Published by :
https://www.ijert.org/
An International Peer-Reviewed Journal

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 14 Issue 12 , December - 2025

A Legal Web Crawling And Reverse Engineering Framework For Economic Data, Analysis, And Modeling: Application To Exchange Rate Depreciation And Stabilization In Madagascar

Andrianaivo Herifidy Malalaniaina
National School of Computer Science,
Doctoral School of Modeling and Computer Science
University of Fianarantsoa
Fianarantsoa, Madagascar

Dimbisoa William Germain
National School of Computer Science,
Doctoral School of Modeling and Computer Science
University of Fianarantsoa
Fianarantsoa, Madagascar

Thomas Mahatody
National School of Computer Science,
Doctoral School of Modeling and Computer Science
University of Fianarantsoa
Fianarantsoa, Madagascar

Abstract - **The collection of data constitutes a fundamental pillar of study, scientific research, and decision-making support mechanisms. In reality, this data collection remains a slow and complex process in the context of scientific study and research. These difficulties stem from the dispersion of sources, the heterogeneity of formats, and legal constraints related to data access. Furthermore, the increasing availability of economic information on the web represents a significant opportunity for analysis and knowledge production. This paper proposes a legal web crawling framework for collecting, analyzing and modeling economic data. The approach is based on reverse engineering, which aims to transform web interfaces into independent and usable data models. The framework combines the use of existing public APIs, such as those from the World Bank and other international institutions, with controlled and compliant crawling mechanisms. The collected data is then subjected to analysis and modelling steps to support decision-making. The results highlight the relevance of this approach for improving the accessibility, legality and reproducibility of economic analyses based on web data. This study adopts a multi-method approach that integrates legal web crawling for data collection, principal component analysis (PCA) supported by correlation analysis to identify the key determinants of currency stabilization, multiple linear regression to assess the marginal effect of each variable on the exchange rate, and an AutoRegressive Integrated Moving Average with eXogenous variables (ARIMAX) model to generate forward-looking exchange-rate projections. This study aims to identify the key determinants of currency stabilization and depreciation in Madagascar, and to propose strategic policy approaches that may support and strengthen monetary stability.**

Keywords - reverse engineering; Legal web crawling; PCA; multiple linear regression; ARIMAX; Exchange rate modeling.

I. INTRODUCTION

Data collection plays a crucial role in scientific research and decision-making processes, as it significantly influences the quality of analyses and the relevance of the results obtained [1]. In the context of data analysis, several methodological approaches can be considered, each enabling the extraction of meaningful insights. Among these data collection methods, field surveys represent a particularly relevant approach, as they allow researchers to gather data directly from the target populations and obtain rich qualitative and quantitative information. However, this process is often time-consuming and costly due to the human resources and time required to reach respondents and ensure the validity of the collected information [2].

In parallel with these constraints, the increasing availability of economic information on the Web represents a major opportunity for analysis and knowledge production. Numerous national and international institutions now publish economic, social, and financial indicators through digital platforms and open web interfaces [3], such as the World Bank [4] and public ministries. These data provide considerable potential for empirical studies and decision-making. However, they are often presented in a heterogeneous and poorly structured manner, having been originally designed for human use, which limits their direct exploitation for analytical purposes. In this context, conventional methods of manual data collection or unregulated scraping prove inefficient in terms of speed, cost, and legal compliance.

Several studies have highlighted the legal and ethical risks associated with automated web data extraction when access

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 14 Issue 12 , December - 2025

rules, terms of use, and data protection principles are not explicitly considered [5][6] .This underscores the need to develop systematic, reproducible, and regulation-compliant data collection methods.

This paper proposes a legal web crawling framework for the collection, analysis, and modeling of economic data. The approach follows a reverse-engineering logic, aiming to transform web interfaces into independent and exploitable data models for economic analysis[7]. The methodological framework combines existing public APIs, such as those provided by the World Bank and other international institutions, with controlled crawling mechanisms that comply with access regulations.

The collected data are subsequently subjected to analysis and modeling steps to support decision-making. The results of this study highlight the relevance of the proposed approach in improving access to economic data, enhancing the legal compliance of data collection processes, and increasing the reproducibility of analyses based on web-sourced data [1].

This paper proposes a new reverse-engineering approach to transform web-based interfaces containing economic data into independent analytical models, from which data can be extracted and economic policy decisions can be inferred. The proposed approach integrates several complementary methods, including legal web crawling, PCA, multiple linear regression, and the ARIMAX model.

II.RELATED WORK

Web data collection and automated scraping have been widely studied in recent years. Y. Bassil (2012) [17] provided a survey of web crawling techniques and their applications, highlighting the limitations of traditional approaches in terms of speed, cost, and legal compliance. E. Ferrara et al. (2014)[1] examined various web data extraction methods, emphasizing the complexity of heterogeneous structures and the need for automated yet compliant solutions.
Legal and ethical risks associated with unregulated web scraping have also been analyzed [5], who stress the importance of respecting terms of use and data protection rules. These studies underline the necessity of developing systematic and reproducible methods for web data collection while adhering to legal frameworks. Regarding the economic modeling approaches applied in Madagascar, J. R. Andrianady et al.(2018)[8], presents the use of a small VAR model to examine the relationship between inflation and money supply in the country. However, this study is limited to only the two variables under consideration. In addition, [9] apply the ARIMAX methodology to analyse and forecast the evolution of key macroeconomic variables in Madagascar. Their work highlights how both domestic and external factors shape the country's economic dynamics and provides useful projections to support economic policy decision-making.

More recent work as [10] and [11] has proposed frameworks that integrate legality, ethics, and efficiency in web data scraping. However, few studies specifically address legal collection of economic data, their analysis using statistical methods such as Principal Component Analysis (PCA), and econometric modeling to support decision-making. This paper presents a reverse-engineering approach that combines lawful data collection, correlation analysis, multiple linear regression, and ARIMAX modeling in order to analyze and model currency stabilization and depreciation in Madagascar.

III. METHODOLOGY

A. Reverse engineering

Reverse engineering [12][13] is an analytical approach that involves examining an existing system in order to understand its internal functioning, structure, and underlying mechanisms, without necessarily having access to its original design. It is based on the observation of the system's inputs, outputs, and behavior to reconstruct an explanatory or predictive model.

In the context of economic data and web-based interfaces [14], reverse engineering aims to extract usable information from digital platforms, organize raw data, and identify implicit relationships among variables. This approach makes it possible to infer decision-making processes or the effects of economic policies from observed data.

Reverse engineering therefore goes beyond merely replicating the studied system, and instead seeks to produce an independent, interpretable model that can be used for analysis, forecasting, or decision support.

B. Data Collection

This study is based on the use of open economic data, which supports a methodological approach focused on publicly available, accessible, and verifiable information. The use of open data enhances transparency and ensures the reproducibility of scientific results.

Data collection relies on a legal web crawling approach [6] [15], combined with institutional APIs. This strategy addresses the limited coverage of APIs while strictly complying with regulatory frameworks governing data access. The crawling process is considered legal as it exclusively targets public content, respects robots.txt directives and platform terms of use, and avoids any extraction of sensitive or restricted data.

The main data sources include government websites, World Bank APIs, and digital platforms of public ministries. Institutional APIs provide direct access to structured and standardized datasets, whereas legal crawling enables the collection of complementary information available through web interfaces when such data are not exposed via programmatic services. The combination of these mechanisms ensures a comprehensive, compliant, and methodologically robust data collection process.

After collection, the data undergo a preprocessing phase involving cleaning, format harmonization, and transformation into homogeneous datasets. Variables are then organized into quantitative categorie to facilitate subsequent statistical analysis and econometric modeling.

The dataset covers several economic dimensions relevant to monetary stabilization in Madagascar. It includes monetary and financial variables, such as the Ariary/USD exchange rate, inflation measured by the consumer price index, money supply (M2), the central bank policy rate, domestic credit to the private sector, and foreign exchange reserves. Agricultural variables are incorporated to capture the sector's structural role, including total agricultural production, agricultural value added as a share of Gross Domestic Product (GDP), and agricultural

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 14 Issue 12 , December - 2025

productivity indicators. In addition, real macroeconomic variables, such as real GDP growth, household consumption, gross fixed capital formation, and unemployment or underemployment rates are used to reflect overall economic dynamics. Finally, external and fiscal variables, including the trade balance, fiscal deficit, external public debt, external aid and transfers, as well as a political governance index, are considered to account for institutional and external influences on monetary stability.

C. Crawling vs scraping

Crawling and scraping [16] are two methods used to extract data from web interfaces. Scraping focuses on the targeted extraction of information from specific web pages. Table 1 provides a detailed comparison between scraping and crawling methods, highlighting their differences in terms of scope, reproducibility, legal compliance, and suitability for scientific research.

Table 1. Comparative Analysis of Web Crawling and Web scraping

Criterion	Web Crawling	Web Scraping	References
Main Objective	Systematic discovery and exploration of web sources	Targeted extraction of specific data	[17]
Scope	Broad and structured (sites, portals, domains)	Limited to pre-identified pages	[1]
Methodological Approach	Reproducible process based on explicit rules	dependent on page structure	[1]
Scientific Traceability	High (documented sources, paths, and rules)	Low to medium	[17]
Legal Compliance	High when respecting robots.txt and terms of use	Variable, often problematic if unregulated	
Ethical Considerations	Integrated into the process	Depends on implementation	[5]
Reproducibility	High	Low to medium	[1]
Interaction with APIs	Complementary	Often redundant	[3]
Adaptation to Open Data	Highly suitable	Poorly structured	[3]
Legal Risks	Low if controlled	High if unregulated	
Recommended Use in Research	Primary method	Secondary, controlled step	[17]
Role in LADEA	Core of the methodological framework	Integrated occasional extraction	[1]

Web data collection is a fundamental step in modern empirical research, particularly in economic studies, as online information can provide real-time indicators and large-scale datasets. Two primary methods are commonly used for extracting data from web interfaces: crawling and scraping.

Web Scraping refers to the targeted extraction of data from specific web pages. It is typically applied when the research objective is narrow, for instance, to collect a defined set of variables from a few sources. While scraping can be efficient for small-scale extraction, it often depends on ad hoc scripts and the structure of individual pages, which limits reproducibility and may raise legal and ethical concerns if access rules are not respected [5].

Web Crawling, in contrast, is a systematic approach that allows the discovery and collection of data across multiple web sources or entire domains. Crawlers follow predefined rules and respect legal constraints, such as robots.txt and terms of service, which ensures compliance and improves reproducibility [1]. Moreover, crawling can be combined with public APIs to access structured datasets, enhancing coverage and reliability of the collected data [3].

The main differences between the two approaches are summarized in Table 1. Crawling is particularly suitable for large-scale studies, longitudinal analyses, and research requiring methodologically robust frameworks, while scraping is more appropriate for quick, focused data extraction tasks.

D. PCA

PCA is a versatile statistical method that reduces a data table to its essential characteristics, called principal components [18].

This multivariate approach is applied in the joint analysis of several indicators, thereby enabling the identification of structural relationships between variables. In this study, Principal Component Analysis is used to identify the variables influencing monetary destabilisation in Madagascar by highlighting factor axes that summarise the correlations between variables, thereby facilitating a comprehensive understanding of the phenomenon under study [19].

In this document, we will establish the correlation matrix between variables in order to visualise and analyse the relationships between different variables. To calculate the correlation matrix, we use the following formula:

$$r_{xy} = \frac{Cov(x,y)}{\sigma x * \sigma y} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 * \sum_{i=1}^{n}(y_i - \bar{y})^2}} \quad \textbf{(1)}$$

such as ,x_i and y_i denote the observations of variables x and y,

- $\bar{x}$ and $\bar{y}$ represent the respective average of the variables,
- σx and σy correspond to the standard deviations of x and y

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 14 Issue 12 , December - 2025

- n denotes the number of observations.

E. Multiple linear regression

Multiple linear regression [20] [21] is an extension of simple linear regression that makes it possible to examine the relationship between a continuous dependent variable and several independent variables (predictors).

The model expresses the dependent variable Y as a linear combination of the explanatory variables Xj, supplemented by an error term. For n observations and p predicators, the equation is written as follows:

$$Yi = \beta 0 + \beta 1 Xi1 + \beta 2\ Xi2 + \ldots + \beta p\ Xip + \varepsilon i \quad \textbf{(2)}$$

Where,

- Yi is the value of the dependent variable for individual i,
- β0 is the intercept,
- βj are the partial regression coefficients,
- εi represents the random error.

The multiple linear regression model represents a fundamental methodological tool for examining the multivariate determinants of the depreciation of the Malagasy Ariary. Unlike bivariate analyses, this approach makes it possible to isolate the net effect of each explanatory factor while controlling for the simultaneous influence of other variables, thereby addressing issues of endogeneity and multicollinearity that are common in macroeconomic time series. Its application provides a rigorous quantification of the structural elasticities between the exchange rate and its underlying economic fundamentals.

F. ARIMAX

The ARIMAX method [9] [22] is an extension of the traditional ARIMA model. It is applied to forecast time series by considering not only the past behavior of the series itself, but also the influence of external variables, known as exogenous variables. The model is generally represented by the following equation:

$$Yt=\mu+ \sum_{i=1}^{r} \beta i\, Xi,t + \Phi_1 Y_{t-1 + \ldots +} \Phi_p Y_{t-p} - \theta_1 \varepsilon_{t-1} - \ldots - \theta_q \varepsilon_{t-q} + \varepsilon t \quad \textbf{(3)}$$

Where Yt represents the value of the series at time t, Xi,t denotes the external (exogenous) variables associated with their coefficients βi, and εt is the error term.

ARIMAX represents a crucial methodological advancement compared to traditional static models. Its distinctive feature lies in the integration of endogenous temporal dynamics with structured exogenous influences. In the context of monetary modeling in Madagascar, this approach allows for the simultaneous capture of the historical inertia of the exchange rate (reflecting market memory) and the contemporaneous impact of economic policy shocks, as formalized by the following equation:

ARIMAX Forecast = Historical Memory + Impact of Current Policies + Anticipated Shocks **(4)**

G. Proposed Approach

This research develops a hybrid sequential methodological framework based on reverse engineering principles. The approach systematically integrates four analytical components: (i) ethical web data collection, (ii) influential variable identification through Principal Component Analysis, (iii) quantification of variable impacts using multiple linear regression, and (iv) future exchange rate forecasting with ARIMAX models. This structured methodology enables comprehensive analysis of monetary stabilization determinants in Madagascar. Our approach is summarised in Figure 1.

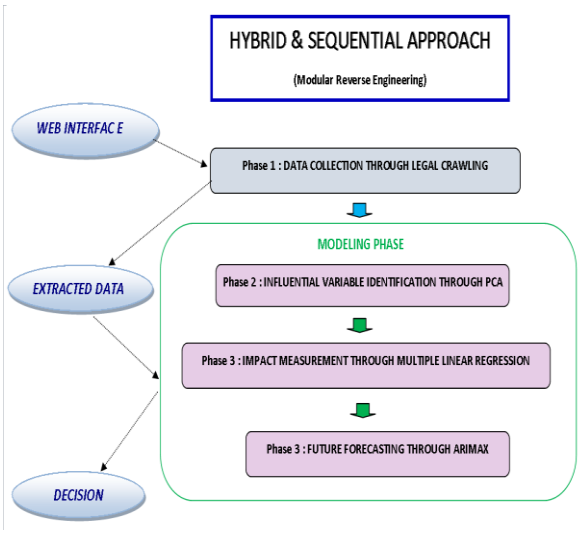

***Figure 1**. Proposed approach*

Figure 1 presents a hybrid and sequential methodological framework grounded in modular reverse engineering. The process starts from a web interface that serves as the main source of economic data. In Phase 1, information is collected through legal web crawling, ensuring ethical and regulatory compliance, and converted from unstructured web content into structured datasets.

The extracted data are then processed within a modeling phase composed of successive analytical steps. Phase 2 applies PCA to identify the most influential variables and reduce data dimensionality. Phase 3 evaluates the effects of these variables using multiple linear regression, enabling the analysis of their relationships with the target variable.

In Phase 4, future values are forecast using the ARIMAX model, which incorporates both historical patterns and exogenous factors identified in earlier stages. This sequential framework supports a gradual transition from raw web data to decision-oriented outputs, resulting in an interpretable and independent model suitable for economic analysis, policy assessment, and forecasting.

H. Computational Tools and Development Environment

This study was implemented using the Python [16] [23] [24] [25] programming language, which was selected for its robustness and its widespread use in economic data analysis and statistical modeling. Python provides a flexible environment that makes it possible to integrate, within a single workflow, the stages of legal data collection, preprocessing, statistical analysis, and econometric modeling.

Specialized libraries such as Pandas, NumPy, Scikit-learn, and Statsmodels were used respectively for time-series structuring, principal component analysis, regression estimation, and the implementation of ARIMAX models. The use of Python

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 14 Issue 12 , December - 2025

ensures methodological transparency, computational reproducibility, and facilitates the replication or extension of the analysis in future research.

IV. RESULTS

A. *Data Extraction and Preprocessing Outcomes*

This subsection presents the results of the legal web crawling process. These include the volume of data collected, the sources of the data, the cleaning procedures applied, and the final structure of the dataset used for analysis.

In this study, a chronological selection of macroeconomic data for Madagascar was carried out, covering the period since 1975. This dataset includes several relevant indicators, such as the exchange rate, inflation, broad money supply (M2) relative to GDP, the policy rate, private credit, reserves, GDP components (agriculture, households, investment), unemployment, the trade balance, the budget deficit, external debt and development assistance, as well as a governance index. This framework makes it possible to observe the joint evolution of these variables over time. It also helps to identify periods of economic tension or structural reform. Finally, it allows for cross-variable analyses, such as examining the relationship between inflation and the exchange rate, or assessing the correlation between governance and GDP growth. The correlation analysis between these variables will be presented in Figure 2.

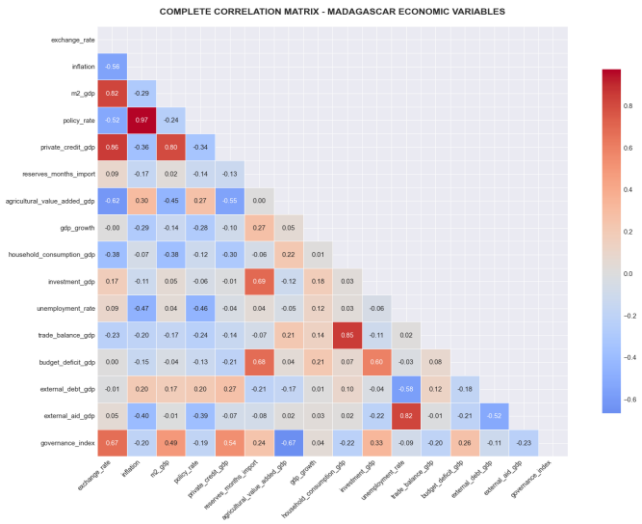

Figure 2. *Correlation matrix of extracted data.*

Correlation analysis*[26], [27]* represents a key step for assessing the existence and strength of linear relationships among the selected macroeconomic variables.

Strong positive correlations associated with the depreciation of the Malagasy Ariary:

- Private credit to GDP ratio (private_credit_gdp) : +0.856, this very strong correlation indicates that rapid growth in credit to the private sector is closely associated with exchange rate depreciation. This may reflect an increase in domestic demand and imports, which places downward pressure on the national currency.
- Broad money M2 / GDP : +0.819, The expansion of monetary liquidity in the economy is strongly linked to currency depreciation. This result is consistent with monetary theory, which suggests that excessive money supply growth may weaken the external value of the currency.
- Governance index +0.669: Although more moderate, this positive correlation may suggest that certain institutional improvements occur alongside financial liberalisation or short-term capital flows, which can increase volatility and exert pressure on the exchange rate.

Negative correlations associated with stability or appreciation:

- Household consumption / GDP :−0.381,

  This moderate negative relationship suggests that consumption-driven domestic demand is associated with greater exchange rate stability or appreciation, likely through more balanced economic growth.

- Policy interest rate :−0.521, An increase in the policy rate is correlated with currency appreciation or stabilization, which is consistent with restrictive monetary policy mechanisms that attract capital inflows and reduce inflationary pressures.
- Inflation :−0.555, The negative correlation indicates that well-controlled inflation is associated with greater exchange rate stability, while higher inflation tends to weaken the currency.
- Agricultural value added / GDP :−0.616, This relatively strong relationship suggests that strengthening the agricultural sector contributes to monetary stability, likely by improving the trade balance and reducing dependence on food imports.

B. *Identification of the most influential variables*

The identification of the most influential variables was carried out using a multiple linear regression model. After estimating the model, the main empirical results are reported in Table 2.

Table 2. *Result of the multiple linear regression*

Variable	Coefficient Effect	Interpretation
m2_gdp	+509.4475	DEPRECIATION
private_credit_gdp	+465.1361	DEPRECIATION
external_debt_gdp	−261.6734	STABILITY
investment_gdp	+183.5770	DEPRECIATION
agricultural_value_added_gdp	−168.0642	STABILITY
trade_balance_gdp	−122.1073	STABILITY
external_aid_gdp	+87.6869	DEPRECIATION

Table 2 presents the contribution of selected macroeconomic variables to either the depreciation or the stability of the

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 14 Issue 12 , December - 2025

Malagasy Ariary. Positive variation values indicate an association with currency depreciation, while negative values are associated with monetary stability.

The regression model shows a strong fit to the data. The R^2 value is 0.947, indicating that 94.7% of the variance in the dependent variable is explained by the independent variables. The adjusted R^2 is 0.923, confirming that the model remains robust when accounting for the number of predictors. The overall model is statistically significant, as shown by the F-statistic of 49.5 and a p-value < 0.001. These results suggest that the model reliably explains the variation in the dependent variable.

C. Sensitivity analysis with combined scenarios

- Inflationary Scenario

In this scenario, increased inflationary pressure combined with a rise in M2/GDP was simulated. The shocks (+50% and +30%) led to a final exchange rate of 5289 MGA/USD, representing a 22.9% increase compared to the current rate. The average impact over the period was 53.7%, showing that monetary expansion under inflationary conditions can significantly depreciate the currency.

Critical Debt Scenario

This scenario considered high external debt and budget deficits. Shocks of +100% and +50% resulted in a final exchange rate of 3940 MGA/USD, a decrease of 8.4% relative to the current rate. The average impact was -52.9%, indicating that elevated debt and deficit levels can exert a stabilizing or strengthening effect on the currency in the simulated environment.

- Agricultural Value Added Scenario

Here, improvements in the agricultural sector and trade balance were examined with shocks of +50% and +30%. The simulation produced a final exchange rate of 3572 MGA/USD, 17.0% lower than the current rate, with an average impact of -51.6%. This suggests that gains in agricultural productivity and trade performance contribute to currency stability or appreciation.

- Improved Governance Scenario

This scenario simulated institutional improvements through higher governance index scores and increased import reserves (+50% and +100% shocks). The final exchange rate reached 4660 MGA/USD, an 8.3% increase, with an average impact of 16.7%. The results highlight that better governance and stronger reserves moderately support currency appreciation.

This scenario will be illustrated in Figure 3.

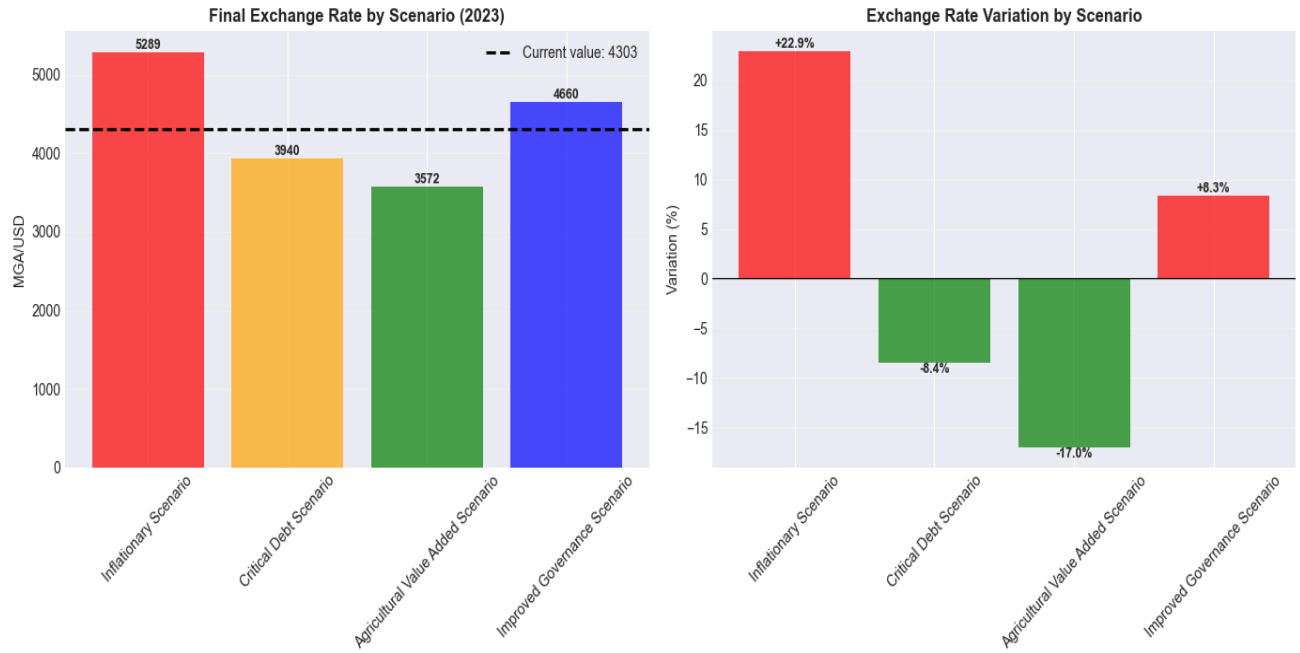

***Figure 3**. Overview of the combined scenarios*

The combined scenario simulations evaluate the effects of different macroeconomic shocks on the exchange rate. The Inflationary Scenario shows that higher inflation and monetary expansion lead to significant currency depreciation. The Critical Debt Scenario indicates that high external debt and budget deficits can have a stabilizing effect. The Agricultural Value Added Scenario demonstrates that improvements in agriculture and trade balance contribute to currency stability. Finally, the Improved Governance Scenario highlights that stronger institutions and higher reserves moderately support currency appreciation.

D. Exchange Rate Projections with ARIMAX

The exchange rate projections for the years 2024 to 2030 were obtained through the utilisation of an ARIMAX model. The baseline scenario, which corresponds to projections under unchanged macroeconomic and policy conditions, is presented in Table 3.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 14 Issue 12 , December - 2025

Table 3. *Baseline Exchange Rate Projection (2024–2030).*

Year	Exchange Rate (MGA/USD)
2024	4498
2025	4667
2026	4838
2027	5011
2028	5186
2029	5364
2030	5545

The projection scenarios were derived from the baseline ARIMAX forecast by applying scenario-specific adjustment factors to the exchange-rate trajectory.

In the pessimistic scenario, increasing positive adjustment rates were applied to the baseline path in order to reflect persistent macroeconomic imbalances, inflationary pressures and depreciation risk. This phenomenon is projected to culminate in a cumulative upward deviation of the exchange rate over the (2024–2030) periods.

Conversely, the optimistic scenario employs moderate and progressive negative adjustment factors, representing structural reforms, policy credibility, and inflation stabilisation, leading to a gradual appreciation of the Ariary relative to the baseline trajectory. In both cases illustrated in Figure 4, adjustments were applied cumulatively over the projection horizon, with controlled stochastic variability to ensure realistic and stable dynamics.

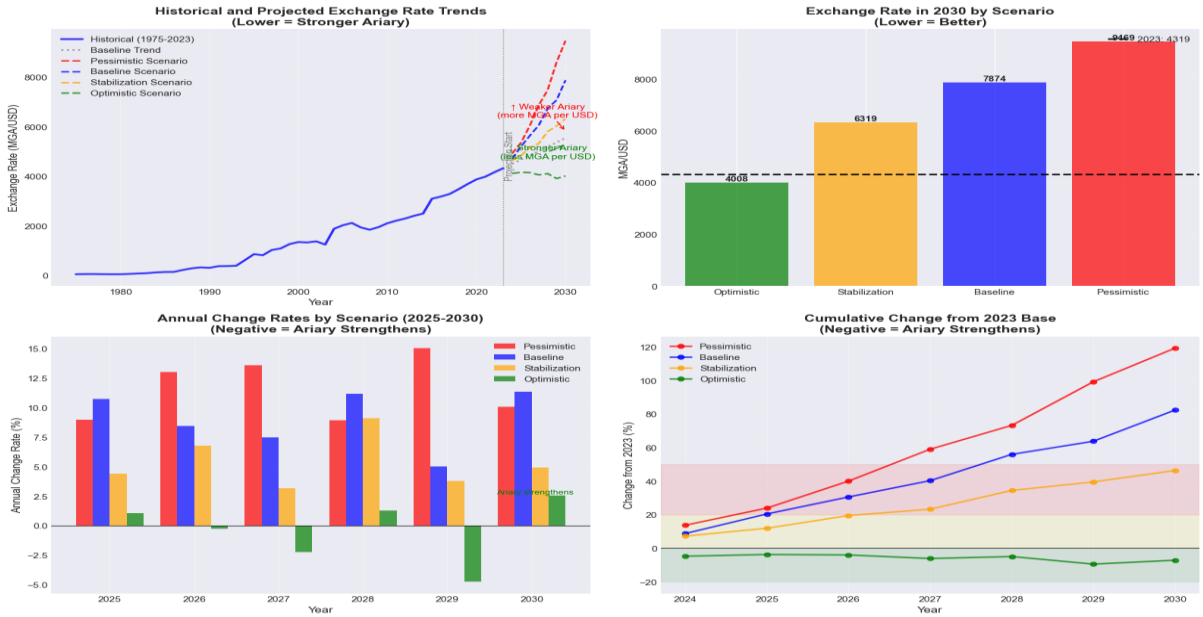

Figure 4. *Historical and projected exchange rates(1975-2023) with pessimistic, stabilization, and optimistic scenarios (2024–2030).*

Based on Figure 4, in addition to the baseline scenario, four alternative scenarios were constructed and analyzed:

- Baseline Scenario

The baseline scenario assumes a continuation of current macroeconomic trends. Under this scenario, the MGA/USD exchange rate exhibits a sustained depreciation throughout the projection period, increasing from 4,696 in 2024 to 7,874 in 2030. This corresponds to a cumulative depreciation of 67.7% and an average annual depreciation rate of 7.7%, indicating persistent exchange rate pressures in the absence of corrective policies.

- Stabilization Scenario

The stabilization scenario reflects the implementation of moderate macroeconomic stabilization policies. The projected exchange rate rises from 4,629 MGA/USD in 2024 to 6,319 in 2030, implying a total depreciation of 36.5% over the period. The average annual depreciation rate is reduced to 4.5%, suggesting that stabilization measures partially mitigate exchange rate volatility but do not fully reverse depreciation dynamics.

- Optimistic Scenario

The optimistic scenario assumes strong economic reforms and effective inflation control. In this case, the exchange rate remains relatively stable, fluctuating between 3,908 and 4,156 MGA/USD over 2024–2030. The Ariary slightly appreciates over the period, with a cumulative change of −2.5% and an average annual appreciation rate of 0.4%, indicating a structurally improved macroeconomic environment.

- Pessimistic Scenario

The pessimistic and baseline scenarios are characterized by high and persistent depreciation, with annual rates of 9.8% and 7.7%, respectively. The stabilization scenario reduces, but does not eliminate, depreciation pressures. Only the optimistic scenario achieves sustained exchange rate stability and modest appreciation, underscoring the critical role of comprehensive

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 14 Issue 12 , December - 2025

reforms and macroeconomic discipline in stabilizing the Ariary.

E. Decision generation

This reverse-engineering framework made it possible to derive decision-oriented insights from the data collected through web crawling, as presented in Table 4.

Table 4. *Model-generated medium- and long-term economic policy targets and strategic measures (2025–2030).*

Area	2025 Objective	2030 Objective	Strategic Measures
Monetary	Inflation below 8%	Inflation between 4–6%	Adjustment of policy interest rates and control of money supply (M2)
Fiscal	Budget deficit below 4% of GDP	Budget deficit below 3% of GDP	Spending efficiency and tax system reform
Sectoral	15% increase in agricultural output value	30% increase in agricultural output value	Agricultural modernization and improved access to sectoral financing
Institutional	Improvement of governance by 10 points	Improvement of governance by 20 points	Enhanced transparency and anti-corruption initiatives
Monitoring	Foreign reserves exceeding 3 months of imports	Foreign reserves exceeding 4 months of imports	Implementation of a key indicator monitoring framework

The table4 highlights the differentiated influence of macroeconomic variables on the dynamics of the Ariary exchange rate. Positive coefficients indicate an association with currency depreciation, while negative coefficients reflect a contribution to exchange-rate stability.

The results show that broad money supply (M2/GDP) and private credit to GDP are the main factors associated with the depreciation of the Ariary. These relationships suggest that rapid monetary and financial expansion may increase inflationary pressures and strengthen import demand, which weakens the external value of the currency. Likewise, investment and external aid appear to be correlated with depreciation, which may reflect dependence on external financing or an orientation of investment toward import-intensive activities.

Conversely, some variables contribute to exchange-rate stability. The central bank policy rate shows a stabilizing effect, indicating that monetary policy plays an important role in anchoring expectations and regulating liquidity. Unemployment, external debt, the trade balance, and agricultural value added are also associated with relative stability, which may be explained by a moderation of domestic demand or structural adjustment effects.

Overall, these results indicate that the dynamics of the Ariary are shaped by a combination of monetary, financial, and structural factors. They also highlight the importance of prudent management of liquidity, credit, and external balances in supporting monetary stability.

V. DISCUSSION

A. Interpretation of Data Extraction and Modeling Strategy

Regarding the reverse-engineering process, which aims to build platform-independent models from web-based data interfaces, this work extracted usable datasets that may serve as specific models for platform-dependent systems (PSM). In addition, a transformed model corresponding to the set of decisions generated was developed as a platform-independent model (PIM). Therefore, this reverse engineering approach corresponds to the inverse process of Model-Driven Architecture (MDA)[3] [15].

In line with the principles of legal web crawling, the data were collected exclusively from open and authorized sources, including the World Bank API, FAOSTAT, and INSTAT Madagascar, all of which comply with formal data-use policies. However, for datasets that are not exposed through APIs, additional legal crawling procedures are required, which represents a current limitation of the present crawling framework.

With respect to decision generation, the study was able to produce structured and interpretable decision outputs in tabular form. Nevertheless, a more advanced perspective would consist in extending these outputs toward a fully specified decision-modeling framework.

B. Interpretation of Reverse Engineering Applied to the Econometric Model

The analysis conducted using the multiple linear regression model and shock simulations made it possible to identify the main macroeconomic factors influencing the dynamics of the Malagasy Ariary exchange rate. The results clearly indicate that certain monetary and financial variables, such as broad money (M2) and private credit as a share of GDP, are strongly associated with currency depreciation. This finding is consistent with economic theory, which suggests that rapid expansion of liquidity and credit can increase import demand and inflationary pressures, thereby weakening the external value of the currency.

In contrast, several structural and institutional variables appear to have a stabilizing effect. The central bank policy rate plays a key role in anchoring expectations and regulating liquidity, while unemployment, external debt, the trade balance, and agricultural value added contribute to reducing exchange rate volatility. These effects highlight the importance of structural adjustments and prudent management of external balances in supporting monetary stability.

The reverse-engineering modeling approach enabled the development of independent models based on data extracted from open sources. This method allowed not only the construction of platform-specific models (PSMs) but also the generation of a consolidated transformed model integrating the inferred decisions (PIM). The extracted decisions provide valuable information set for economic analysis and planning, although further progress could be achieved by directly modeling the decision-making process itself.

Overall, the findings show that exchange rate dynamics result from a combination of monetary, financial, and structural factors. Quantitative modeling and scenario simulations

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 14 Issue 12 , December - 2025

therefore offer a useful tool for assessing the impact of monetary and financial policies and for anticipating the effects of economic shocks on currency stability.

VI. CONCLUSION

This paper proposes a reverse-engineering–based framework designed to transform multi-source web interfaces into independent analytical system models for economic analysis. The study introduces a legal crawling framework that combines API-based access with controlled web data collection to support the extraction, modeling, and forecasting of economic indicators.

The empirical results demonstrate that the proposed approach ensures faster and more secure data collection while maintaining legal compliance and data structuring. PCA effectively identifies the most influential variables, multiple linear regression quantifies their impact, and the ARIMAX model enables reliable forecasting. Scenario and sensitivity analyses further enhance the evaluation of optimistic and pessimistic economic conditions, allowing the construction of combined policy scenarios.

The framework successfully extracts relevant data from heterogeneous web sources and transforms them into decision-oriented outputs in tabular form. This multidisciplinary contribution is particularly relevant for the Malagasy context, as it supports evidence-based monetary analysis and contributes to the design of a more stable and effective monetary system.

Nevertheless, the study presents certain limitations. Legal crawling remains constrained by API availability, motivating the adoption of a standardized Legal-Aware Data Extraction Approach (LADEA) that extends beyond API-based methods. In addition, although decision tables are generated automatically, their analytical value could be improved through formal decision modeling using Decision Model and Notation (DMN).

Future research will focus on extending the proposed framework to other economic sectors, such as industry and services, integrating natural language processing techniques for qualitative web data analysis, and developing automated dashboards for enhanced visualization and decision support. Further work will also emphasize the full implementation of LADEA and the formal modeling of extracted decisions using DMN to strengthen both analytical rigor and policy relevance.

REFERENCES

[1] E. Ferrara, P. De Meo, G. Fiumara, and R. Baumgartner, ‘Web data extraction, applications and techniques: A survey’, Knowl.-Based Syst., vol. 70, pp. 301–323, Nov. 2014, doi: 10.1016/j.knosys.2014.07.007.

[2] J. Frechtling, ‘An overview of quantitative and qualitative data collection methods’, 2002 User-Friendly Handb. Proj. Eval., pp. 43–62, 2002.

[3] D. N. B. Dahibhate, D. S. Sadlapur, V. More, and M. K. Ingale, ‘Global Perspectives on Open Science: Bridging Research, Access and Innovations’, 2025.

[4] F. Willekens, ‘Programmatic access to open statistical data for population studies’, Demogr. Res., vol. 49, pp. 1117–1162, 2023.

[5] V. Krotov, L. Johnson, and L. Silva, ‘Tutorial: Legality and Ethics of Web Scraping’, Fac. Staff Res. Creat. Act., Dec. 2020, doi: https://doi.org/10.17705/1CAIS.04724.

[6] S. S. Sohail et al., ‘Crawling Twitter data through API: A technical/legal perspective’, May 22, 2021, arXiv: arXiv:2105.10724. doi: 10.48550/arXiv.2105.10724.

[7] D. Cai, S. Yu, J.-R. Wen, and W.-Y. Ma, ‘VIPS: a Vision-based Page Segmentation Algorithm’, Vis..

[8] J. R. Andrianady, ‘Application of a simple VAR model on the inflation and money relationship in Madagascar’, University Library of Munich, Germany, 2018.

[9] A. Josué, F. M Randriamifidy, M. HP Ranaivoson, and T. Miora Steffanie, ‘Econometric Analysis and Forecasting of Madagascar’s Economy: An ARIMAX Approach’, 2023.

[10] M. A. Brown, A. Gruen, G. Maldoff, S. Messing, Z. Sanderson, and M. Zimmer, ‘Web Scraping for Research:Legal, Ethical, Institutional, and ScientificConsiderations’,Dec. 19, 2024, arXiv: arXiv:2410.23432. doi: 10.48550/arXiv.2410.23432.

[11] M. A. Brown, A. Gruen, G. Maldoff, S. Messing, Z. Sanderson, and M. Zimmer, ‘Web scraping for research: Legal, ethical, institutional, and scientific considerations’, Big Data Soc., vol. 12, no. 4, p. 20539517251381686, Dec. 2025, doi: 10.1177/20539517251381686.

[12] G. A. Di Lucca, A. R. Fasolino, F. Pace, P. Tramontana, and U. De Carlini, ‘WARE: a tool for the reverse engineering of Web applications’, in Proceedings of the Sixth European Conference on Software Maintenance and Reengineering, Mar. 2002, pp. 241–250. doi: 10.1109/CSMR.2002.995811.

[13] L. Favre, ‘MDA-based reverse engineering’, Reverse Eng. Adv. Appl., 2012, Accessed: Jul. 15, 2024.

[14] H. M. Andrianaivo, W. G. Dimbisoa, and T. Mahatody, ‘Reverse Engineering Approach for Analyzing and Transforming Graphical User Interface Source Code into Class Diagrams’, J. Comput. Commun., vol. 13, no. 03, pp. 86–102, 2025, doi: 10.4236/jcc.2025.133007.

[15] J. Gu, ‘An Empirical Study on the Judicial Regulation of Data Crawling Unfair Competition’, Int J Educ Humanit, vol. 9, pp. 61–66, 2023.

[16] R. Mitchell, Web Scraping with Python: Collecting More Data from the Modern Web. O’Reilly Media, 2018.

[17] Y. Bassil, ‘A Survey on Information Retrieval, Text Categorization, and Web Crawling’, arXiv.org, 2012.

[18] M. Greenacre, P. J. F. Groenen, T. Hastie, A. I. D’Enza, A. Markos, and E. Tuzhilina, ‘Principal component analysis’, Nat. Rev. Methods Primer, vol. 2, no. 1, p. 100, Dec. 2022, doi: 10.1038/s43586-022-00184-w.

[19] I. T. Jolliffe and J. Cadima, ‘Principal component analysis: a review and recent developments’, Philos. Trans. R. Soc. Math. Phys. Eng. Sci., vol. 374, no. 2065, p. 20150202, Apr. 2016, doi: 10.1098/rsta.2015.0202.

[20] M. Tranmer and M. Elliot, ‘Multiple linear regression’, Cathie Marsh Cent. Census Surv. Res. CCSR, vol. 5, no. 5, pp. 1–5, 2008.

[21] G. K. Uyanık and N. Güler, ‘A study on multiple linear regression analysis’, Procedia-Soc. Behav. Sci., vol. 106, pp. 234–240, 2013.

[22] B. M. Williams, ‘Multivariate Vehicular Traffic Flow Prediction: Evaluation of ARIMAX Modeling’, Transp. Res. Rec. J. Transp. Res. Board, vol. 1776, no. 1, pp. 194–200, Jan. 2001, doi: 10.3141/1776-25.

[23] R. Mitchell, Web Scraping with Python. LinkedIn, 2020.

[24] W. Python, Python Releases Window, vol. 24, 2021.

[25] D. M. Thomas and S. Mathur, ‘Data Analysis by Web Scraping using Python’, in 2019 3rd International conference on Electronics, Communication and Aerospace Technology (ICECA), Jun. 2019, pp. 450–454. doi: 10.1109/ICECA.2019.8822022.

[26] M. Ezekiel, ‘Methods of correlation analysis.’, 1930.

[27] S. Senthilnathan, ‘Usefulness of correlation analysis’, Available SSRN 3416918, 2019.