

Authoring Departments:

Monetary and Capital Markets Department

Information Technology Department

Namibia

A Complete Forecasting Framework for the Central Bank Autonomous Factors

May 2022

Prepared By

Romain Lafarguette (MCM, Mission Chief), Kei Moriya (ITD), Kevin Sheppard, Anastasios Panagiotelis (external experts)

Technical ASSISTANCE Report

For Official Use Only

DM5 Reference: 7198239

Mission ID(s): 21MMG6700 

Project ID(s): MCM\_NAM\_2019\_02

DISCLAIMER

*The contents of this report constitute technical advice provided by the staff of the International Monetary Fund (IMF) to the authorities of Namibia (the “TA recipient”) in response to their request for technical assistance. This report (in whole or in part) or summaries thereof may be disclosed by the IMF to IMF Executive Directors and members of their staff, as well as to other agencies or instrumentalities of the TA recipient, and upon their request, to World Bank staff, and other technical assistance providers and donors with legitimate interest, including members of the Steering Committee of AFRITAC, unless the TA recipient specifically objects to such disclosure (see* [Operational Guidelines for the Dissemination of Technical Assistance Information](http://www.imf.org/external/np/pp/eng/2013/061013.pdf)*). Publication or Disclosure of this report (in whole or in part) or summaries thereof to parties outside the IMF other than agencies or instrumentalities of the TA recipient, World Bank staff, other technical assistance providers and donors with legitimate interest, including members of the Steering Committee of AFRITAC, shall require the explicit consent of the TA recipient and the IMF’s AFR Department*

Contents Page

Glossary 5

Preface 6

Executive Summary 7

I. Introduction 10

II. Forecasting Framework 12

III. Forecasting Currency in Circulation 17

IV. Forecasting the State Account Balance 23

V. Forecasting Net Foreign Assets 29

VI. Recommendations and Conclusion 34

VII. Bibliography 36

Figures

[Figure 1. Analytical Balance Sheet of a Central Bank (Autonomous Factors in Red, Policy Position in Green, Other Items in Blue) 13](#_Toc97303916)

[Figure 2. Family of Individual Models Used in the IMF Forecasting Framework 15](#_Toc97303917)

[Figure 3. Software Infrastructure 17](#_Toc97303918)

[Figure 4. Currency in Circulation Over Time 18](#_Toc97303919)

[Figure 5. Seasonality Patterns Analysis: Daily (left) and Distance to Pay Day 19](#_Toc97303920)

[Figure 6. Out-of-Sample Performance by Models, Using the RMSE Criterion, at Different Horizons. Results are Similar for the MAE Metric 20](#_Toc97303921)

[Figure 7. Out-of-Sample Models fit for Currency in Circulation 21](#_Toc97303922)

[Figure 8. Best Single Model Forecast With Confidence Interval 22](#_Toc97303923)

[Figure 9. Forecasts Derived from Best Models Average 23](#_Toc97303924)

[Figure 10. State Account Balance Seasonal Patterns 24](#_Toc97303925)

[Figure 11. Quarterly Seasonality in Namibia 25](#_Toc97303926)

[Figure 12. Out-of-Sample Performances, SAB Forecasts 26](#_Toc97303927)

[Figure 13. Out-of-Sample Fit, SAB Forecasts 27](#_Toc97303928)

[Figure 14. Diversity in SAB Forecasts Across the Range of Models 28](#_Toc97303929)

[Figure 15. Point Forecasts and Confidence Interval for SAB 29](#_Toc97303930)

[Figure 16. NAD Flows Over Time 30](#_Toc97303931)

[Figure 17. Change in NAD Net Flows 31](#_Toc97303932)

[Figure 18. Coverage Ratio of NFA Flows Forecasts 32](#_Toc97303933)

[Figure 19. Out-of-Sample Hit-Ratio Coverage by Model at 95 percent Confidence Interval at Different Horizons 33](#_Toc97303934)

Contents Page

Tables

1. Key Recommendations 9

Appendixes

I. Data and IT Infrastructure Requirements 37

II. Statistical Forecasting Models and Methods 40

III. Evaluating Model Performances 46

IV. Model Selection and Combination 47

V. Software Documentation 48

VI. User R Guide 58

Glossary

|  |  |
| --- | --- |
| AIC | Akaike’s Information Criterion |
| ARIMA | Auto Regressive Integrated Moving Average |
| ARMA | Auto Regressive Moving Average |
| BIC | Bayesian Information Criterion |
| BON | Bank of Namibia |
| CiC | Currency in Circulation |
| ELM | Extreme Learning Machine |
| ETS | Exponential Smoothing Model (E: Error Term, T: Trend, S: Seasonal) |
| EWMA | Exponentially Weighted Moving Average |
| GARCH | Generalized Auto-Regressive Conditional Heteroskedasticity |
| IMF | International Monetary Fund |
| MAE | Median Average Errors |
| MCM | Monetary and Capital Markets Department |
| MLE | Maximum Likelihood Estimator |
| MOU | Memorandum of Understanding |
| NAD | Namibian Dollar |
| NFA | Net Foreign Assets |
| OLS | Ordinary Least Squares |
| RMSE | Root Mean Squared Errors |
| SARIMA | Seasonal Auto Regressive Integrated Moving Average |
| SAB | State Account Balance |
| SES | Simple Exponential Smoothing |
| TBATS | Trigonometric Seasonal, Box-Cox Transformation, ARMA Residuals, Trend and Seasonality |

# Preface

At the request of the Bank of Namibia, a Monetary and Capital Markets (MCM) Department remote mission took place from May to June 2021 to assist the authorities in forecasting central bank liquidity for Namibia.

The mission met with Diina Hamutumwa, Nicholas Mukasa, Aloys Mwashekele and Liina Joseph (all at the Market Operations Department). The mission wishes to thank the colleagues of the Bank of Namibia for their cooperation and productive discussions.

# Executive Summary

**The IMF provided a full-fledged liquidity forecasting to help the BON forecasting the autonomous factors using a model-based approach. The BON currently only forecasts currency in circulation using the past five-day moving-average to forecast at the near future. Although robust and simple to implement, this approach is backward-looking and does not take into account seasonal patterns. Besides, the BON doesn’t forecast the state account balance nor the net foreign assets. A more comprehensive and systematic forecasting approach would therefore be beneficial to the BON.**

**The IMF framework forecasts the three main autonomous factors (currency in circulation, state account balance and net foreign assets), and provides an aggregation method to reconcile the three factors into a single liquidity metric. The forecasting framework relies on the most updated statistical methods and** tests different families of forecasting models, with an integrated selection of the best performing model based on different metrics of out-of-sample performance. The framework also allows to automatically reparametrize the models to account for structural breaks and new data developments.Technically, the frameworkgoes beyond the projection of currency in circulation and instead allows the BON to use statistical models to forecast the liquidity position stemming from the three main autonomous factors, including the net foreign assets and the state account balance. Second, the framework implements dynamic model selection and model averaging to improve forecasting accuracy and reduce modeling risk. Third, the model automatically reconciles the three autonomous factors forecasts into a single metric and estimates the probabilistic confidence interval of the projection at a given tolerance level, which can then be used to calibrate the operations. These features are key to articulate the liquidity forecasts with liquidity management in a consistent and integrated operational approach.

**The IMF forecasting framework automatically tests a battery of models and allows the authorities to arbitrage between forecasting accuracy and modeling complexity.**While dynamic model selection and model averaging usually outperform the results of single individual models, they come at the cost of extra complexity. The central bank might prefer to use the best individual model performers, as single models are simpler to use and understand. The IMF framework automatically estimates, tests and presents the results of a battery of models, making it easy for the authorities to choose the most appropriate model.

**The IMF team found that the Seasonal Auto Regressive Integrated Moving Average (SARIMA) model is the most appropriate models to project currency in circulation, TBATS[[1]](#footnote-2) should be used for the state account balance, while an exponentially weighted moving average (EWMA) is the recommended model for projecting the net foreign assets. Based on a thorough testing of multiple models, the mission found that simple models actually performs very well and offer the best compromise between robustness and accuracy. Model averaging and dynamic model selection marginally improve accuracy, at the cost of extra complexity. However, it would be important to reassess the performance by validating all models on updated data at regular interval, as more complex models might perform better in the future, if the data patterns evolve.**

**The current BON forecasting approach for CiC is around 50 percent less accurate than the best performing models proposed by the IMF for currency in circulation. The IMF team has backtested the BON current approach using out-of-sample performance metrics, and benchmarked it against the models included in the IMF framework. The backtesting exercise suggests using a simple 5-day moving average instead of more appropriate models could lead to a deterioration in accuracy – as measured by Root Mean Squared Errors (RMSE) – of around 50 percent (the results are similar when using a median-based metric). This result clearly demonstrates the interest of using a more modern and comprehensive forecasting framework.**

**The mission recommends using the forecasting framework daily and validating all the models every quarter or a few weeks after a structural break. The mission has operationalized the framework that can be executed within seconds for daily uses.** Models have to be validated at a regular interval, for instance every quarter, or a few weeks after a major structural break, since the data patterns might have changed substantially. Model validation is an important step to make sure that the choice of forecasting models is still the most appropriate. The forecasting framework has been entirely programmed under R, the leading free and open-source statistical software used in the forecasting industry. The IMF experts have designed a user-friendly infrastructure to help the authorities producing liquidity forecasts smoothly and quickly. The framework is divided into three main files, that handle: (i) data pre-processing; (ii) model validation; and (iii) forecasting. Splitting model validation and forecasting allows to gain time and increases flexibility, as model validation is time consuming and doesn’t need to be done every day; while the forecasting file can be executed within seconds. The programming infrastructure is entirely automatized and produces charts and tables on the fly for easy inspection and interpretation, with dynamic HTML reports. The code is fully documented, using best programming practices to limit errors.

**The next step is to incorporate these liquidity forecasts to calibrate the monetary operations.** This requires (i) estimating the optimal structural liquidity surplus – based on an estimation of the demand for commercial banks reserves at the central bank, (ii) calibrating the different monetary instruments and (iii) conduct operations accordingly. The authorities should also improve on data collection, data reporting and data infrastructure to ease operational work, reduce bottlenecks and mitigate operational risk.

**Table 1. Key Recommendations**

| **Recommendations and Authority Responsible for Implementation** | **Priority** | **Timeframe1** |
| --- | --- | --- |
| Implement the forecasting framework and obtain the set of forecasts, leveraging on the full range of models (BON) | High | Near-term |
| Validate the forecasts at a regular interval (BON) | Medium | Medium-term |
| Improve data collection, IT infrastructure, and data treatment following the good practices outlined in the TA report (BON) | Medium | Medium-term |
| Publish the methodology to strengthen the transparence and the credibility of the central bank (BON) | Medium | Medium-term |
| Follow-up with estimating the demand for commercial banks reserves at the central bank, estimating the optimal level of excess liquidity and calibrating the operations according to the liquidity forecasts (BON). The IMF can help the BON with this endeavor | High | Medium-Term |

**1***Near term: < 12 months; Medium term: 12–24 months.*

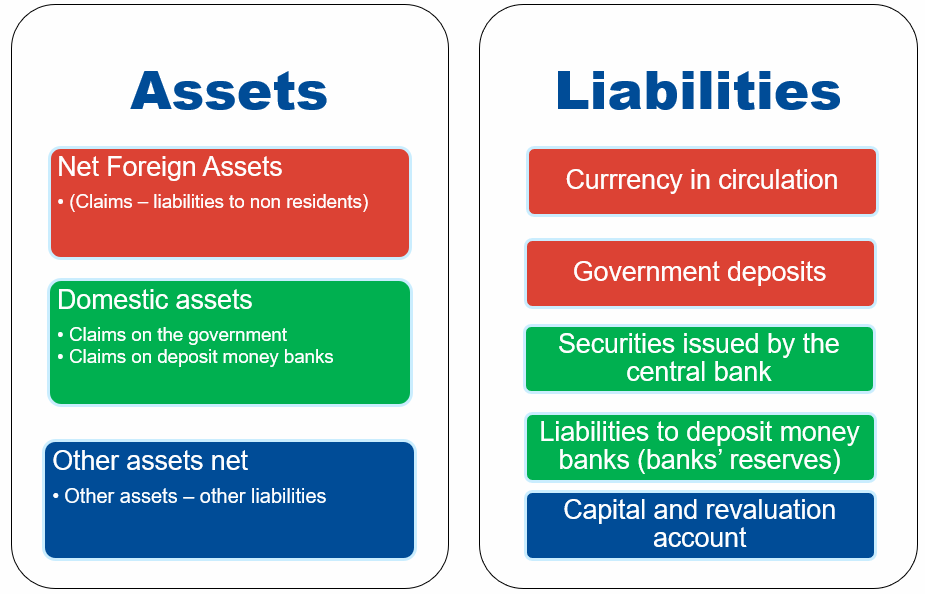
# Introduction

1. **The Bank of Namibia (BON) operates a fixed currency peg to the South African Rand since 1993.** TheNamibia dollar can be converted at a 1:1 exchange rate with the Rand. The final objectives—price stability and financial stability—are achieved by stabilizing imported inflation from the main trading partner, as Namibia is a small open economy.
2. **There are no capital controls in Namibia, and foreign assets have to constantly adjust to preserve the parity with the Rand, complicating BON structural liquidity management.** Most local banks are subsidiaries of South African banks, and FX transactions are settled at t+0 and often takes place with South African counterparts. Capital is therefore fully mobile, which complicates liquidity management, as capital flows could translate into large domestic liquidity swings in Namibia, possibly harmful to financial stability.
3. **The BON operational framework relies on a comprehensive set of monetary instruments** **to manage the excess liquidity.** On top of the level of foreign reserves, which anchors the credibility of the peg, the BON main policy instrument is the interest rate paid on excess reserves deposited at commercial banks’ settlement accounts at the BON. The interest rate on excess reserves is changed at the discretion of the central bank and is used as “a prime intervention instrument to control short-term capital flows” (Bank of Namibia, 2020). The Bank of Namibia also offers repurchase transactions, with seven-day, overnight and intraday repos. These facilities are designed to help banks meeting their short-term liquidity requirements. The Bank of Namibia also issues bills, for sterilization and liquidity management purposes, with maturities of 7, 14, 21, 28, and 56-day.
4. **Accurate forecasts of the autonomous factors are crucial to calibrate the central bank operations adequately.**The central bank aims to maintain a stable level of excess liquidity to guarantee a smooth transmission of monetary policy while preserving financial stability. An excessive level of structural liquidity surplus would hamper market development, thereby undermining monetary policy transmission, while an insufficient level would spur interbank rate volatility and endanger financial stability. Therefore, it is critical for the central bank to forecast the autonomous factors accurately to be able to anticipate and maintain an adequate level over time.
5. **The central bank currently lacks a liquidity forecasting framework for operational use.**The liquidity position of the central bank depends, in the case of Namibia, of three main autonomous factors: the currency in circulation, the state account at the central bank, and the net foreign assets. Currently, the BON only projects currency in circulation based on past five-day moving average, and doesn’t factor seasonality nor trend analysis. Besides, the CiC forecasts are realized once per month, which is not frequent enough for operational purposes. The Namibian Treasury shares its own forecasts on future balances of the Treasury account at the central bank, that is used directly by the BON. However, the BON doesn’t benchmark nor correct the Treasury forecasts from potential bias. Finally, the level of the BON net foreign assets is very volatile due to the fixed exchange rate arrangement and full capital mobility. The BON currently doesn’t forecast net foreign assets, but can, from time to time, use exogeneous information on diamond sales to anticipate some capital flows. There is therefore ample room to develop a full-fledged liquidity forecasting framework that can be used on daily basis for operations.
6. **The IMF has developed a new central bank liquidity forecasting framework and applied it on Namibia’s data.**This new framework leverages on the most recent advances of the statistical literature and is fully automatized. It encompasses the forecasts of the three main autonomous factors – CiC, SAB and NFA – and incorporates a reconciliation algorithm which aggregates the three forecasts into a single liquidity metric. Rather than relying on a single model, the IMF team has tested many different forecasting models, appropriate for each autonomous factor. The models are then benchmarked using out-of-sample performance metrics, which is considered as best practice in forecast evaluation. The framework is based on a two-tier approach, with the estimation of single models and of model combination/dynamic model selections for improving accuracy and reducing modeling risks. The codes are operationalized via a split between model validation and forecasting, with the forecasting part executing within seconds for daily uses.
7. **The current BON forecasting framework is estimated to be 50 percent less accurate than the best performing models used by the IMF for currency in circulation. The IMF team has backtested the BON current approach using out-of-sample performance metrics, and benchmarked it against the models included in the IMF framework. The backtesting exercise suggests using a simple 5-day moving average instead of more appropriate models could lead to a deterioration in accuracy – as measured by Root Mean Squared Errors (RMSE) – of around 50 percent.**
8. **The mission recommends using the framework daily and validate the models every month.**The models validation results suggest that, for CiC, SARIMA with trigonometric seasonality is the best individual model to use; for SAB, the TBATS model performs very well. For the NFA, the model with the best out-of-sample performance is a simple Exponential Weighted Moving Average (EWMA). These models offer an excellent ratio between robustness and accuracy, and are simple to deploy and estimate. However, the BON can improve its forecasts further by using model averaging or dynamic model selection, at the cost of extra complexity though.
9. **The rest of the technical report is organized as follows.**Section II presents the main principles of the forecasting framework designed by the IMF team. Section III discusses the modeling of currency in circulation, while section IV and V cover the state account balance and the net foreign assets. These three sections are articulated around a presentation of: (i) data patterns and data pre-processing; (ii) model validation via out-of-sample performance; and (iii) forecasting. Section VI discusses the results and formulates the recommendations and concludes. Appendix I covers the best practice for data management and reporting; Appendix II presents the forecasting models and the statistical methods used during the TA mission; Appendix III presents the software infrastructure documentation and Appendix IV is a short user guide for the R language.

# Forecasting Framework

1. **The mission designed a complete forecasting infrastructure to project the three main autonomous factors and reconcile them into a single liquidity metric.** The central bank balance sheets comprises: (i) items that the central bank controls, which are used to reach the policy position, for instance securities issued by the central bank or bank reserves, (ii) autonomous factors, which varies outside of the control of the central bank;[[2]](#footnote-3) and (iii) other items. The central bank has to forecast the autonomous factors to adjust its policy instruments accordingly, and maintain the optimal level of structural liquidity in the banking system. The framework designed by the IMF covers these the three main autonomous factors: the net foreign assets, the currency in circulation and government deposits.

Figure 1. Analytical Balance Sheet of a Central Bank (Autonomous Factors in Red, Policy Position in Green, Other Items in Blue)

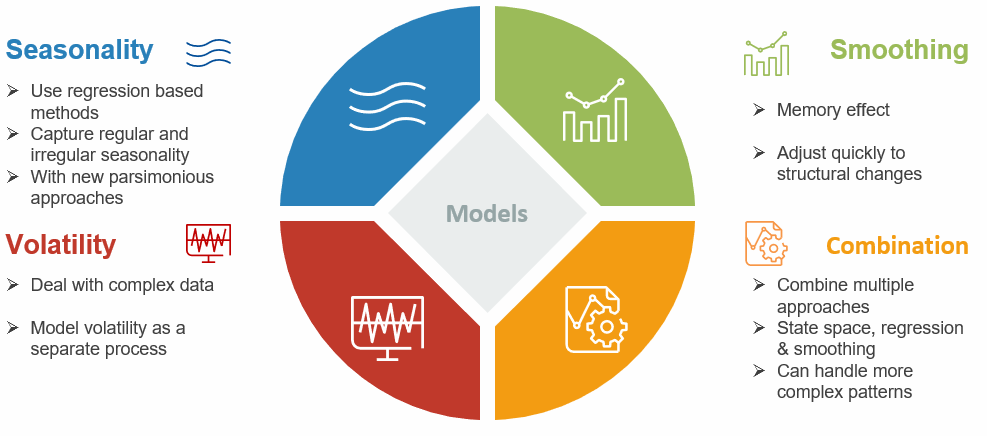


Source: IMF staff

1. **Forecasting accurately the autonomous factors has important operational and financial stability implications for the central bank.** First, accurate forecasts of the autonomous factors are necessary to calibrate central bank operations precisely. In particular, it helps to calibrate structural operations, such as the issuance of sterilization bills or medium-term repurchase agreements, and therefore reduces the frequency and volumes of fine-tuning operations. Second, the central bank can smooththe impact of large liquidity swings stemming from the autonomous factors and therefore strengthens financial stability. Third, accurate forecasts help to maintain an optimal level of liquidity which optimize monetary policy transmission while preserving market development, as excessive amount of liquidity surplus are detrimental to the interbank market.

1. **The modeling framework needs to account for the different macro-financial drivers impacting the autonomous factors.** Currency in circulation results primarily from the transactional demand for banknotes and coins. The overall economic activity—trade, transactions, etc. —as well as seasonal factors (holidays, religious festivities, etc.) and structural factors (digitalization) are impacting currency in circulation. The state account balance depends on the Treasury cash-flows managements (variation in revenues and expenditures) as well as non-central bank borrowing from domestic and international markets. Finally, central bank net foreign assets are primarily impacted by current and capital flows,but also structural domestic factors such as the degree of dollarization and   
   de-dollarization. This variety of drivers is calling for a differentiated treatment in the forecast of each autonomous factor.
2. **The mission implemented a flexible, multi-model forecasting framework.** Because currency in circulation is strongly impacted by seasonal factors, potentially combined with a trend, the framework deploys time-series models specifically adapted to treat such patterns. Seasonal Auto Regressive Integrated Moving Average (SARIMA), exponential smoothing models (ETS) and Trigonometric Seasonal, Box-Cox Transformation, ARMA residuals, Trend and Seasonality (TBATS) models are therefore estimated for currency in circulation. Regression-based methods such as SARIMA can capture both regular and irregular seasonality, with more or less parsimonious approaches (depending on the type of modeling between binary and trigonometric seasonality for instance). Smoothing-based types of models, such as ETS, weights more recent more than older ones, and are therefore more robust to structural breaks than models weighting observations equally. More complex models, such as the TBATS, combine multiple approaches (state space, regression and smoothing) to handle more complex patterns. The state account balance also exhibits seasonal patterns and the mission uses the same type of seasonal models as for currency in circulation. The main difference between SAB and CiC though is that the Namibia Treasury shares with the central bank its projection on future Treasury cash flows that the central bank can include to fine-tune its SAB projections. Finally, the net foreign assets are typically extremely volatile in fixed exchange rate arrangements, as they reflect the constant capital flows adjustment necessary to preserve the fixed parity with the South African rand. To project the net foreign assets, models with a focus on volatility modeling are necessary, and the framework deploys Exponentially Weighted Moving Average (EWMA) and Generalized AutoRegressive Conditional Heteroskedasticity (GARCH) models. The specification and estimation procedure of these models are presented in the Appendix II.

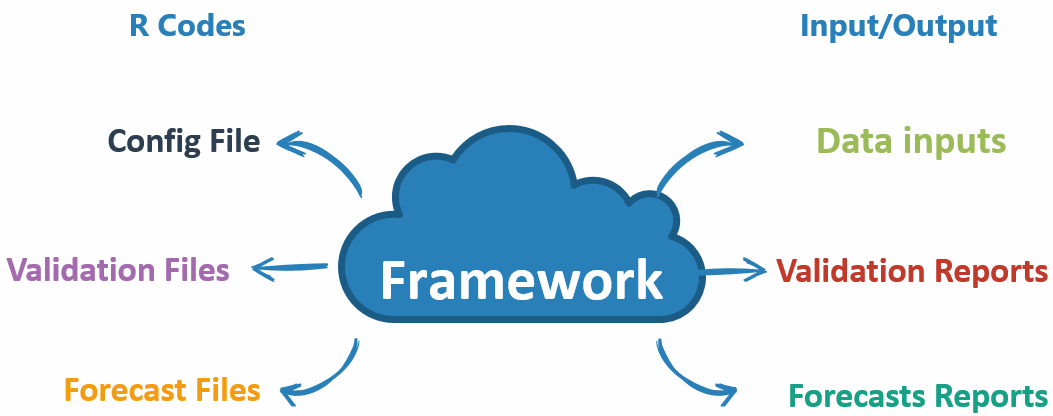
Figure 2. Family of Individual Models Used in the IMF Forecasting Framework



Source: IMF staff

1. **The diversity of models allows the IMF team to decide among statistical trade-offs using a data-driven approach.** Forecasting entails some important statistical trade-offs in the choice of models, particularly between forecastability/robustness and complexity/accuracy. Forecastability allows models to handle different data patterns equally well and to perform reasonably well on average, with limited risk of under-performance. Models robusts to outliers and structural breaks typically exhibit a strong forecastability with limited risk of data overfit. The objective is to rely on more complex models only if there is sufficient evidence of benefits to forecasting. More complex models are typically better at fitting to training data and capture more complex patterns, due to their increased flexibility in parametric estimation. However, this extra flexibility can lead to overfitting and substantially lower quality of forecasts. Another side effect of the increased complexity/number of parameters is a potentially increased volatility of the forecasts over time, over different sample origins, as new data become available. It is difficult to know *ex-ante* which models to select to optimize these tradeoffs, and the IMF forecasting framework put in place a data-driven decision approach, where a battery of different models – from simple to complex ones – are estimated and tested. The best performers models, based on specific criteria presented below, can therefore be used as the primarily projection method for forecasting.
2. **The modeling diversity is increased further by the estimation of a two-tier empirical strategy, where single models are tested on the first layer and dynamic selection/model averaging are estimated on the second one.** The single models described above (SARIMA, ETS, TBATS, GARCH, etc.) are first estimated on the data and constitute the first modeling tier. The second-tier builds on the first-tier to improve the forecasting accuracy of the single models, either by combining them or selecting them. A combination of models averages the forecasts of different models (potentially with different weights) to improve the overall accuracy; some models might be performing better than others over certain periods, and aggregating them reduces the modeling risk (i.e., trusting a single model for the forecasts). On the other hand, model selection consists in selecting the most appropriate model based on its recent performance. For instance, if last week a given model outperformed all the alternatives, it will be chosen for this week, etc. The model selection method quickly adjusts to structural breaks by shifting towards the most appropriate model when necessary.
3. **The framework uses a battery of forecasting performance metrics and information criteria to select the most appropriate model.**The team uses out-of-sample performance metrics to benchmark the models and choose the most appropriate one. Using out-of-sample metrics is crucial to reduce the risk of overfitting and provides an accurate picture of the model's actual performance (see Hyndman and Athanasopoulos (2021) for a thorough discussion). The IMF framework currently proposes three different out-of-sample performance metrics: Root Mean Squared Errors (RMSE), Median Average Errors (MAE) and the Hit-Ratio. These metrics are presented in detail in Appendix II.
4. **The framework reconciles the forecasts of the three autonomous factors to provide a single liquidity metric forecast.** The total autonomous liquidity position is a combination of the three autonomous factors: liquidity position = NFA – CiC – SAB. While it is possible to infer directly the liquidity position by naively summing/subtracting the three forecasts, the literature has developed methods to handle this issue more efficiently. The idea is to combine the subcomponents single forecasts (bottom-up) with the forecast of the overall liquidity position (the “base”) by different algorithms (OLS, min Trace, etc.). This approach, called forecast reconciliation, is explained in details in Hyndman et al. (2011).
5. **The framework also reflects on estimation risk by providing the confidence interval associated with each forecasting model.** Themodelingframework delivers, on top of the forecasted future values, the confidence interval associated with the forecasts. For single models, the confidence intervals are usually derived through the asymptotic properties of the models. For averaged and reconciled models, the confidence interval are obtained via bootstrapping. Confidence intervals represent an important information for the forecaster, as they indicate the accuracy of the forecasts and gauge the risk of error. Confidence intervals can also be used to calibrate specific central bank instruments that can help the market to absorb the forecasting errors, for instance, by appropriately calibrating required reserves with averaging.
6. **The software infrastructure automatizes the testing and forecasting processes in three main steps for fast execution and easy use for central bankers.** The framework is based on high-quality, open-source, free R codes. The codes are split in three files. The config file allows the user to easily change the parameters, such as the forecasting horizon, the training sample, the holidays, etc. The validation file estimates and tests each model out-of-sample and presents the performance metric for the user to decide upon the best model. Finally, the forecasting file produces the forecasts depending on the parameters chosen by the user. Splitting the codes between forecasting and validation helps speeding up the forecasting process, as model validation can take a few hours to execute and does not need to be estimated every day. The forecasting file alone can execute within a few seconds. The IMF team recommends validating the models every quarter, or a few weeks after a major structural break. The validation and the forecast files also produce results reports, encompassing HTML-based dynamic plots for data and results visualization. The main output is also exported on Excel tables for being used for operational purposes. Data input is excel-based, with pre-filled templates, for easy and convenient data feed.

Figure 3. Software Infrastructure

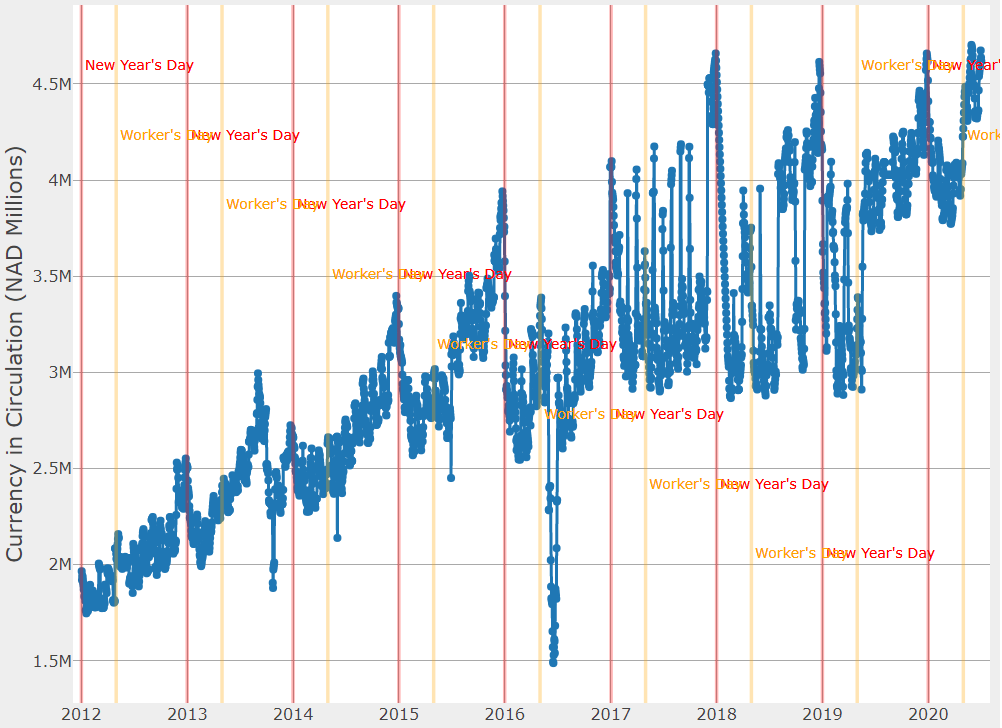


Source: IMF staff

# Forecasting Currency in Circulation

1. **The treatment of missing values matters to model seasonality properly.** It is important to ensure that the time series is continuous in terms of dates, with no omissions. Continuity is essential for the modeling of any potential seasonality in the time series. With omitted dates, the periodicity between observations is broken, making the estimation of the seasonal terms very difficult. Therefore, when encountering missing values, the IMF team imputed them. There are multiple ways to impute missing values. The mission relies on linearly interpolating from neighboring values. Although this is not necessarily the best approach, the scarcity in missing values suggests that the method's robustness and simplicity counterbalance any minor losses in accurate imputations. Another alternative would be to provide indicator variables for each of the missing values to the models. Mathematically this is equivalent to excluding the observations. The IMF team also **removed duplicated dates (by default, the lowest of the two available values is taken as the main value) and dropped data incorrect entries during weekends. Appendix I presents the good practices in terms of data management, data treatment and data reporting that the BON should follow.**
2. **The IMF team treated structural breaks in CiC time series. The CiC series contains structural break on January 4, 2016, and May 20, 2019. Both of these structural breaks are handled by including a dummy variable equals to zero before the break and one after the break. In effect, this allows for a level shift in CiC. Note that structural breaks have been immediately spotted via visual inspection as clear rupture in the time series. Going forward, the BON could use a statistical approach (e.g. via Chow, CUMSUM, Wald or Likelihood ratio tests, etc.) to systematically detect and correct for the structural breaks, which complicate and potential bias models’ estimation.**

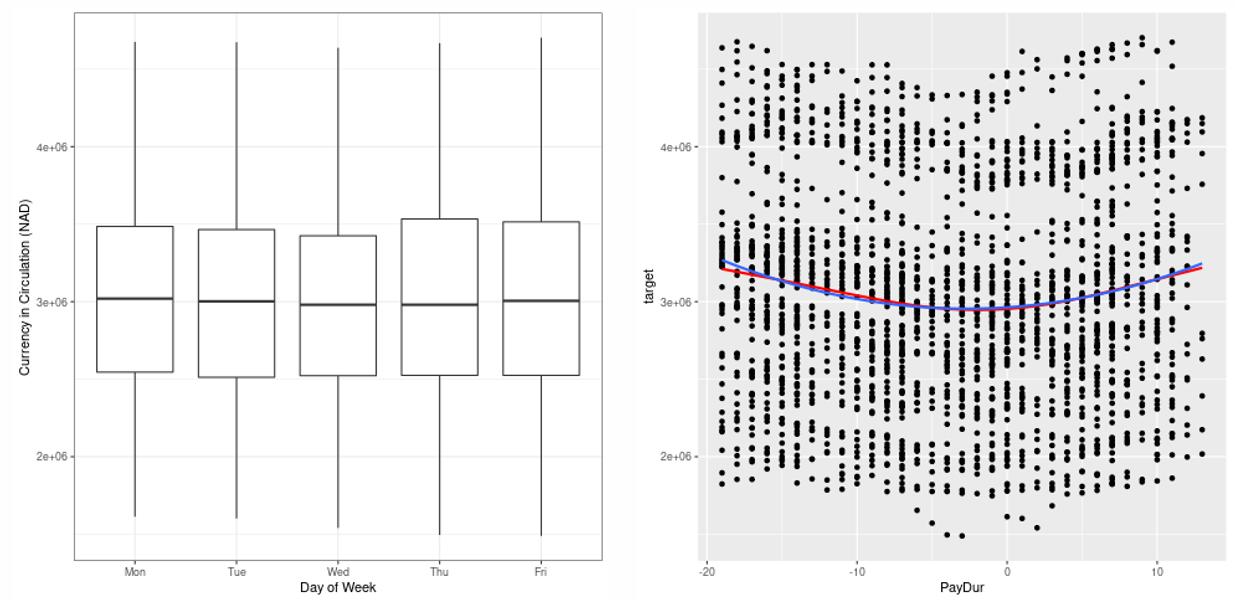
Figure 4. Currency in Circulation Over Time



Source: BON and IMF staff

1. **Currency in circulation in Namibia exhibits monthly seasonality around pay day, at the end of the month.** Figure 5 **shows an analysis of the seasonal patterns for currency in circulation at daily and monthly frequency. While there is no clear pattern at the daily frequency (for instance, a surge in currency in circulation on Friday), the pay day (the 28 of each month) generates clear seasonal patterns.**

Figure 5. Seasonality Patterns Analysis: Daily (left) and Distance to Pay Day

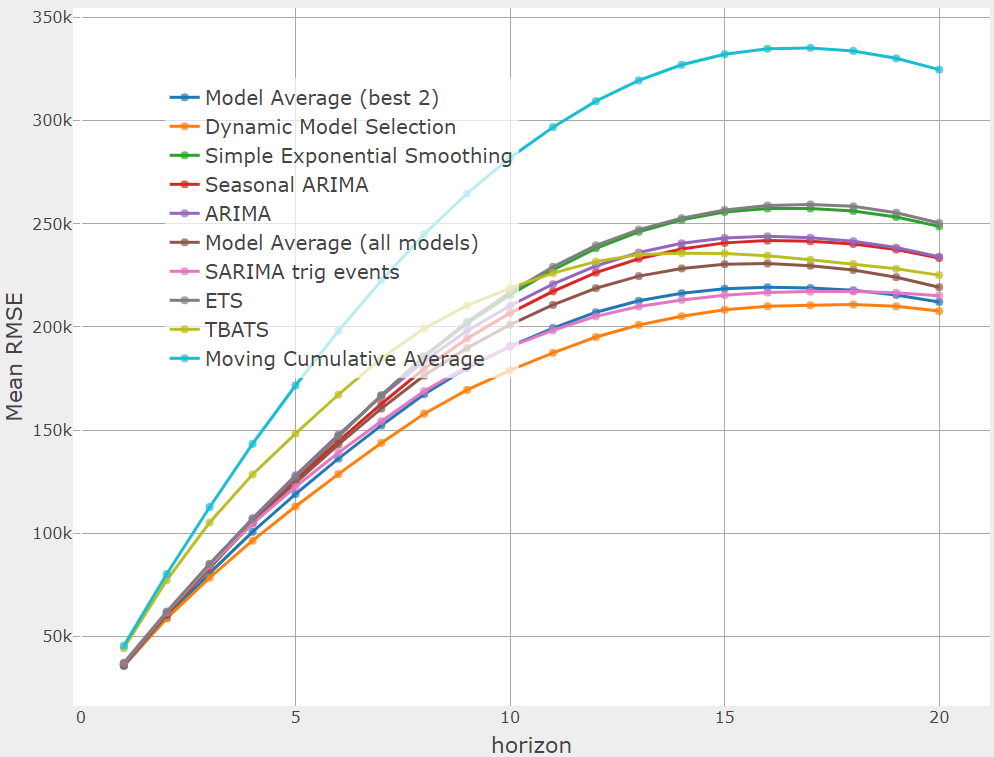


Source: IMF staff

1. **These seasonality patterns should be captured by appropriate statistical models. Given the regularity of pay days in Namibia, models incorporating seasonality treatment performs typically well. These models include Seasonal ARIMA (also called ARIMA with regression), TBATS (Trigonometric Seasonal, Box-Cox Transformation, ARMA residuals, Trend and Seasonality) and ETS with seasonal treatment. A presentation of these models is available in Appendix II.**
2. **The IMF model validation confirms that the SARIMA model with trigonometric seasonality is the best single model in terms of performance, while dynamic model selection has the best overall accuracy; the current methodology used by the BON is around 50 percent less accurate than these two models. The framework tests the different models out-of-sample over the period 2019–11–29 to 2020–07–02 and computes RMSE and MAE as performance metrics over nine different models: seven single models (ARIMA, SARIMA with binary seasonality, SARIMA with Trigonometric Seasonality and Events, ETS non-seasonal, ETS seasonal, TBATS) as well as two composite models: model average (top two single models) and dynamic model selection. The results are similar for both the RMSE and the MAE metrics:**

* **The dynamic model selection is the model with the lowest RMSE and MAE across all horizons. Because the framework is always using the best performing model over the recent periods, dynamic model selection usually performs very well if the data is stable enough, without major structural breaks. The downside of dynamic model selection is that it complicates the use of the framework, as the models used can switched by horizon and by estimation period. This approach is appropriate for users only interested in performance.**
* **The model averaging performs second, and is parametrized to average the two best individual models for each horizon. Interestingly, the average of the two best models always outperform single models, suggesting that combining even the best single model with a slightly less performing one is beneficial, as it increases modeling diversity, reducing modeling risk and forecast errors.**
* **Among single models, the SARIMA with trigonometric seasonality and events is the best performing model. It is a standard model used in the literature for forecasting seasonal patterns. The use of trigonometric terms makes is quite parsimonious and reduce parametric noise while increasing accuracy. Should the BON want to use a single consistent model over time, it should implement a SARIMA.**
* **The model currently used by the BON (the moving cumulative average), is around 50 percent less accurate than the SARIMA trigonometric model with events, at a 10-day horizon.**

Figure 6. Out-of-Sample Performance by Models, Using the RMSE Criterion, at Different Horizons. Results are Similar for the MAE Metric

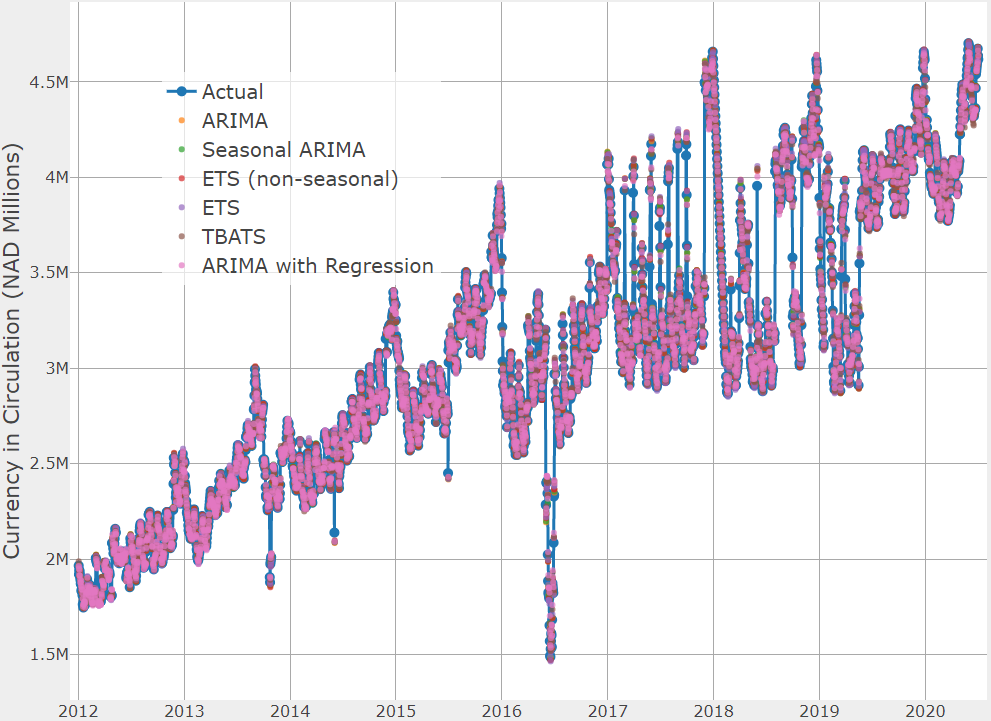


Source: IMF staff

1. **The performance gain offered by dynamic model selection and model averaging is not substantial and the BON could keep a simple SARIMA as the main forecasting tool. Except at long horizons**—**over two weeks**—**the performance of a SARIMA is very closed to the dynamic model selection, suggesting that the accuracy gains of moving to second-level models (averaging and selection) do not compensate for the complexity costs.**

1. **The out-of-sample fit of the currency in circulation models are accurate over the one-month horizon. Figure 7 presents the out-of-sample fits, realized over a one-month horizon, for each model. The data in blue are the true data. Generally, the forecasting models have a quite good quality of fit, except during sudden peaks and through – for instance over 2017-2018 - that are typically difficult to anticipate with relatively simple linear models.**

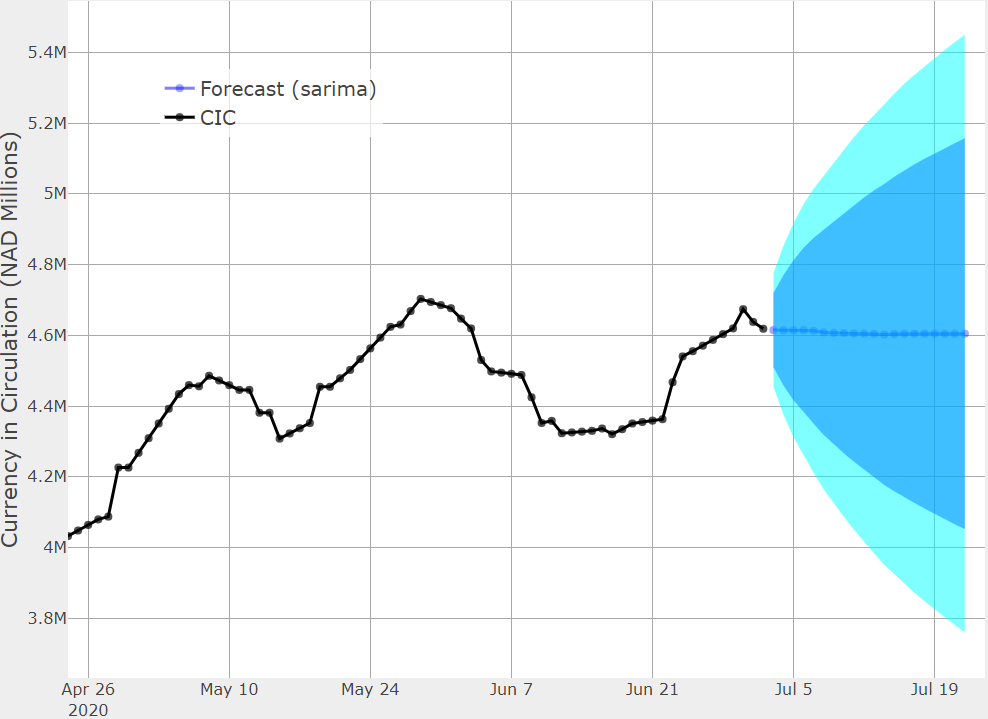
Figure 7. Out-of-Sample Models fit for Currency in Circulation



Source: IMF staff

1. **The projected confidence interval helps the authorities to assess the forecasting risk level.** The forecasting framework provides both the point forecasts and the associated confidence interval over the forecasting horizon. Figure 8 presents the projection of the best single model forecast (a SARIMA), with confidence interval at 80 percent (dark blue) and   
   95 percent (light blue). Although the confidence interval seems quite large, it reflects the large variation experienced over 2017–2018 (see Figure 7).

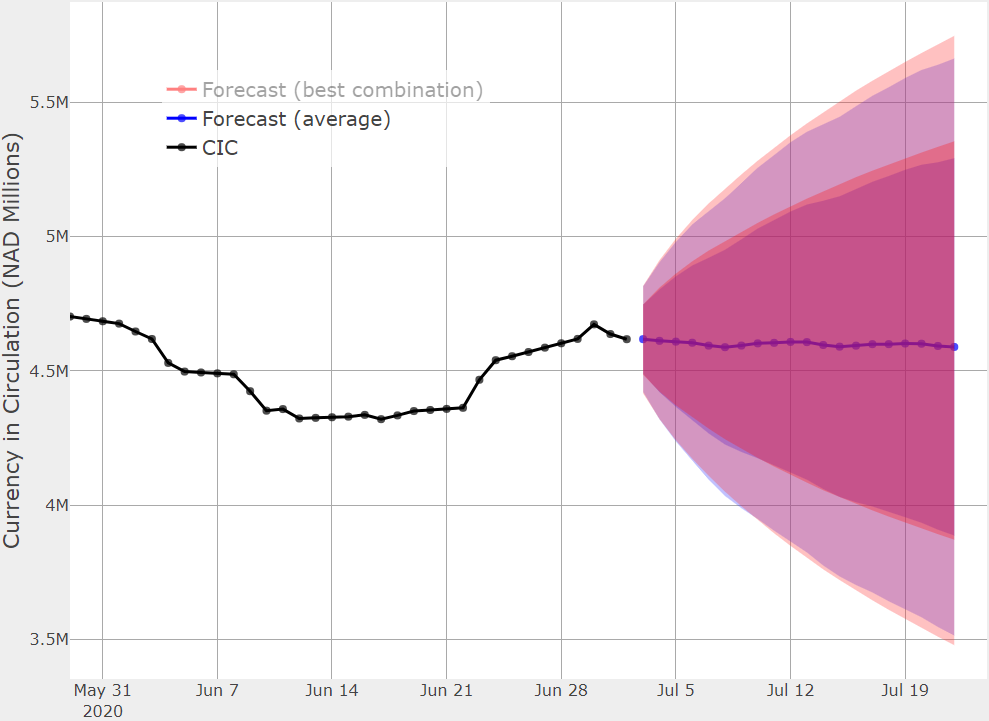
Figure 8. Best Single Model Forecast With Confidence Interval



Source: IMF staff

1. **The forecasts can also be obtained using aggregate models, for instance models average.** The framework offers the possibility to average models or to dynamically select them to forecast currency in circulation. Figure 9 presents the forecast of the model average, including the confidence interval bands at 80 and 95 percent respectively.

Figure 9. Forecasts Derived from Best Models Average

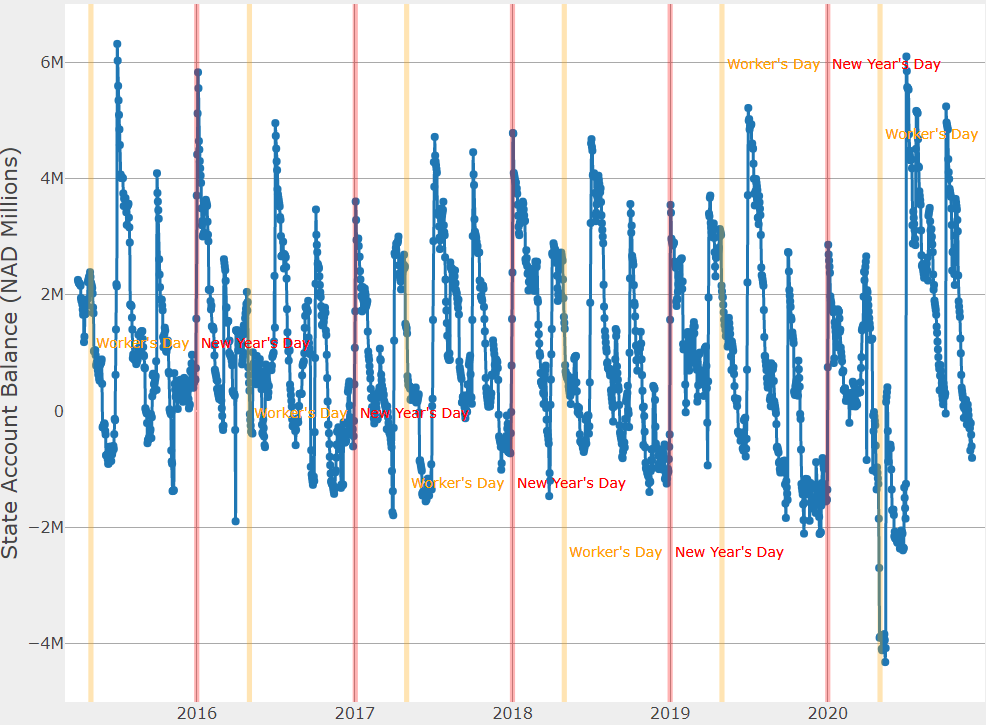


Source: IMF staff

# Forecasting the State Account Balance

1. **Although the Treasury provides forecasts on the state account balance to the central bank, it is beneficial for the BON to forecast the state account balance independently. The central bank has signed a Memorandum of Understanding with the Treasury, and the exchange of information between the two institutions is frequent and of good quality.** That being said, the central bank should produce its own SAB forecasts, based on historical time series and statistical models. First, there is an operational risk, as the Treasury might encounter some issue and not be able to deliver the forecasts on time for the operations. Second, the central bank should be able to assess the deviation of SAB forecasts from statistical patterns to enquire with the Treasury the reason of this discrepancy. For instance, it might be the case that the Treasury forecast is far away from the central bank one, for ad-hoc and specific reasons. Benchmarking Treasury forecasts and assessing the Treasury forecasting error is a good practice for central banks to improve their own liquidity forecasts. Going forward, the central bank could integrated the Treasury forecasts into its own forecasting framework, for instance, by using the Treasury forecasts as an exogeneous regressor or as a Bayesian prior. However, it is important to correct first for potential systematic biases in the Treasury forecasts before using them in the model. This requires to systematically archive and store the different vintages of the Treasury forecasts over time to be able to conduct statistical analysis on them.
2. **The team proceeded to a similar data treatment on the SAB time series as the one done for CiC. The SAB time series suffers from the same issues as the CiC time series. Hence, missing values are interpolated, duplicated observations are removed, the models are fit on a five-day week, etc.** Figure 10 **presents the time series of state account balance over the last six years.**

Figure 10. State Account Balance Seasonal Patterns



Source: IMF staff

1. **The SAB exhibits very regular quarterly seasonality, driven by taxation payment.** Figure 11 **presents the regression fit of a quarterly trend on the state account balance time series. The fit is very good and confirms the quarterly seasonal patterns. The SAB spikes at the beginning of the quarter (when taxes are collected) and linearly decline over time (as money is spent).**

Figure 11. Quarterly Seasonality in Namibia

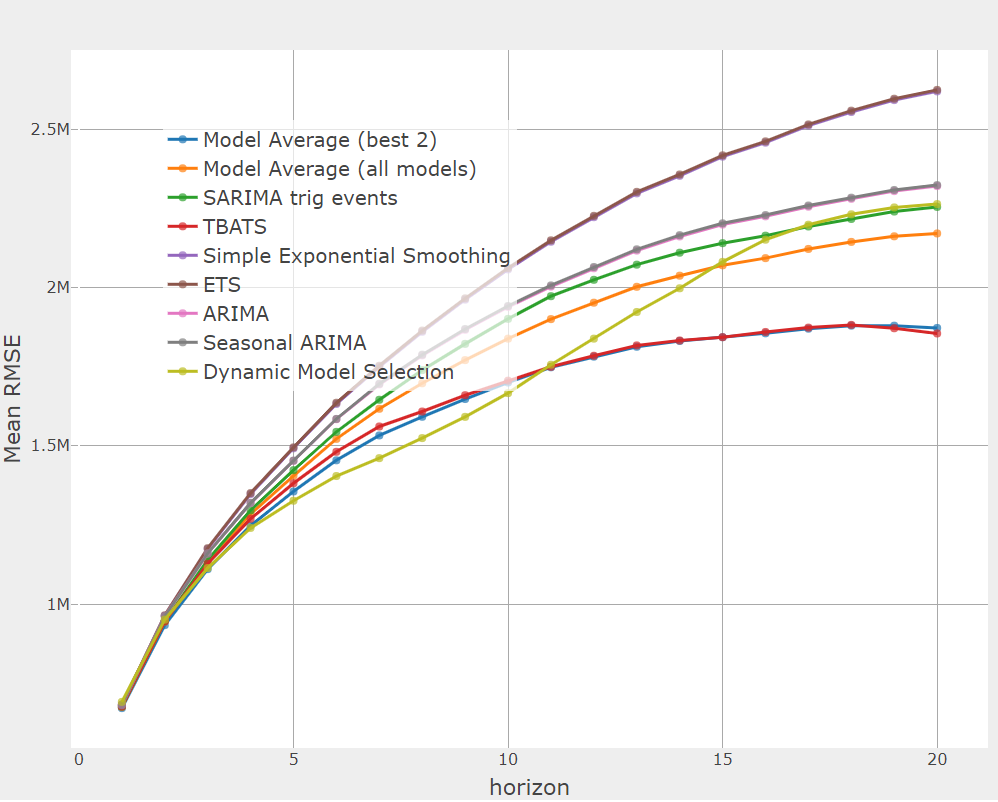
Chart, scatter chart

Description automatically generated

Source: IMF staff

1. **Model validation suggests that, among single models, the TBATS model is the best performer.** The TBATS model is a model capturing different features, including non-normality, seasonality, trend and recursive errors. Model averaging is a very solid performer over the time horizon, and is the best performer at horizons below 10 days. These results, presented in Figure 12, shows the “crossing” between models at short and long-term horizons. The out-of-sample performance analysis follows the same principles and metrics as those done for CiC.

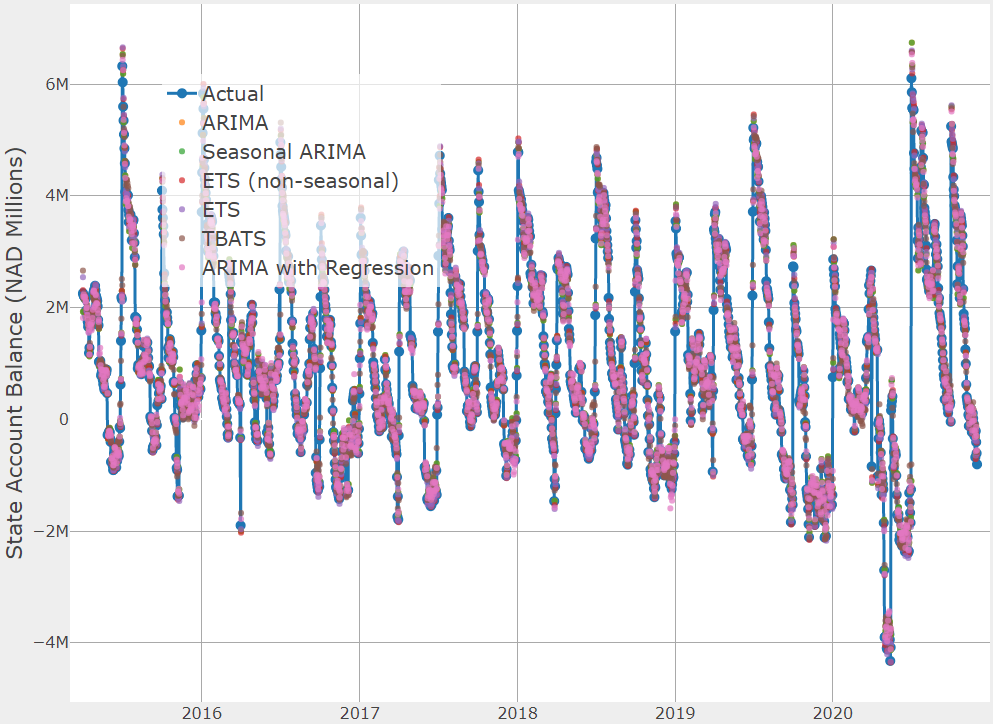
Figure 12. Out-of-Sample Performances, SAB Forecasts



Source: IMF staff

1. **The SAB forecast models offer a relatively accurate out-of-sample fit and captures well most of the quarterly seasonality.** Figure 13 presents the out-of-sample fits of the different models, including with the actual value (blue line with dotted markers). The quality of the out of sample fit is even better than for CiC, as the big spikes due to quarterly taxation are quite well anticipated.

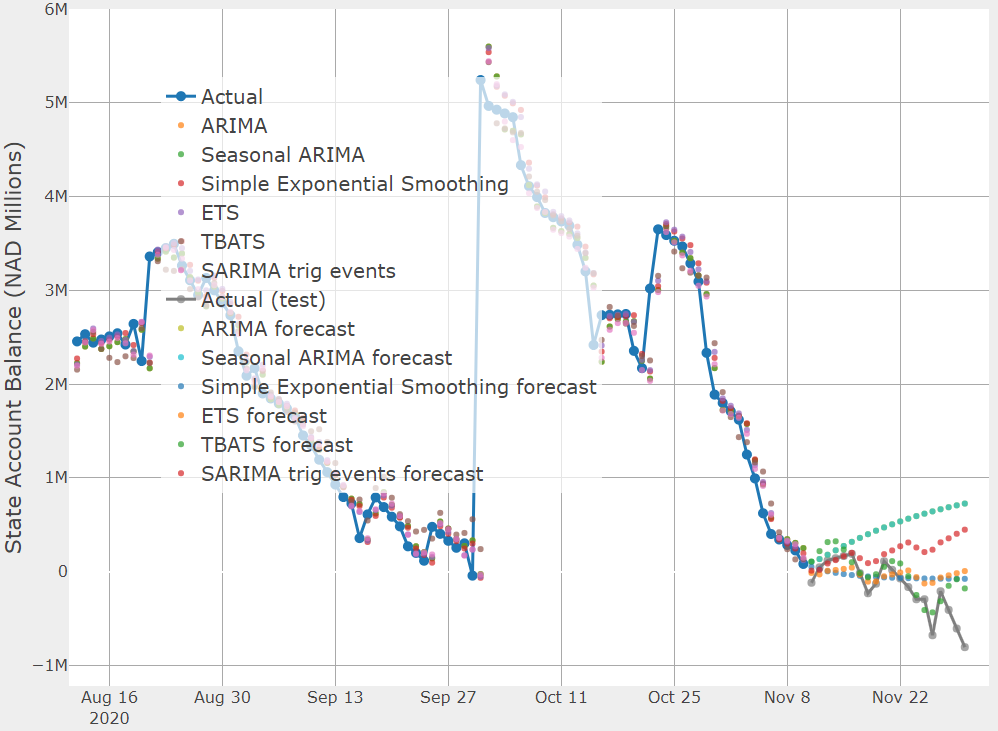
Figure 13. Out-of-Sample Fit, SAB Forecasts



Source: IMF staff

1. **SAB forecasts exhibit a relatively large dispersion over models, suggesting a potential benefit of using model averaging to reduce modeling risk.** Figure 14presents the out-of-sample forecasts produced by the all range of models, which shows a relatively wide dispersion. Some of these models can have poor performance and should be discarded. However, among the best performing ones (as presented in Figure 12), an average of the two best models would help mitigating modeling risks, as it the average is the first performer for horizons below 10 days. Above 10 days though, using a unique TBATS, without averaging, offers better performances.

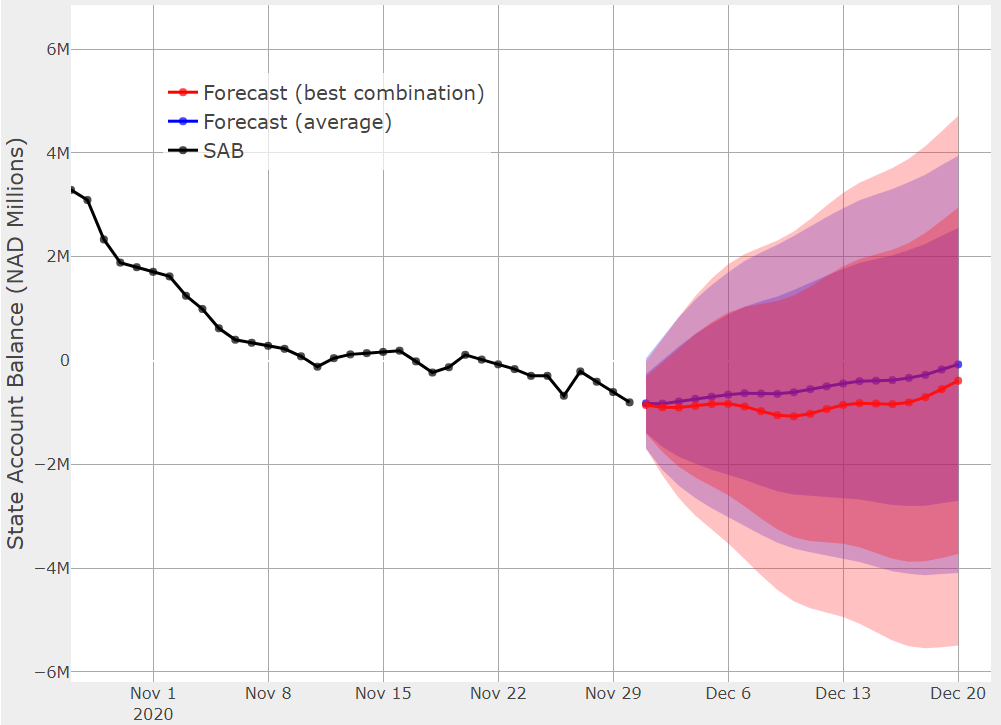
Figure 14. Diversity in SAB Forecasts Across the Range of Models



Source: IMF staff

1. **As for CIC, SAB point forecasts and confidence interval are produced by the IMF framework and can be used to assess the deviation from the Treasury forecasts.** Figure 15presents the forecasting output of the IMF framework. The advantage of producing confidence interval is that the model can now benchmark the probability of the Treasury forecast to be significantly different from the seasonal forecast (ETS, TBATS, etc.), at a given risk threshold, depending on the confidence interval. A value in outside of the   
   95 percent range for instance could trigger a discussion and check-up by the central bank to understand the deviation.

Figure 15. Point Forecasts and Confidence Interval for SAB

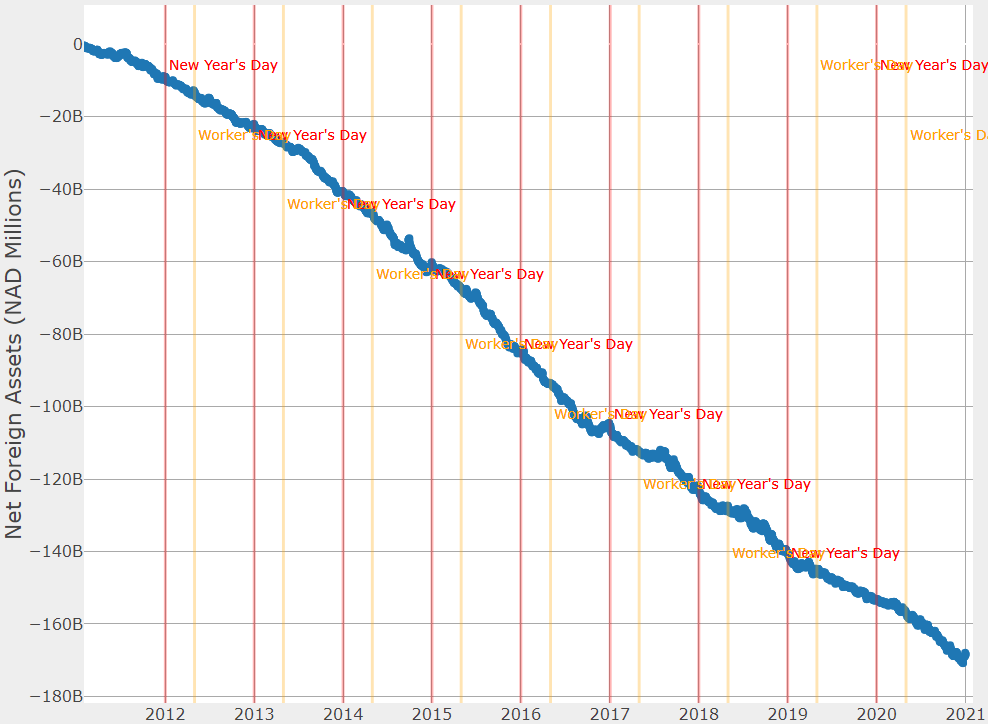


Source: IMF staff

# Forecasting Net Foreign Assets

1. **The BON provided Namibian Dollar (NAD) flows instead of NFA per say, and the mission works directly from the flows.** Forecasting by the stocks (NFA) or by the flows are two valid options, but NFA are necessary to operate forecast reconciliation for the total liquidity position.[[3]](#footnote-4)
2. **As for CIC and SAB, the IMF treated the flow data, which are of relatively good quality.** Similar data treatment were applied as described previously, with only a few observations with problems.
3. **While there is a clear trend in net flows, there are no important seasonality or day-of-week effects in the flows data.** At the difference of CiC—driven by cash demand during special events—and SAB (quarterly payment), there are no important seasonal patterns for currency flows. However, the net currency flows in Namibia have been constantly declining over time to reach NAD -180 billion at the end of the period, as shown in Figure 16.

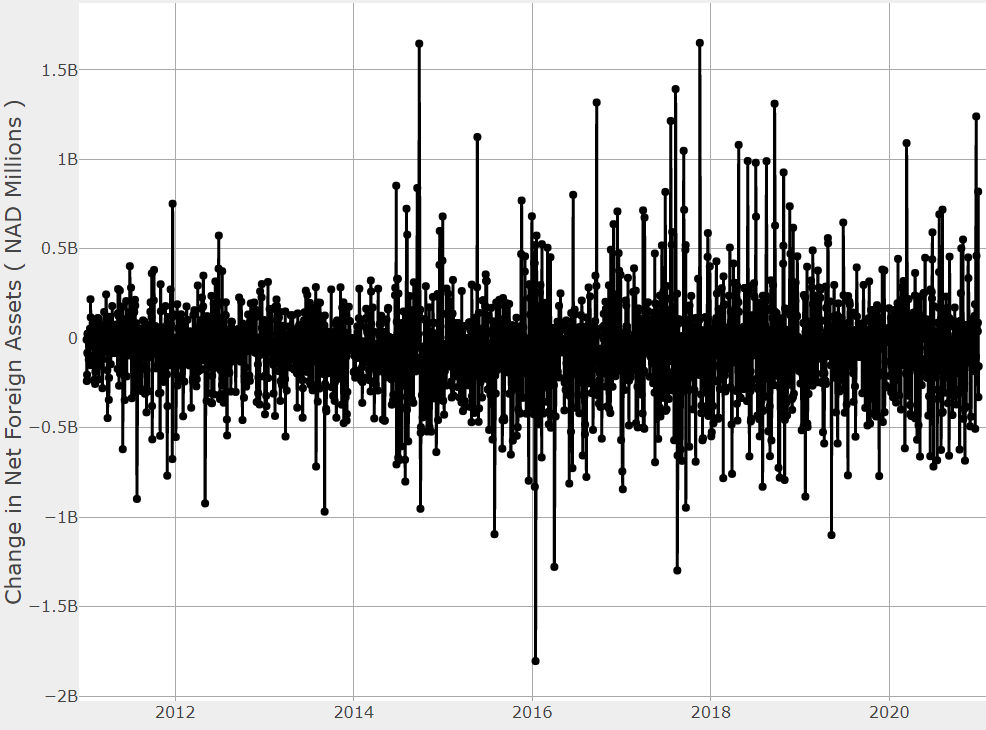
Figure 16. NAD Flows Over Time



Source: BON and IMF staff

1. **The NAD flows data experienced a structural break in mid-2014, with a sharp increase in conditional heteroscedasticity.** After this date, the day-to-day volatility of the net flows increased substantially, as shown in Figure 17.

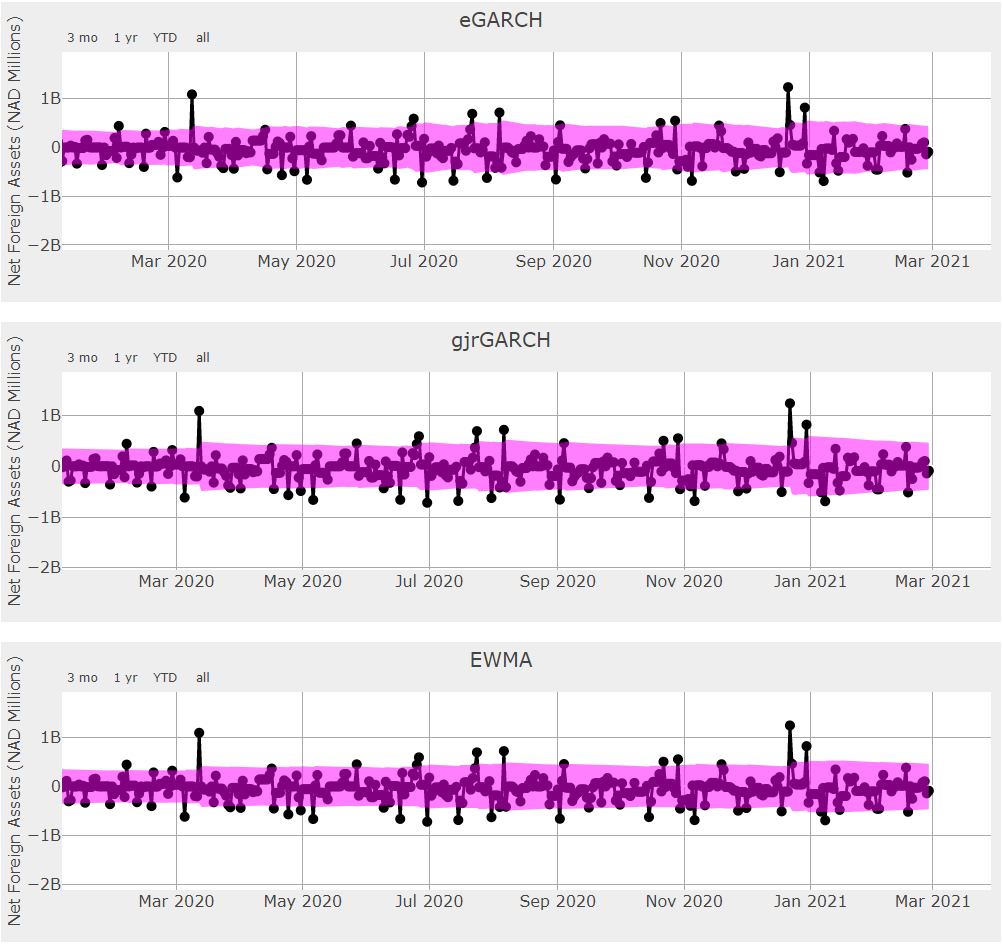
Figure 17. Change in NAD Net Flows



Source: IMF staff

1. **Contrary to CiC and SAB forecasts, the metric of interest for NFA flows is the quality of forecast coverage instead of the point forecast distance.** The NFA flows are extremely volatile, and, in this context, point forecasts might not be the most informative statistics for a central banker. The IMF team prefers to focus on forecasting coverage, which is defined as the frequency of realized values falling within a certain interval, per model. With highly volatile time series, standard performance metrics can be misleading as the best performing model might “luckily” hit a few times well but miss a lot. Central bankers should be more conservative and might prefer a model with more regularity, hence the use of coverage metrics via the hit ratio indicator instead of point forecasts performance metrics such as the RMSE or the MAE. Figure 18 presents the coverage ratio of different models used to forecast NFA flows. The pink area is the confidence interval of the model at 95 percent, while the black dots are the realized observations falling outside this confidence interval.

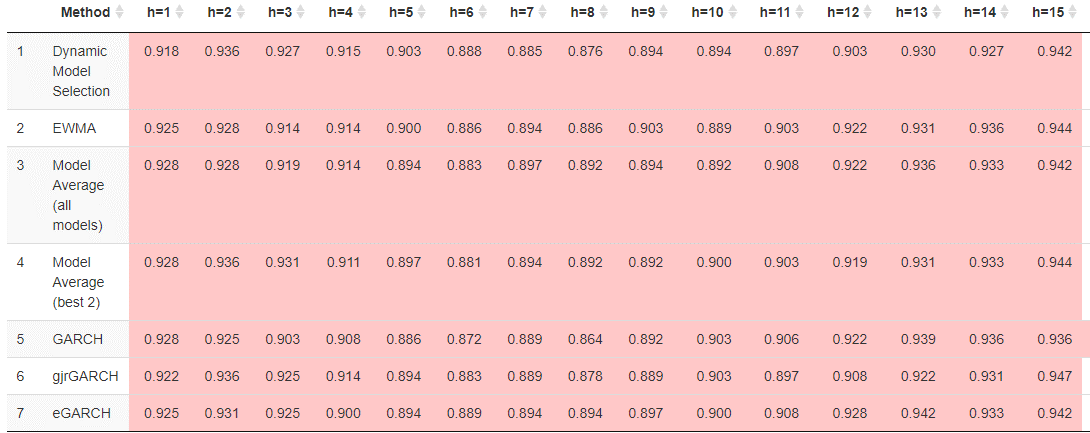
Figure 18. Coverage Ratio of NFA Flows Forecasts



Source: IMF staff

1. **The EWMA is the model with the best compromise between accuracy and simplicity for forecasting the NFA flows.** Using the hit-ratio coverage metric, the IMF team found that all models perform quite well with limited difference in performance among them. Hence, it makes sense to favor the simplest ones. Among single models, the EWMA has solid performances at 95 percent coverage, as shown in Figure 19. This result is consistent both with the structural break and the highly volatile nature of NFA flows, as a simple EWMA tends to include short-term persistence, which underweight old observations and tends to provide a reasonable guess by using a weighted average of the most recent values. The performance pattern is also consistent with findings on other countries NFA forecasting exercise, especially small open economies operating a fixed exchange rate arrangements without capital controls. Note that the RMSE and MAE metrics, although less appropriate for this type of time series, also indicate that the EWMA is a very strong performing model.

Figure 19. Out-of-Sample Hit-Ratio Coverage by Model at 95 percent Confidence Interval at Different Horizons



Source: IMF staff

1. **As for CiC and SAB, the IMF framework provides both the point forecast and the confidence interval.** However, the point forecasts are less informative and considering the confidence interval is more reasonable given the highly volatile nature of the series.
2. **Given that only flows and not stock are available, the IMF team has not been able to reconcile the forecasts of the three autonomous factors and retrieve a single liquidity metric.** Because it is not possible to add stock to flows, the IMF team will follow up with the authorities to obtain the NFA stock and re-run the analysis. The framework delivered to the authorities, however, is already able to handle the reconciliation, provided that the right data are inputted. The advantage of using a reconciliation approach to compute a single metric, instead of summing the three autonomous factors directly is that it improves the quality of the total liquidity forecast. The reconciliation approach leverages the information across the individual forecasts (the variance-covariance matrix of the error terms) to reduce both the in-sample error fit and the out-of-sample forecasting errors.

# Recommendations and Conclusion

1. **The best individual forecasting models are: SARIMA for CiC, TBATS for SAB and EWMA for NFA. The models have been validated thoroughly via a rigorous data treatment, multiple models tests, out-of-sample performance metrics, etc. Individual models actually performs very well and offer the best compromise between robustness and accuracy. Model averaging and dynamic model selection marginally improve accuracy, at the cost of extra complexity. However, it would be important to reassess the performance by validating all models on updated data at regular interval, as more complex models might perform better in the future, if the data patterns evolve.**
2. **The mission recommends using the forecasting framework daily and validating all the models every quarter or a few weeks after a structural break. The mission has operationalized the framework that can be executed within seconds for daily uses.** Models have to be validated at a regular interval, for instance every quarter, or a few weeks after a major structural break, since the data patterns might have changed substantially. Model validation is an important step to make sure that the choice of forecasting models is still the most appropriate. The forecasting framework has been entirely programmed under R, the leading free and open-source statistical software used in the forecasting industry. The IMF experts have designed a user-friendly infrastructure to help the authorities producing liquidity forecasts smoothly and quickly. The framework is divided into three main files, that handle: (i) data pre-processing; (ii) model validation; and (iii) forecasting. Splitting model validation and forecasting allows to gain time and increases flexibility, as model validation is time consuming and does not need to be done every day; while the forecasting file can be executed within seconds. The programming infrastructure is entirely automatized and produces charts and tables on the fly for easy inspection and interpretation, with dynamic HTML reports. The code is fully documented, using best programming practices to limit errors and bugs.
3. **The authorities have now a full-fledged infrastructure to generate forecasts of the three autonomous factors.** The framework is readily operational and has been backtested on real BON data. The BON should have the free and open-source R software installed on the computer to execute the IMF codes. We remain at the disposal of the BON to answer any questions they might have or to organize follow-up sessions to train further their staff.
4. **The authorities should publish the methodology and the forecasts of the autonomous factors to help market participants manage their liquidity better.** The autonomous factors not only impact the balance sheet of the central bank, they also affect the whole money market. While some market participants might be producing forecasts of their own, it is helpful that the central bank releases its own forecasts to help all the participants to improve their liquidity management better. The publication should also include the forecasts evaluation (produced automatically by the IMF forecasting framework) and quality testing for sharing a complete and robust analysis with the market.[[4]](#footnote-5) The central bank might want to first test the model for a few weeks before releasing the forecasts. The methodology should also be released, as it contributes to raising the technical level among market participants and reinforces the central bank's credibility.

1. **The authorities should improve data collection, data reporting and data infrastructure to ease operational work, reduce bottlenecks and mitigate operational risk.** Currently, the data are exchanged and manipulated on Excel spreadsheets, which lacks proper versioning and access rights. Using spreadsheets severely increases the risk of data manipulation and computations errors, and consumes an excessive amount of time. The IMF framework is entirely coded under a programming language – R - and could be readily interfaced with a proper database management system (for instance a free and open source system such as MongoDB, MySQL, MariaDB, etc.). They should also get better equipped in hardware. Appendix I provides technical recommendations on the best practices in IT infrastructure.
2. **The next steps should include follow-up TA missions by the IMF to integrate the liquidity forecasts with the monetary operations.** The BON now has a full-fledged liquidity forecasting framework, using the most recent models in the statistical literature, and is able to forecast the three autonomous factors – and not only CiC. The next step is to incorporate these forecasts to calibrate the monetary operations, which requires (i) to estimate the optimal structural liquidity surplus – based on an estimation of the demand for commercial banks reserves at the central bank, (ii) to calibrate the different monetary instruments based on their operational features and (iii) conduct operations accordingly. The IMF can assist the BON with these endevaors.

# Bibliography

Bank of Namibia. (2020): “Namibia’s Monetary Policy Framework,” *Bank of Namibia Operational Publications*, 1–19.

Gray, S. (2008) “Liquidity Forecasting Handbook,” *Centre for Central Banking Studies*, Bank of England

Hyndman, R. J., and Khandakar, Y. (2008). Automatic Time Series Forecasting: The Forecast Package for R. *Journal of statistical software*, 27, 1-22.

Hyndman R., R. Ahmed, G. Athanasopoulos and H. Shang (2011) “Optimal combination forecasts for hierarchical time series” *Computational Statistics and Data Analysis* 55(9), 2579-2589.

Hyndman R. and G. Athanasopoulos (2021), “Forecasting: Principles and Practice Second Edition”, *Otexts book*, 1-442.

Kourentzes, N., Barrow, D., and Petropoulos, F. (2019). Another Look at Forecast Selection and Combination: Evidence from Forecast Pooling. *International Journal of Production Economics*, 209, 226-235.

Schaechter, A. (2000). “Liquidity Forecasting,” *IMF Working Paper* 00/17

Tashman, L. J. (2000). Out-of-Sample Tests of Forecasting Accuracy: An Analysis and Review. *International Journal of Forecasting*, 16(4), 437-450.

Veyrune, M. R. M., & Guo, S. (2019). “Autonomous Factor Forecast Quality: The Case of the Eurosystem.” *IMF Working Paper* 19/296.

## Appendix I. Data and IT Infrastructure Requirements

Liquidity forecasting is a crucial component of the operational framework of any central bank. It should rely on accurate, updated, and comprehensive data, as well as an adequate software and hardware infrastructure. Implementing a complete and modern data and IT infrastructure will help the BON scaling up its modeling capacity in the future. The importance of a solid data and modern IT infrastructure is crucial not only for liquidity forecasting but also for pricing BON instruments, calibrating operations, managing collateral, monitoring market developments, implementing early-warning signals, etc. This Appendix reviews the data and IT infrastructure requirements for modern and efficient financial modeling.

**Data Coverage:** Liquidity forecasts are done on: (i) currency in circulation; (ii) state account at the central bank; and (iii) net foreign assets. Although these three items are available as central bank balance sheet items, their data must be updated daily and readily available to the monetary policy department for forecasting purposes. Granular data are essential for accurate and robust forecasting:

* *Currency in Circulation:* Daily data on currency issued by the central bank and currency deposited by commercial banks at the central bank. The decomposition by banknotes and coins denomination would help refine the forecasts by incorporating different denomination dynamics.
* *State Account:* The Treasury should share with the central bank its data on future expenditure and revenues items, in particular for large and exceptional items, to help the central bank anticipating significant liquidity disruptions (e.g., large flows). Also, more granular data on the revenues and expenditure of the government would allow the central bank to improve its forecast of the state account.
* *Net Foreign Assets:* The BON should receive information about large FX payments—particularly on diamond sales, which impact the central bank's net foreign assets.
* *General Financial and Economic Data:*BON staff, particularly research analysts and econometricians, should have access to a wide range of financial and economic data for their daily work, including liquidity forecasting and other topics. The most frequently cited databases are Bloomberg Market Data Feed™, Haver DLX™, EPFR Global Data™, Thomson Reuters Datastream™, and Refinitiv™. It is essential to access data from different providers—not only Bloomberg—to increase data coverage. For instance, Haver DLX™ and EPFR Global Data™ are helpful complements to Bloomberg on granular central banks data and daily capital flows information.

**Database Management:** The BON should maintain a central database repository that the monetary policy department can access at any time, with adequately managed access rights to preserve data integrity. The dataset should follow the best industry standards regarding identifiers, data storage and accuracy, software access, and management (e.g., an Oracle SQL database). Notably, the monetary policy department should have direct access to the data to use modern statistical software (e.g., Python and R, see below) to directly query from the database. Direct data access allows for easy automation, data retrieval, cleaning, and management, to gain time and minimize operational risk.

**Data Reporting:** Data from external providers should be reported automatically and carefully checked to avoid inconsistencies and other issues. More generally, the statistics department should operate an automatic data control process to detect large movements in the data (for instance, more than one standard deviation) and raise a warning so that the BON statisticians could check the input. Moreover, the monetary policy department should operate its own automatic routine check on the data it receives and communicate with the statistics department when needed.

**Harmonization:** Forecasting relies on time series. It is not enough that the dataset is recent and updated. The time series should also cover a sufficiently long period in the past to adequately fit the models. Given the evolving reporting standards and infrastructure, there is a need for statistical harmonization to ensure that the time series are consistent and comparable across time.

**Software Requirements:**Operating forecasting models and estimating financial models require a modern and complete statistical and programming toolkit. Given the strengths and weaknesses of the different available software, it is strongly recommended that the two most important of them, Python and R, are installed on BON computers. They can be completed by other software such as Matlab or Eviews. While R is a specialized statistical language commonly used by academics, Python has a broader scope and is widely used in the financial industry, including central banks. R has a wide range of packages for forecasting, statistical inference, and Bayesian estimation. Python is the leading language used for data management, data visualization, simulated methods, machine learning, big data, kernel estimation, etc. Packages such as reticulate or rpy2 allow for interoperability between the two languages.

Furthermore, it is crucial that the BON staff can use an integrated development environment (IDE) for both Python and R to work in the best conditions. Therefore, installing the most updated RStudio and Python Anaconda complete distributions is strongly recommended, as both offer a full suite of tools for designing and implementing codes for scientific computing.

Finally, developing and collaborating on programming projects necessitate tracking changes and sharing the codes internally to reduce operational risks and foster collaboration among staff. Therefore, it is strongly recommended to use a versioning software such as git for this purpose, with an internal Github repository for the BON. Archiving codes on Github tremendously reduces operational risks and helps programming teams to work together seamlessly. For instance, the IMF modelers use git internally for their programming projects.

Importantly, RStudio, Python Anaconda, and Git/Github are open-source software at the best international standards for data science and programming and are used by many financial institutions and large companies worldwide.

**Hardware Requirements:**Likewise, the BON should be equipped with adequate IT hardware to operate the forecasting and modeling infrastructure. To reduce operational risk, increase computing power, guarantees continuous data backup, and facilitate cooperation among staff, the BON should manage a high-performance computing server infrastructure. The computing servers should be accessible by the team from their BON laptop—even remotely if possible—and offer substantial computing power (both CPU and GPU) and massive fast-access memory. The BON should install Python Anaconda and RStudio on the servers and connect the servers to the BON data repository to allow staff to run operations from the servers. Running monetary operations from the servers (e.g., liquidity forecasting, modeling, etc.) has many advantages. It reduces operational risks, fastens operations, saves staff time, and offers more flexibility to use time-intensive and memory-consuming algorithms. Finally, granting staff access to cloud computing allows them to scale up large and complex projects efficiently.

## Appendix II. Statistical Forecasting Models and Methods

*This appendix presents the specification of the statistical models used for forecasting the autonomous factors. Details about data treatment, the interpretation of the results, etc. are available in the main text.*

### Regression-Based Models

**ARIMA Models**

**The Autoregressive Integrated Moving Average family of models is a flexible class of models used for time series forecasting in a wide range of settings.** The ARIMA model is defined as:

Here is the backshift operator that lags a variable, i.e. ,,etc. The order of differencing is typically equal to 1 (or in rare cases 2) for nonstationary series and 0 for stationary series. The term is known as the autoregressive (AR) polynomial or order and the term is known as the moving (MA) polynomial or order . The term is a random ‘noise’ or ‘innovation’ term. The nomenclature ARIMA(p,d,q) is used to describe an ARIMA model, for example an ARIMA model with , and would be referred to as an ARIMA(2,1,2) model.

**Model Selection (ARIMA Order)**

**For all ARIMA models, choices of the orders must be made.** The estimation is done using the stepwise algorithm of Hyndman and Kandakhar (2008):

1. Find using KPSS test.
2. Estimating four initial models and choose the best of these four models.
3. Expand the candidate model set by considering models that have or differing from current best by 1.
4. Iterate until no improvement made.

The criterion for selection is the Akaike Information Criterion corrected for small sample size (AICc). The Algorithm implemented in auto.arima function within the forecast package in the R software environment. The same algorithms can be modified and applied to seasonal ARIMA and seasonal ARIMA with regression described below.

**Seasonal ARIMA**

**An important extension to ARIMA models is seasonal ARIMAs (SARIMA), which allows for the modelling of patterns that repeat themselves every observations.** In general, SARIMA take the form:

where , and are the orders of the *seasonal* AR component, *seasonal* differencing and *seasonal* MA component. The nomenclature ARIMA(p,d,q)(P,D,Q)[m] is used to describe such models, for instance an ARIMA(1,0,0)(0,1,1)[5] model would be equivalent to

Seasonal ARIMA models of this form are only capable of explicitly capturing only one form of seasonality. In the Currency in Circulation data in particular (and possibly for the other autonomous factors), there are potentially multiple seasonalities. The candidate choices for are shown in the table below.

|  |  |
| --- | --- |
| m | Type |
| 7 | Weekly |
| 30 | Monthly |
| 354 | Yearly (Civic Calendar) |
| 365 | Yearly (Hijri Calendar) |

Fortunately, since the ARIMA model is able to easily incorporate covariates, there are alternatives to capturing these multiple seasonalities.

**Seasonal ARIMA with Regression**

**Dummy Variables**

**Seasonality can be modelled using dummy variables.** This approach is particular well suited when the length of the seasonal pattern is short and when the pattern is not necessarily smooth. For example flexible day of week effects can be modelled using only four variables of the form

With similar dummies can be defined for Tue, Wed, Thur, etc. These dummies are then included in a vector of covariates and the ARIMA model has the same specification as previously with replaced by . It is important to be clear that an ARIMA model with dummies (or other covariates) would be specified, in the example of an AR(1) as

and not

as is often the case. The first specification, which is the one that is used to allow for easier interpretation of the coefficients and avoids issues with using stationary predictors (such as dummies) as covariates for non-stationary data.

**Fourier Terms**

**Where is large, using dummies leads to an impractically large number of dummy variables.** Instead as long as the seasonal pattern in smooth, Fourier terms can be used instead. For these are defined as

and

The choice is considered for monthly, and yearly effects with the length of year in both the civic and religious calendars considered.

### Smoothing-Based Models

**ETS Models**

**The Error Trend Seasonality (ETS) class of forecasts is built up from simple exponential smoothing,** as per the following:

where is the forecast at time and is a smoothing parameter. This can also be written in terms of latent components via

which is useful for extending the model to account for trend and seasonality. Trend can be incorporated via

where is the forecast horizon, is the latent trend component and a smoothing parameter. Alternatively, a damped trend can be modelled using a damping parameter as

Seasonality can be model in an additive fashion via

where is a latent seasonal term and is a smoothing parameter. Seasonality can be accounted for in a multiplicative fashion via

To make these forecasting models statistical models they can be rewritten with error terms. For example, simple exponential smoothing with additive errors is given by:

Errors can also be incorporated multiplicatively. Different Error, Trend, Seasonal (ETS) models combine:

* Additive Error (A) or Multiplicative Error (M).
* No trend (N) Additive trend (A) or Additive damped trend (Ad).
* No seasonality (N) Additive seasonality (A) or Multiplicative seasonality (Ad).

The nomenclature ETS(.,.,.) is used, for example ETS(A,N,A) has additive errors, no trend and additive seasonality. Models can be selected via the AICc.

A disadvantage of the ETS family of models is that covariates cannot be as easily incorporated into ETS models—as is the case for ARIMA models—making it difficult to control for structural breaks and holiday effects.

### Advanced Models

**TBATS Model**

**The TBATS model incorporates many of the features of the models already introduced.** Seasonality and trend are handled via exponential smoothing (using trigonometric terms for the former), a Box-Cox transformation is used and ARIMA innovations. This allows seasonality to change over time. A particularly attractive feature of the TBATS model is the ability to handle multiple calendars—it was initially proposed for modelling electricity demand in Turkey which is a time series also influences by both the civic Gregorian calendar and religious Hijri calendar.

### Volatility Models

**Volatility models are appropriate for forecasting the net foreign assets, given the highly volatile nature of the time series.** Three classes of model are fitted.

**GARCH**

The most popular family of conditional volatility models is the GARCH model. The variance is modelled as

**EGARCH Model**

The specification of the eGARCH model is given by

where

An advantage of this specification is its asymmetry since the sign and magnitude of innovations have different effects on the variance.

**GJR GARCH Model**

The GJR-GARCH specification is given by

where if and if

Similar to eGARCH, this specification allows for asymmetric effects.

## Appendix III. Evaluating Model Performances

**Alternative forecasts performance are benchmarked via out-of-sample rolling origin forecasts.**The team produced 20 day ahead forecasts and retained the last 100 days as a test set. The rolling origin evaluation is implemented as follows. The training set specifies and trains all models and produces forecasts for the first twenty days of the test set. First, the out-of-sample forecast accuracy is recorded. Next, the training set is increased by one observation so that the models are re-specified and re-trained. The process is repeated until all the observations in the test set are exhausted.

**The mean Root Mean Squared Error (RMSE) and Median Average Errors (MAE) across forecast origins are the metric to evaluate forecast accuracy.** Quadratic errors track the mean of the distribution of the forecast target, preferring unbiased forecasts, but are sensitive to extreme values. For this reason, the RMSE and MAE across origins are preferred. If the two metrics deliver similar results, it implies that the forecasts are stable across forecast origins. On the contrary, when there is a large difference between the RMSE and the MAE, particularly with substantially higher mean RMSE over the median RMSE, this implies that the forecasts are unreliable across periods, and there are events of extreme errors.

The RMSE at a given horizon is defined as:

The MAE at a given horizon is defined as:

Where is and is the h-step ahead forecast of .

## Appendix IV. Model Selection and Combination

**The rolling origin setup supports dynamic model selection across forecast periods.** The user can identify forecasts that perform better locally by tracking the accuracy of past forecast origins. The performance tracking uses the mean RMSE of the t+1 to t+5 forecasts over the last five forecast origins. To increase the reliability of model selection, the IMF experts have used a window of 5-period origins. The random part of the actuals heavily influences errors from a single origin in the evaluation period. As the evaluation window expands to include more periods, some of this randomness is canceled out. The selection of 5 periods is a compromise of some mitigation of randomness while still retaining the recent observations. Increasing the window to more periods expands the focus of the selection to older observations. The advantages of using more than a single period are discussed further in the forecasting literature when the merits of rolling origin evaluation are explained in details (Ord et al., 2017; Tashman, 2000).

**Forecast combination is used as an alternative selection.** The forecasting literature has shown that forecast combination is generally beneficial over the use of a single forecast. Forecast combination does not require selecting a single model and therefore minimizes the risk of choosing and relying on a single potentially inappropriate forecast. The team uses equal weights in combining the forecasts. Equal weights limit the estimation errors in the combination weights that can affect negatively the performance of the combined forecast. Forecast combination builds both on the substantial weight of empirical evidence in the forecasting literature, and the more recent theoretical understanding on the effect of the estimation uncertainty of the combination weights to the final forecast. Kourentzes et al. (2019) discuss forecast combination extensively.

**Forecast pooling produces a single forecast from a set of projections, both for dynamic model selection and combination. The pooling selects only well-performing forecasts**. Forecast pooling is used to offset the variability of the final forecast both from the selection and the combination described above. The notion of pooling suggests that instead of using all forecasts considered, the pool is restricted to the top-performing models. Forecast pooling is somewhat at odds with the theoretical understanding that forecast combination performs best when there is substantial diversity in the base forecasts to be combined. However, the team reconciles these two approaches by limiting the pool of models to those that capture partial or complete seasonal information, excluding the Naïve (Random Walk) from the pool of models. These models capture the information in the time series in different ways, which introduces diversity. Kourentzes et al. (2019) provide an overview of the ideas of forecast pooling.

## Appendix V. Software Documentation

*The documentation describes the R code used during the TA mission on forecasting the autonomous factors, what it aims to do, how to run it, and the underlying models and outputs.*

**Language**

The code is written in the R language, and information is available at   
<https://www.r-project.org/>. R is an open-source language, meaning that all source code is available and is made of contributions by many people worldwide. It is also freely available, so the language and other libraries are also available to anybody with an internet connection free of charge.

The team chose the R language for this project because R has many statistical and forecasting packages that are easy to use “out-of-the-box.” Abundant pre-coded packages allow implementation easily under R, such as ARIMA, exponential smoothing, and other models.   
R can also write out into familiar formats like csv and Excel and create visualizations that can be exported to PDF format, making the output readable to a large audience of non-specialists.

**Environment**

For most users of the R language, RStudio (<https://www.rstudio.com/>) is the de facto environment for reading, writing and executing R code. RStudio is also free.

and is used by statisticians, economists, and other professionals all over the world. To install RStudio, please visit <https://www.rstudio.com/> and install the version corresponding to the operating system.

To run the code provided by the IMF team, it is recommended to use R version 3.6.3 so that the library version is the same. For a Windows environment, the IMF provides a folder containing the necessary libraries to run the code. The user should place this code in the same folder as the R source code, and this will ensure that the version of all libraries also matches.

**Setup**

The framework consists of several R scripts with file extension .R, several R Markdown files with extension .Rmd, a file config.R containing the configuration parameters, and a file utils.R which contains functions used throughout the R scripts.

Before running the code, the config file config.R should be edited to configure all necessary parameters that are passed into the R scripts. The next section provides an overview of the script corresponding to each stage, and explains what the code does as well as what the outputs are.

**R Scripts**

This section provides an explanation of each stage of the framework. For each stage, there is one script to be run, with the behavior of the script being controlled by the parameters in config.R, and possibly by the output of previous stages.

The four stages are

1. Preestimation/setup
2. Validation
3. Forecast
4. Reconciliation

In all cases, preestimation should be run to set up the models. The second stage, validation, is not strictly necessary, but to evaluate models based on past performance, it is recommended to run this stage. The results of this stage will be used by the next stages of forecast and reconciliatoin.

The third stage, forecast will generate forecasts based on the most recent data. If validation has been run, these results will be used to select the best model based on past performance.

The final stage, reconciliation, can only be done if data for all autonomous factors are available. The code will take the forecasts of each autonomous factor as well as the overall net liquidity and reconcile the forecasts.

**Preestimation/Setup**

**Data Preparation**

The first step is to prepare the data. A folder data should be created and **csv files** containing the data should be prepared. As convention, the csv files should be named CIC.csv, NFA.csv, SAB.csv, and:

* The first three lines should be used for comments and will be ignored when reading in
* The fourth line should be the column names, for example date, CIC. **The date column should always be named date.**
* All remaining rows should be the **date in format YYYY-MM-DD** and the value of the series. Dates with no data (due to weekends or no data availability) can be filled with N/A or not be included. The program will fill in and interpolate any missing dates.

Once these files are prepared, go to config.R and make sure that the series that have data are included in SERIESLIST (see below for explanations of all configuration parameters).

**Running preestimate.R**

The next step in preestimation is to run the script preestimate.R. Within RStudio, open the file preestimate.R and click on the Source button at the top right of the Source panel. The script has three parts:

* Generation of a file containing dates of structural breaks (edit configuration parameters SB\_PERM, SB\_TRAN\_START, and SB\_TRAN\_END in config.R, see also below for a description of these parameters).
* Generation of a file containing Fourier terms (edit FOURIERFREQ and FOURIERK).
* Generation of a file containing holidays (edit HOLIDAYLIST for fixed-date holidays and create a file holidays.csv for moving holidays).

When the script is finished, three .RDS files should be created in the folder output/preestimation, these will be used for models that use external regressors.

**Validation**

The validation step is not strictly necessary, but provides input on which models had the best performance by comparing forecasts to past data. For each validation date, a forecast is made for HORIZON days and compared to data, and metrics such as RMSE and MAE are calculated across different horizons and validation dates.

To set the dates to run validation on, the configuration parameters NVALIDATIONS, VALIDATION\_END VALIDATION\_FREQ and VALIDATION\_MIN should be set. VALIDATION\_END should be the last date to run validation over, while VALIDATION\_FREQ specifies the number of days between validation dates, so that in the end NVALIDATIONS validation dates are available. VALIDATION\_MIN ensures that the data has a minimum number of days included. Also, note that any validation date that is less than HORIZON days away from the last data point will be removed (see below for further details of these parameters). The parameter VAL\_METRIC which is either set to RMSE or MAE will determine how the models are ordered for past performance.

Once these parameters are set up, run the script validate.R which will run the forecasts for the generated validation dates and compare models. The models to run over are specified by the configuration parameters CIC\_MODELS, SAB\_MODELS, NFA\_MODELS, AGG\_MODELS (see below for detailed descriptions.)

The script validate.R will produce several outputs. For each validation period and model that is run, a model output file in the form output/model\_fits/[series]\_[model]\_[validation date].rds will be created, where [series], [model] and [validation date] will be appropriately substituted. These files can be read in later so that models do not need to be reestimated.

Another file output/forecasts/[series]forecats\_[validation date].csv will also be created, containing the forecasts for all models run for this series at this validation date.

The validation stage is meant to batch process past validations. The most time-consuming part of this framework is estimating the models, and for this reason, **if the model output file exists, the program will only read in those results and not run the estimates again.** Care is needed as if the script files or configuration parameters are changed, they may change the expected behavior of the models, but if the output files already exist they will not be updated. The output files can either be manually removed, or the script reset.R can be sourced to remove all files.

At the end of the validation script, it will generate a report based on R Markdown, a document format that allows generation of Word, HTML, PDF and other outputs. The source of the validation report is report\_validate.Rmd.

The output validation report will appear as reports/validatoin/validation\_[series]\_[validation date].html. It will contain:

* A table of RMSE values, with highlights for models that were selected using combinations of models, as well as showing the model chosen with parameters like CIC\_SELECTED in **bold**.
* A chart of RMSE values agains horizon averaged over all validation periods.
* Same chart and table for MAE.
* A table of Hit Ratio, where hit ratio is defined as the percentage of historical data that fell outside of the 95-percentile of the forecasted range.
* A chart of the series of interest, together with holidays shown as vertical bands.
* A chart showing the data together with the latest models overlayed.
* For series that use non-GARCH-type models, the latest models and their forecasts are shown as well as the actual data for the latest validation date. For series using GARCH-type models, the first difference in the data, which is modeled, is shown, along with conditional variances for each GARCH model.
* A description of each model that was estimated.

**Forecast**

For most cases this stage is the most important, and it provides forecasts for the last date of data. To run this stage source the file forecast.R which will estimate all models specified in CIC\_MODELS and similar parameters.

At the end of the script a report is generated, using the source file report\_forecast.Rmd, and the output will appear in reports/forecast/forecast\_[series]\_[forecast date].html. The forecast report will show:re

* A table of forecasted dates along with forecasted values and variances.
* A chart of the historical data up to the forecasted date along with the selected model for this series and the 80- and 95-percentile ranges.
* A chart showing the forecasts due to the average of all models
* the best of combination of two models based on historical performance followed by tables showing the forecast values.
* The latest forecasts for all models When the best combination for average of two models is chosen, the latest validation results are read in. **If the validation is more than 30 days before the current forecast date, a warning will appear in the report.**

**Reconciliation**

If all series CIC, SAB and NFA are available in SERIESLIST, the aggregate of these can be reconciled into a consistent result. When reconciling the forecasts from each component, the models specified by CIC\_SELECTED etc., will be used so it is important to set these properly.

There are four methods of reconciliation, Base (Unreconciled), Bottom up, OLS and MinT. The script reconcile.R will run these methods to reconcile the forecasts, and at the end will generate a reconciliation report which will appear in reports/reconcile/report\_reco\_[forecast date].html.

The reconciliation will be done on the latest common date to all series, so it is important that all series have the same date. If any of the series have data available beyond this date, a warning will appear in the report.

The reconciliation report will show:

* The latest reconciled results along with their components
* Charts showing historical data ending at the reconciled date together with forecasts and their ranges for all reconciliation methods
* Tables and charts of RMSE, MAE and a table for Hit Ratio, similar to the validation report.

Note that the validation period will use the same validation parameters VALIDATION\_END, NVALIDATIONS, VALIDATION\_FREQ, VALIDATION\_MIN as in the validation stage, so having run the validation stage will save a lot of time as the models will then not have to be reestimated. The forecasts from the validation dates are reconciled, then compared to the historical data to determine which reconciliation method performed the best. The criteria for this is given by REC\_METRIC and is RMSE or MAE.

The method selected with REC\_METHOD will be shown in **bold** in the tables, while the best performing method based on the validation results will be highlighted.

For more details on the reconciliation parameters, see below.

**Config Parameters**

**Global Parameters**

These parameters specify the global behavior of the program, for example the series to model, the horizon to forecast, as well as the name of the agency doing the forecasting, the currency name and units, etc.

* + - **SERIESLIST**

A named list containing the series (CIC, SAB, NFA, AGG) with names like Currency in Circulation as descriptions. Comment out or remove any series that do not have data, and for those included, make sure that the file data/[series].csv exists.

* + - **HORIZON**

The forecast horizon, or how many days should be forecasted. For example, if a forecast is needed for the next 14 days, this should be set to 14.

* + - **NCORES**

The number of cores to use in case of parallel computations. **Currently not used.**

* + - **AUTHORITY, CURRENCY, CURRENCY\_SCALE**

AUTHORITY is the name of the authority running the framework, such as “European Central Bank”. This will appear in the output of the HTML reports as the author.

CURRENCY is the currency of the data, such as “US dollars”, and CURRENCY\_SCALE is the scale, such as “Millions”. These also show up in the HTML reports.

* + - **DATASTARTDATE**

A parameter to specify whether to start the data at a given date. This is always used when reading in the data, so is a way to globally control when the data starts. If set to NULL, all available data will be used. This may be useful for example when there is a structural break in the data and data before a given date is not useful for modeling.

**Preestimation Parameters**

* + - **HOLIDAYLIST, HOLIDAYFILE**

HOLIDAYLIST is a named list to specify holidays with fixed dates. For example for New Year Day (January 1st), the date of January 1st is represented as “01-01” in the name of the list, and the actual entry is “New Year Day”. Any fixed date holidays can be added in this fashion, and the preestimate.R script will use these dates and generate the holiday regressors file. The Islamic holidays of Eid al Adha and Eid al Fitr are special holidays that move around but if they are included in this list, a program will calculate the dates and add them to the external regressors.

HOLIDAYFILE is for special holidays that are not fixed dates, and if such holidays exist, these can be added into a csv file so that holiday regressors can be generated for these holidays as well. The format of this csv file should be as follows:

* The first two lines will be ignored and treated as comments.
* The third line should contain the column names of date, Type.
* All rows after that should be a date in format YYYY-MM-DD, followed by the name of the holiday, such as “Holy Week.”
  + - **HOLIDAY\_RDS**

The output RDS file (a binary format used by R) for the holiday regressors. There should not be a need to modify this.

* + - **SB\_PERM, SB\_TRAN\_START, SB\_TRAN\_END**

These parameters are for structural breaks in the data. SB\_PERM is a list of dates where a level shift occurred, for example if there was a sudden expansion on a given date, and this expansion was maintained. In the modeling process, if external regressors are used all dates included will be used. If no such dates exist, set to NULL.

SB\_TRAN\_START and SB\_TRAN\_END are pairs of start and end dates for transitory sturctural breaks, for example if there is a fixed period where elevated levels existed. If there are no such structures set this to NULL.

* + - **SB\_RDS**

The output RDS file for structural breaks, there should not be any need to modify this.

* + - **FOURIERFREQ, FOURIERK**

FOURIERFREQ represents the long-term periodicity in the data (daily and weekly periods are modeled separately). For example, if there are known monthly, quarterly, or annual periods, these can be modeled as a combination of trigonometric functions. FOURIREFREQ specifies the length of these frequencies, and FOURIERK specifies how many terms to keep.

* + - **DAILYDUMMIES\_RDS**

The output RDS file for periodic regressors, there should not be any need to modify this.

* + - **SEASONALITIES**

The actual terms of FOURIERFREQ to be used when modeling external regressors. Therefore, this should be a subset of FOURIERFREQ.

**Validation Parameters**

These parameters are used to specify the behavior of the validate.R script.

* + - **NVALIDATIONS, VALIDATION\_END, VALIDATION\_FREQ, VALIDATION\_MIN**

These parameters determine which dates to run validation on. *Validation* is to run forecasts over past dates where data for the forecasted periods already exist, and to compare the results. The parameters here are used together in a single function calc\_forecast\_dates() to calculate dates to run validation over. The parameter VALIDATION\_END specifies the final date of validation to use, and from there NVALIDATIONS previous validation dates are calculated, each of which are VALIDATION\_FREQ days apart. For example if VALIDATION\_END is given as as.Date("2021-01-01"), NVALIDATIONS is 10, and VALIDATION\_FREQ is 7, the validation dates will be the 10 dates starting on 2020-10-30 and ending on 2021-01-01, each day being 7 days apart.

This setup is useful as forecasts may be made on specific dates of the week for a given horizon, so that the validations are based on actual past forecast dates. In the case that daily forecasts need to be validated, VALIDATION\_FREQ can be set to 1.

The parameter VALIDATION\_MIN ensures that there is a minimum number of days in the data to do modeling. If for any of the validation dates calculated using the above parameters, if the number of days to model is less than this parameter, this validation date is removed from the results of the calc\_forecast\_dates() function.

* + - **VAL\_METRIC**

When validating past performance, the models must be compared using a metric. The VAL\_METRIC specifies whether to use RMSE or MAE to compare model performance. This metric is then used for the Dynamic Model Selection of best performing model over the validation dates.

**CIC Parameters**

These parameters specify the behavior when running over CIC data. They are used in the validation, forecast, and reconciliation stages.

* + - **CIC\_MODELS**

A named list of models to run to model CIC. The abbreviated name like regarima is given in the list, with each model having a name like ARIMA with Regression, mostly for reporting purposes.

When running the forecasts, these models will be run.

* + - **CIC\_SB**

Structural breaks (if any) to specify for CIC. Should be given in the form ‘sb\_perm\_1’ where the number is the order within the SB\_PERM parameter.

* + - **CIC\_SELECTED**

The model selected by the user to highlight, used in the reports for forecasts and for reconciliation.

**SAB Parameters**

These parameters are similar to the CIC parameters but applied for modeling SAB.

* + - **SAB\_MODELS**

A named list of models to run to model SAB, same structure as CIC\_MODELS.

* + - **SAB\_SB**

Structural breaks (if any) to specify for SAB, same structure as CIC\_SB.

* + - **SAB\_SELECTED**

The model selected by the user to highlight, same structure as CIC\_SELECTED.

**NFA Parameters**

These parameters are similar to the CIC parameters but applied for modeling NFA. One difference is that typically NFA will not be modeled well with models like ARIMA, ETS or TBATS, so GARCH-type models will be used.

* + - **NFA\_MODELS**

A named list of models to run to model SAB, same structure as CIC\_MODELS.

* + - **NFA\_SB**

Structural breaks (if any) to specify for SAB, same structure as CIC\_SB.

* + - **NFA\_SELECTED**

The model selected by the user to highlight, same structure as CIC\_SELECTED.

**AGG Parameters**

These parameters are for forecasting the aggregate of CIC, SAB and NFA.

* + - **AGG\_MODELS**

A named list of models to run to model SAB, same structure as CIC\_MODELS. Similar to NFA, AGG models will typically GARCH-type. ### AGG\_SELECTED The model selected by the user to highlight, same structure as CIC\_SELECTED.

**Reconciliation Parameters**

* + - **REC\_METHOD**

The chosen method for reconciliation by the user to be highlighted.

* + - **REC\_METRIC**

Similar to VAL\_METRIC, the metric for selecting the best reconciliation method, should be either RMSE or MAE. Past reconciliation results will be compared to actual data based on this metric and ranked. The best performing method will be highlighted in the reconciliation report.

**Conclusions**

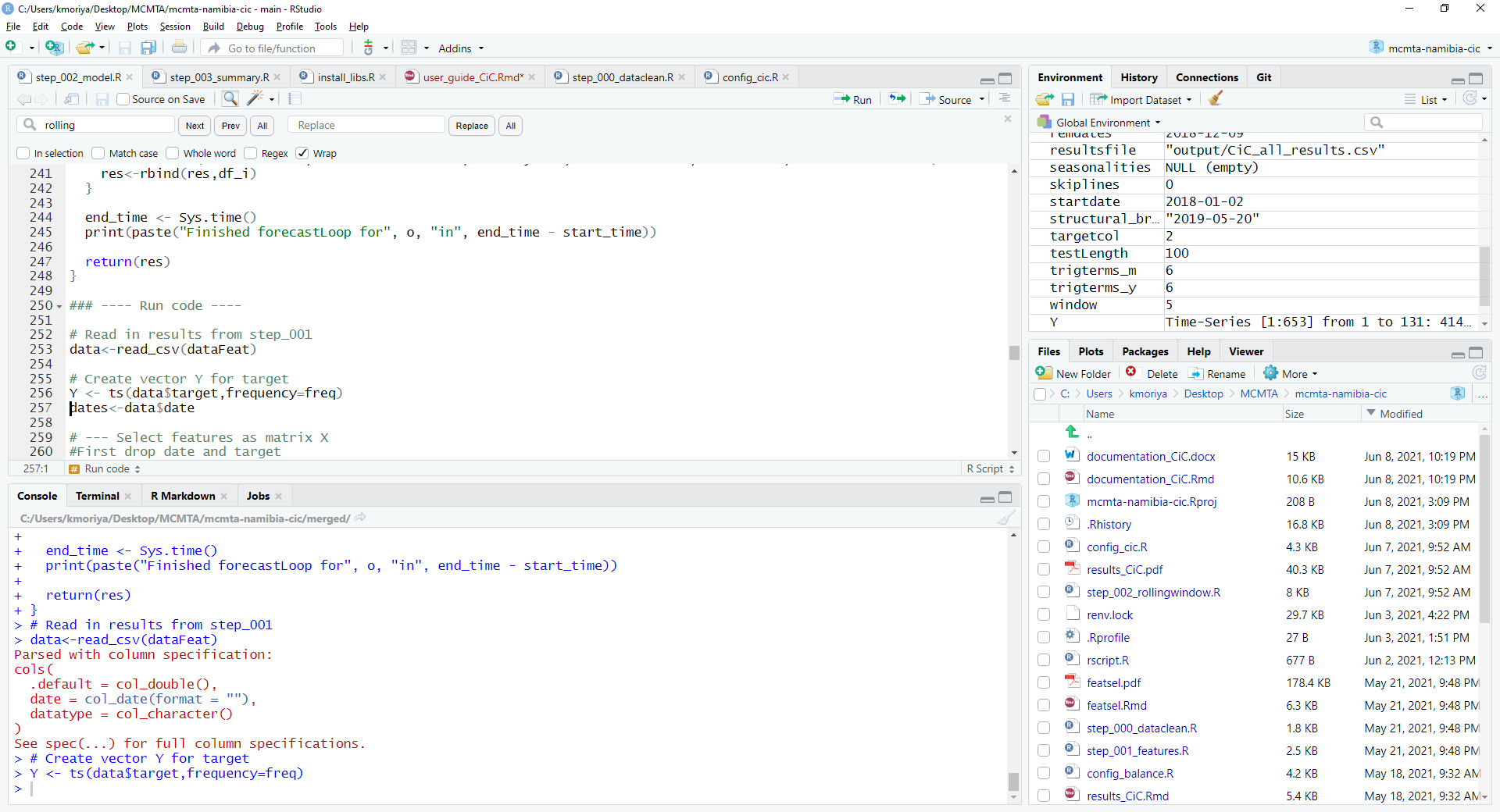
This document describes how to configure and run the liquidity forecasting framework developed by the IMF’s Monetary and Capital Markets Department. For any questions, the contact information is available on the first page.

This document is also generated by R Markdown, similar to the reports that are generated at different stages. The source code is the file UserGuide.Rmd and after opening in RStudio, clicking on the Knit button will generate the output Word document.

## Appendix VI. User R Guide

**Using RStudio**

The code used for this project is in the R programming language, and it is easiest to use RStudio for reading, writing, and executing R code. Once RStudio is installed, the application should look like the following:



Source: IMF staff

Note: Screenshot of RStudio.

The upper left shows the source code, typically in the form of R scripts with extension .R. The lower left side shows the console, where R commands can be executed. The upper right panel shows available variables and additional useful information, while the bottom right panel shows files in the current folder, installed packages and other information.

More information on RStudio and how to use it is available at <https://www.rstudio.com/> and other online resources, many of which are free.

**Running the Code**

To run the code, first open the R script of interest. The R code will have a line like

source('config.R')

so if necessary edit the config file. Details of the config file options are in the documentation file.

Once this is done, the user can execute each code line by moving the cursor to the line of interest and typing Ctrl + Enter. The line(s) run will be echoed in the console panel in the bottom left, and the console will show any output or errors as well.

The user can execute the entire R script either with Ctrl + A (select all) followed by   
Ctrl + Enter, or using the Source button at the top right of the source code panel.

As the code is executed, various outputs such as csv, Excel and PDF files will be generated, which will assist in understanding what the code is doing and how each model works.

For help with the code, please contact Kei Moriya ([kmoriya@imf.org](mailto:kmoriya@imf.org)) and Romain Lafarguette ([rlafarguette@imf.org](file:///C:\Users\kmoriya\OTmp\rlafarguette@imf.org)).

1. TBATS: Trigonometric Seasonal, Box-Cox Transformation, ARMA Residuals, Trend and Seasonality [↑](#footnote-ref-2)
2. We follow the literature and we define the autonomous factors as the factors that (i) are not under the control of the central bank and (ii) impact banking liquidity. The main autonomous factors include currency in circulation, net foreign assets and the state account balance. “*The autonomous factors are the factor not under central bank direct control*.” (Bank of England, 2008), “*Autonomous factors refer to factors beyond the control of the central bank*” (Norges Bank, 2019), “*Autonomous supply of liquidity (i.e., all supply factors that are in the shortrun beyond the control of the central bank)*” (IMF 2000). Other items on the central bank balance sheet that impact liquidity are either used for monetary policy purposes (e.g. OMO), or can be perfectly anticipated by the central bank (other domestic assets not held for monetary purposes) and therefore don’t need to be forecasted. Other items might be outside of the central bank control but don’t impact liquidity (for instance, the reevaluation account) [↑](#footnote-ref-3)
3. The IMF will follow-up with Namibia to incorporate NFA series. [↑](#footnote-ref-4)
4. See Veyrune and Guo (2019) for a quality test of the ECB autonomous factors forecasts. [↑](#footnote-ref-5)