. Here, the Bootstrap-after-bootstrap approach emerges as a valuable tool. By generating multiple bootstrap samples from historical export data, this approach enables the simulation of diverse scenarios and the assessment of prediction variability. Moreover, through the computation of prediction intervals or confidence intervals based on these samples, analysts can gain insights into the range of potential export values, thereby informing more robust decision-making processes amidst uncertain circumstances.

Model selection

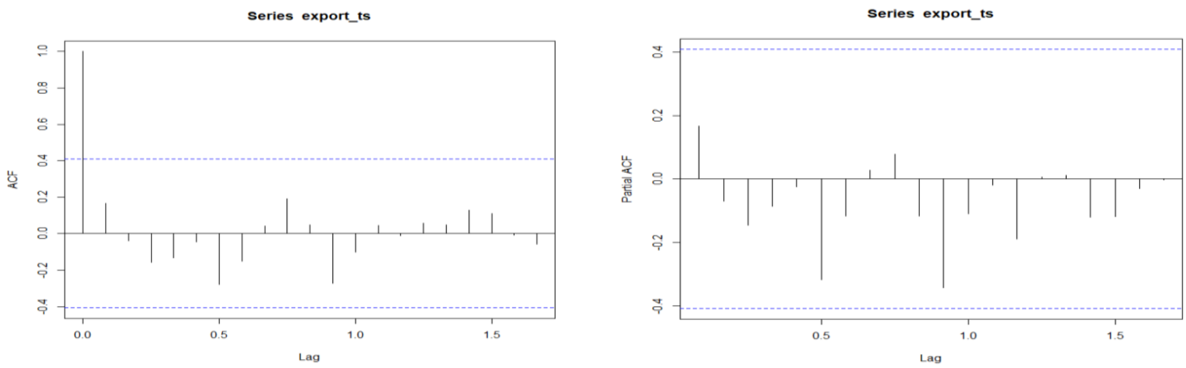
To utilize statistical criteria such as the Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), or Hannan-Quinn Information Criterion (HQ) in predicting the export values of oil and gas in Indonesia using an autoregressive (AR) model (AR(p)), several steps must be followed. Firstly, a range of AR models with varying orders p (representing the number of lagged terms) needs to be specified for consideration. Following model fitting, the AIC, BIC, and HQ criteria are computed for each model, reflecting the trade-off between goodness of fit and model complexity. Lower criterion values signify better model performance. Next, comparisons are made across all AR(p) models considered, with the model exhibiting the lowest values of each criterion being chosen as the optimal AR(p) model for predicting export values. Utilizing this selected model, predictions of future export values can be generated using techniques like one-step ahead or multi-step ahead forecasting. Subsequently, the predictive performance of the chosen AR(p) model is assessed through methods such as backtesting or cross-validation, ensuring the accuracy and reliability of predictions against observed export values. Lastly, refinement of the modeling process may be undertaken iteratively, considering alternative specifications or additional predictors to enhance predictive accuracy. By systematically applying these steps and leveraging statistical criteria like AIC, BIC, and HQ, analysts can effectively select the most appropriate AR(p) model for predicting export values of oil and gas in Indonesia, thereby improving the quality and reliability of forecasts.

Results and Discussions

To generate both point forecasts and prediction intervals, this research utilized bias-correction techniques through the bootstrap-after-bootstrap technique within autoregressive time series models. The application of these techniques is essential as relying solely on point forecasts lacks comprehensive insight. The absence of prediction intervals makes it challenging to evaluate the precision and uncertainty associated with the estimates (De Livera et al., 2011; Hyndman & Athanasopoulos, 2018; Mahmudah et al., 2023).

In prediction analysis using an autoregressive model, it is essential to ascertain the suitable AR order by employing statistical criteria like the Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), or Hannan-Quinn Information Criterion (HQ). These criteria aid in the identification of the optimal order (p) for the Autoregressive (AR) model. AIC is preferred because it focuses on predicting accuracy while penalizing complexity to a lesser extent. BIC imposes a more significant penalty based on sample size, making it excellent for selecting the correct model from a set, especially in large-sample cases. HQ, a compromise between AIC and BIC, serves as an alternative when neither AIC nor BIC produces satisfying results. AIC and BIC are the most often utilized of these; AIC is selected for optimizing predictive performance, whilst BIC is chosen to ensure the most parsimonious model, which is very important when dealing with large amounts of data. Models with the lowest AIC, BIC, or HQ values are deemed the most appropriate. Reduced values signify a superior balance between the quality of model fit and its complexity. The results of the analysis revealed that the order p=2 yielded the minimum AIC, BIC, and HQ values, with AIC=9.833775; BIC= 10.03125, and HQ=9.883440, respectively. Therefore, AR(2) is the best model to predict the Indonesian export value of oil and gas.

Furthermore, understanding the stationarity of time series data is crucial in time series analysis, particularly when applying autoregressive models. Stationarity ensures that the statistical properties of the data, such as mean and variance, remain constant over time. In the context of forecasting and time series analysis, the Autocorrelation Function (ACF) and the Partial Autocorrelation Function (PACF) plots are the most essential tools to check the nature of the serial correlation. These graphs also determine the sort of time series model to apply based on the observed correlation. ACF and PACF plots aid in assessing stationarity by revealing patterns of correlation and partial correlation between data points at different lags. Positive values show that data at the time are positively correlated to their previous values. This implies that if the time series is positive (or negative) in a specific time period, the corresponding lag time will also exhibit positive values (or negative values). Inversely, negative values mean that the current time results are anti-proportional or inversely related to their past values. This means that in an inverted yield curve, as long-term yields rise, short-term yields fall, or if short-term yields rise, long-term yields fall. In Figure 2, the ACF plot demonstrates the correlation between observations at different time lags, while the PACF plot focuses on the correlation between observations while controlling for the effects of previous lags.



(a) (b)

Figure 2. a) ACF plot; b) PACF Plot

In the presented ACF plot above, the autocorrelation values remain within the dotted line region, indicating that the time series data maintains a stationary behavior. Sampling variability can cause autocorrelation estimates to deviate slightly. If the overall ACF plot shows a sharp decline after lag 1 and other autocorrelation values are within the confidence area, the data can still be considered stationary. This stationary nature is a favorable condition for autoregressive modeling, as it ensures that the statistical properties of the data remain constant over time, enhancing the reliability of predictions based on the autoregressive model. The results of prediction based on the traditional AR (2) model for the following five periods (months) are presented in the table 1.

Table 1. Forecasting results using traditional autoregressive model

Periods	**Point forecast**	**5%**	**95%**
h1	1.34	1.11	1.69
h2	1.34	1.11	1.69
h3	1.34	1.11	1.69
h4	1.34	1.11	1.69
h5	1.34	1.11	1.69

To acquire bias-corrected forecasts and bootstrap prediction interval estimates, this research employed varying numbers of bootstrap replications (B=100, 250, 500, 1000, and 10,000). The selection of specific values for B involves a careful consideration of trade-offs between accuracy, computational resources, robustness, comparability with existing research, and the precision of estimates. The values chosen in this research study likely reflect a balance between these factors based on the specific context and objectives of the analysis. The forecasting analysis was executed for the following five periods, aligning with the five months succeeding

November 2023. The results of bootstrap-after-bootstrap prediction intervals and bootstrap bias-corrected point forecasts are presented in the table 2.

Table 2. Forecasting results using bootstrap-after bootstrap approach

Periods	B=100			B=250			B=500			B=1,000			B=10,000		
	Point forecast	5 %	95 %	Point forecast	5 %	95 %	Point forecast	5 %	95 %	Point forecast	5 %	95 %	Point forecast	5 %	95 %
h1	1.26	1.12	1.62	1.26	1.10	1.66	1.26	1.12	1.66	1.26	1.12	1.65	1.26	1.12	1.65
h2	1.26	1.12	1.70	1.26	1.13	1.72	1.26	1.09	1.71	1.26	1.09	1.69	1.26	1.10	1.69
h3	1.25	1.09	1.66	1.25	1.10	1.68	1.25	1.09	1.69	1.25	1.09	1.71	1.25	1.09	1.71
h4	1.24	1.07	1.70	1.24	1.10	1.72	1.24	1.09	1.74	1.24	1.08	1.72	1.24	1.08	1.73
h5	1.23	1.07	1.71	1.23	1.04	1.75	1.23	1.05	1.74	1.23	1.07	1.75	1.23	1.08	1.74

The forecasting results presented in Table 1 and Table 2 demonstrate variations in both point forecasts and prediction intervals, reflecting differences in the methodologies employed. Table 1 showcases outcomes derived from a traditional autoregressive (AR) model, where a constant point forecast of 1.34 is observed across all forecast periods (h1 to h5). The values in Table 1 tend to be the same for all periods (h1 to h5) because the table utilizes results from a traditional autoregressive model. In an autoregressive model, predictions for each period are based on previous values in the time series, with the model estimating consistent parameters from past observations. In the context of Table 1, the consistent prediction values across all periods indicate that the autoregressive model may not account for changing trends, seasonal patterns, or other external factors that could affect the observed values over time. This suggests that the model might be too simplistic to capture the complexity of the observed data or that the data may not exhibit significant variation among the observed time periods. Table 2 employs the bootstrap-after-bootstrap approach, allowing for the generation of forecasts with varying numbers of bootstrap replications (B=100, 250, 500, 1,000, and 10,000). Here, the point forecasts exhibit slight fluctuations across different values of B, ranging from 1.23 to 1.26, albeit within a narrow margin. Likewise, the prediction intervals display minor variability, with lower bounds ranging from 1.04 to 1.12 and upper bounds spanning from 1.62 to 1.75 across different B values. Notably, higher B values generally lead to narrower prediction intervals, indicating improved precision. These differences underscore the impact of methodological choices on forecasting outcomes, with the traditional AR model offering consistent forecasts while the bootstrap-after-bootstrap approach introduces slight variability, albeit with the potential for enhanced precision with increased bootstrap replications.

From Table 2, the values of 5% and 95% referred to prediction intervals that encompass 90% of the prediction distribution, with 5% on the lower side (lower bound) and 95% on the upper side (upper bound). The results of the prediction analysis using bias-corrected forecasts and bootstrap prediction intervals in Table 2 indicated consistent figures, with no significant differences observed. All point forecasts generated from varying numbers of bootstrap replications indicated the consistency of the results.

In Table 2, the 5% and 95% values denoted prediction intervals that cover 90% of the prediction distribution, allocating 5% to the lower side (lower bound) and 95% to the upper side (upper bound). The prediction analysis results, employing the bootstrap-after-bootstrap technique and bootstrap prediction intervals, as shown in Table 2, exhibited uniform figures with no notable disparities. The coherence in results was evident across all point forecasts derived from different numbers of bootstrap replications (Hyndman & Athanasopoulos, 2018; Mahmudah et al., 2023). Moreover, Table 2 illustrated the robust stability of the model utilized for prediction analysis, indicating that it was not responsive to fluctuations in bootstrap data. The point forecasts exhibited a notable degree of consistency, regardless of the specific value of B employed.

Furthermore, the results of the predictive analysis on point forecasts for the upcoming 5 periods suggest a decline in the export value of oil and gas in Indonesia, with the predicted values being 1.26, 1.26, 1.25, 1.24, and 1.23, respectively. The estimate starts with the export value remaining stable at $1.26 billion for the first two quarters, indicating a short plateau before the fall begins. By the third period, there is a modest reduction as the value falls to $1.25 billion, confirming the start of a downtrend. This pattern continues into the fourth period, with export value dropping to $1.24 billion. By the fifth period, the reduction has progressed to the lowest estimated value of $1.23 billion. These figures represent the anticipated values for each period and imply a downward trend in the export values over the specified time frame. The numerical values signify the model's predictions for the upcoming months, indicating a potential contraction in the export values of oil and gas based on the established forecasting model.

Figure 3 illustrates the representation of point forecasts and prediction intervals with different numbers of bootstrap replications: B=100, 250, 500, 1000, and 10000. This research utilized quantile estimates, explicitly focusing on quartiles derived from the estimated distribution. The purpose of employing prediction intervals was to evaluate the coverage probability within the specified range under the distribution (Hyndman & Athanasopoulos, 2018; Mahmudah et al., 2023). In this research, a 95% prediction interval was utilized, determined by extracting the 2.5% and 97.5% quantiles from the forecast distribution. The adoption of this 95% prediction interval aligns with its common use in forecasting analysis, alongside the frequently employed 80% prediction interval (Chamdani et al., 2019; De Livera et al., 2011; Hyndman & Athanasopoulos, 2018).

In Figure 3, the blue line corresponds to the point forecasts for the anticipated data spanning the subsequent 5 periods (months). Concurrently, the red line signifies the prediction intervals. It is noteworthy that as the prediction analysis indicated increased uncertainty, the prediction intervals tended to expand. The visualization of prediction intervals also serves as a means to assess the effectiveness of forecasting models. As observed in Figure 3, the consistent alignment of point forecasts within the prediction intervals suggests that the model adeptly captures the variability inherent in the data.

Moreover, as depicted in Figure 3, a distinct pattern emerges concerning the smoothness of prediction intervals at different levels of B. For B=100 and B=250, the intervals exhibit a relatively less smooth trajectory. This implies that with a lower number of bootstrap replications, the intervals tend to be more jagged and less consistent. This phenomenon is likely attributed to the limited sampling variability captured by a smaller number of bootstrap samples. Contrastingly, as the number of bootstrap replications increases to B=500 and B=1000, a discernible improvement in smoothness becomes apparent. The prediction intervals manifest a notably smoother and more continuous pattern, suggesting enhanced stability and reliability in capturing the underlying variability of the data. This phenomenon aligns with the fundamental principle of bootstrap methods, where a more significant number of replications leads to a more accurate representation of the population distribution. Noteworthy is the substantial increase in smoothness observed when B=10000. At this level of replications, the prediction intervals reach their peak smoothness. This observation underscores the asymptotic behavior of bootstrap methods, indicating that as the number of replications approaches infinity, the estimation becomes increasingly precise, resulting in smoother and more refined prediction intervals (Chernick & LaBudde, 2014; Mahmudah, 2023; Thombs & Schucany, 1990).

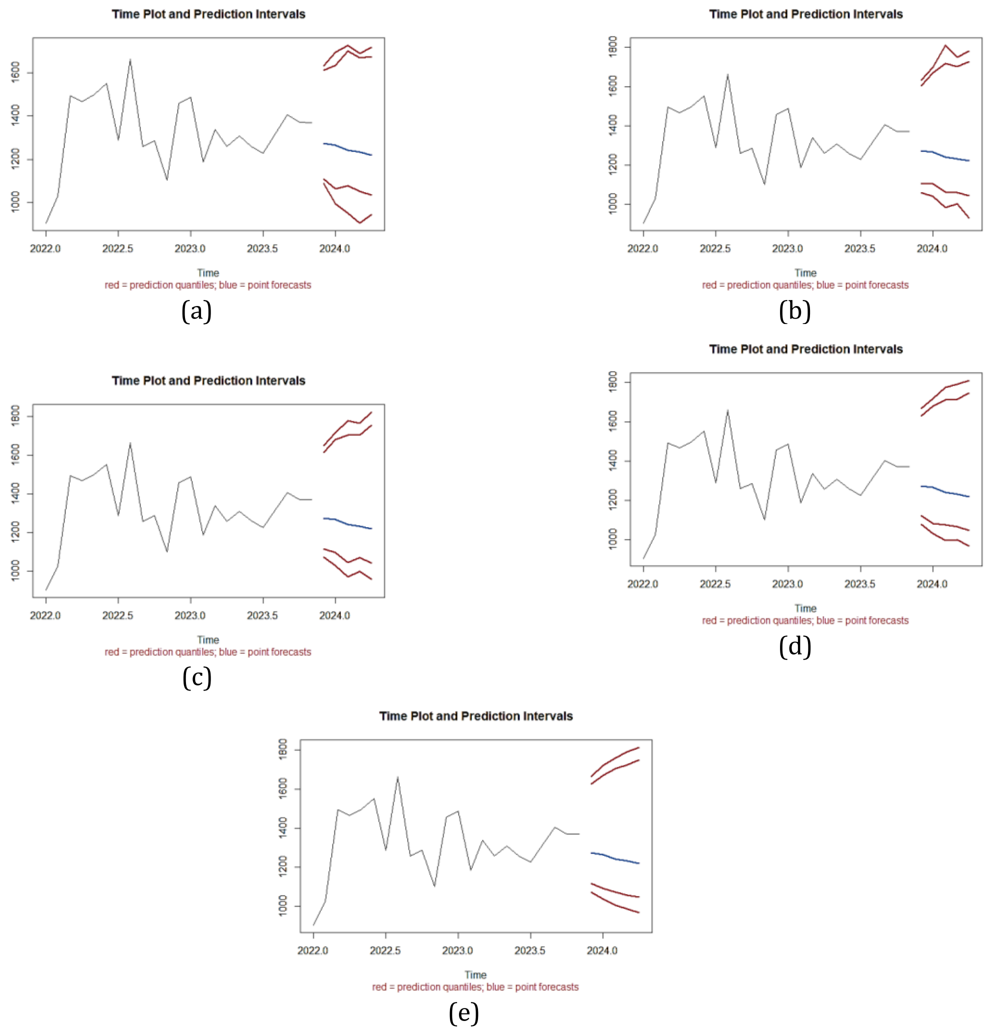


Figure 3. a) Time plot and prediction intervals (B=100); b) Time plot and prediction intervals (B=250); c) Time plot and prediction intervals (B=500); d) Time plot and prediction intervals (B=1000); e) Time plot and prediction intervals (B=10000)

These findings have significant implications for the application of bootstrap-after-bootstrap in autoregressive models, particularly when analyzing the Indonesian value of exports in the oil and gas sector. The choice of an optimal number of bootstrap replications is crucial, as it directly influences the smoothness and reliability of prediction intervals, ultimately impacting the accuracy of model predictions and the robustness of statistical inferences (Hyndman & Athanasopoulos, 2018; Kim, 2001; Masarotto, 1990).

In conclusion, Figure 3 provides a visual representation of the relationship between the number of bootstrap replications and the smoothness of prediction intervals. The insights gained from this analysis contribute to the refinement of the Bootstrap-after-bootstrap methodology, offering researchers and practitioners a nuanced understanding of how the choice of B influences the stability and precision of autoregressive model predictions in the specific context of Indonesian oil and gas exports.

Conclusion

In conclusion, this research utilized a hybrid methodology integrating autoregressive models and the bootstrap-after-bootstrap approach to predict the value of oil and gas exports in Indonesia. The analysis identified the AR(2) model as the optimal choice based on minimized AIC, BIC, and HQ values. The investigation into varying numbers of bootstrap replications revealed that higher replication numbers, particularly B=10000, resulted in the smoothest prediction intervals, emphasizing the precision of the bootstrap-after-bootstrap method. Furthermore, this research's predictive analysis for the upcoming five periods indicated a projected decline in the export value of oil and gas in Indonesia. This insight underscores the importance of accurate forecasting in anticipating economic trends and making informed decisions. Overall, the findings of this research highlight the effectiveness of the bootstrap-after-bootstrap approach in enhancing the accuracy of parameter estimates and providing robust prediction intervals. The method's ability to offer a precise and conservative evaluation of future uncertainties is particularly valuable for decision-makers in the context of Indonesia's oil and gas export market. Future research may explore additional complexities and external factors to refine predictive models further and contribute to a more comprehensive understanding of the dynamics influencing oil and gas export values. This includes evaluating the influence of geopolitical changes like trade sanctions and international relations, as well as technical improvements that may change production efficiencies or introduce feasible renewable alternatives. Furthermore, assessing the impact of environmental policies, such as emissions laws, and broader economic conditions, such as global recessions or growth periods, could provide more information. The research may also take into account trends in consumer behavior toward more sustainable energy sources, as well as the resilience of supply chains to disruption. Collectively, these elements can improve predictive models and provide a more comprehensive picture of global energy markets.

Conflicts of interest

The authors declare that there are no conflicts of interest.

References

Alqahtani, A., & Klein, T. (2021). Oil price changes, uncertainty, and geopolitical risks: On the resilience of GCC countries to global tensions. *Energy*, *236*, 121541. https://doi.org/https://doi.org/10.1016/j.energy.2021.121541

Bozorg, M., Bracale, A., Carpita, M., De Falco, P., Mottola, F., & Proto, D. (2021). Bayesian bootstrapping in real-time probabilistic photovoltaic power forecasting. *Solar Energy*. https://doi.org/10.1016/j.solener.2021.07.063

Chai, J., Zhang, X., Lu, Q., Zhang, X., & Wang, Y. (2021). Research on imbalance between supply and demand in China's natural gas market under the double-track price system. *Energy Policy*, *155*, 112380. https://doi.org/https://doi.org/10.1016/j.enpol.2021.112380

Chamdani, M., Mahmudah, U., & Fatimah, S. (2019). Prediction of Illiteracy Rates in Indonesia Using Time Series. *International Journal of Education*, *12*(1), 34–41. https://doi.org/10.17509/ije.v12i1.16589

Chernick, M. R. M. R., & LaBudde, R. A. R. A. R. A. (2014). *An introduction to bootstrap methods with applications to R*. John Wiley & Sons.

Clements, M. P., & Kim, J. H. (2007). Bootstrap prediction intervals for autoregressive time series. *Computational Statistics & Data Analysis*, *51*(7), 3580–3594.

Cunado, J., Gupta, R., Lau, C. K. M., & Sheng, X. (2020). Time-varying impact of geopolitical risks on oil prices. *Defence and Peace Economics*, *31*(6), 692–706. https://doi.org/https://doi.org/10.1080/10242694.2018.1563854

De Livera, A. M., Hyndman, R. J., & Snyder, R. D. (2011). Forecasting time series with complex seasonal patterns using exponential smoothing. *Journal of the American Statistical Association*, *106*(496), 1513–1527. https://doi.org/https://doi.org/10.1198/jasa.2011.tm09771

Demirbas, A., Omar Al-Sasi, B., & Nizami, A.-S. (2017). Recent volatility in the price of crude oil. *Energy Sources, Part B: Economics, Planning, and Policy*, *12*(5), 408–414. https://doi.org/https://doi.org/10.1080/15567249.2016.1153751

El-Gamal, M. A., & Jaffe, A. M. (2018). The coupled cycles of geopolitics and oil prices. *Economics of Energy & Environmental Policy*, *7*(2), 1–14. https://doi.org/https://www.jstor.org/stable/27030624

Errouissi, R., Cardenas-Barrera, J., Meng, J., Castillo-Guerra, E., Gong, X., & Chang, L. (2015). Bootstrap prediction interval estimation for wind speed forecasting. *2015 IEEE Energy Conversion Congress and Exposition (ECCE)*, 1919–1924.

Friedrich, M., Smeekes, S., & Urbain, J. P. (2020). Autoregressive wild bootstrap inference for nonparametric trends. *Journal of Econometrics*. https://doi.org/10.1016/j.jeconom.2019.05.006

Hyndman, R. J. R. J., & Athanasopoulos, G. (2018). *Forecasting: principles and practice*. OTexts. https://otexts.com/fpp3/

Kilian, L. (1998). Confidence intervals for impulse responses under departures from normality. *Econometric Reviews*, *17*(1), 1–29. https://doi.org/https://doi.org/10.1080/07474939808800401

Kim, J. H. (2001). Bootstrap-after-bootstrap prediction intervals for autoregressive models. *Journal of Business and Economic Statistics*, *19*(1), 117–128. https://doi.org/10.1198/07350010152472670

Kim, J. H., & Shamsuddin, A. (2020). A bootstrap test for predictability of asset returns. *Finance Research Letters*, *35*, 101289.

Mahmudah, U. (2023). Robust Prediction Intervals for Indonesian Inflation: A Bias-Corrected Bootstrap Approach. *Journal of Fundamental Mathematics and Applications (JFMA)*, *6*(2). https://doi.org/https://doi.org/10.14710/ihis.v%vi%i.20502

Mahmudah, U., Surono, S., Wahyu Prasetyo, P., & E Haryati, A. (2023). Forecasting educated unemployed people in Indonesia using the Bootstrap Technique. *Journal of Mahani Mathematical Research*, *12*(1), 171–182. https://doi.org/10.22103/JMMR.2022.19368.1239

Masarotto, G. (1990). Bootstrap prediction intervals for autoregressions. *International Journal of Forecasting*, *6*(2), 229–239.

Noguera-Santaella, J. (2016). Geopolitics and the oil price. *Economic Modelling*, *52*, 301–309. https://doi.org/https://doi.org/10.1016/j.econmod.2015.08.018

Razmi, F., Azali, M., Chin, L., & Habibullah, M. S. (2016). The role of monetary transmission channels in transmitting oil price shocks to prices in ASEAN-4 countries during pre-and post-global financial crisis. *Energy*, *101*, 581–591. https://doi.org/https://doi.org/10.1016/j.energy.2016.02.036

Shang, H. L. (2018). Bootstrap methods for stationary functional time series. *Statistics and Computing*, *28*(1), 1–10. https://doi.org/10.1007/s11222-016-9712-8

Staszewska-Bystrova, A., Staszewska-Bystrova, A., & Staszewska-Bystrova, A. (2011). Bootstrap prediction bands for forecast paths from vector autoregressive models. *Journal of Forecasting*, *30*(8), 721–735. https://doi.org/10.1002/for.1205

Su, C.-W., Qin, M., Tao, R., Moldovan, N.-C., & Lobonţ, O.-R. (2020). Factors driving oil price——from the perspective of United States. *Energy*, *197*, 117219. https://doi.org/https://doi.org/10.1016/j.energy.2020.117219

Thombs, L. A., & Schucany, W. R. (1990). Bootstrap prediction intervals for autoregression. *Journal of the American Statistical Association*, *85*(410), 486–492.