Journal of Economics, Finance and Management Studies

ISSN (print): 2644-0490, ISSN (online): 2644-0504

Volume 08 Issue 05 May 2025

Article DOI: 10.47191/jefms/v8-i5-01, Impact Factor: 8.317

Page No: 2546-2557

Random Forest-Based Analysis of Risk and Importance in Selected Macroeconomic Imbalance Procedure (MIP) Scoreboard Indicators: A New Perspective on CLIFS Relevance

Doina Liliana Badea[1], Adina – Maria Voda[2]

[1,2]Bucharest University of Economic Studies, Department of Statistics and Econometrics, Romania

ABSTRACT: The relevance of Macroeconomic Imbalance Procedure (MIP) Scoreboard indicators in predicting financial crisis risk is explored in this study, with particular attention paid to how they relate to the Country-Level Index of Financial Stress (CLIFS). We apply Random Forest and Gradient Boosted Trees models in TIBCO Statistica to data from 2002 to 2021, providing a novel method of research by combining country-level financial stress indices with indicators of internal and external imbalances. The findings demonstrate that while Gradient Boosted Trees provide better predictive performance, Random Forest offers steady error rates with less overfitting. However, the latter tend to underestimate extreme values and overestimate small ones, highlighting areas for improvement. This paper's novelty lies in its interdependent analysis of these indicators, presenting a new framework for studying macroeconomic imbalances and crisis risks. The findings suggest that additional factors influencing CLIFS remain unexplored, providing avenues for further research to refine predictive models for financial crises.

KEYWORDS: CLIFS, MIP Scoreboard, Financial Stability, Machine Learning, Random Forest, Gradient Boosted Trees, Economic Modeling, European Union

I. INTRODUCTION

A. Background of the study

Macroeconomic Imbalance Procedure (MIP) has been implemented as an early-warning mechanism for sovereign crises in the European Union (EU). According to the European Commission (EC), the MIP was introduced during the economic and financial crisis, which started in 2007-2008 with the aim to identify potential risks early on, preventing the emergence of bad macroeconomic imbalances, and correcting the ones that are already in place, that way minimizing the risks of new crises occurrence. After 2008-2009 when countries faced large current account deficits and real estate bubble inflated, became relevant that imbalances in one country could affect others throughout a contagious evolution. This led the EC to develop the procedure, which came into force in December 2011 as part of the 'six-pack' of legislative acts, which strengthens the monitoring of macroeconomic policies in the EU and the euro area. To detect potentially harmful imbalances a scoreboard has been developed consisting of stock and flow indicators mix, which can capture both short-term rapid deteriorations as well as the long-term gradual accumulation of imbalances. The surveillance Scoreboard for the macroeconomic imbalances' detection consists of fourteen headline indicators related to external and internal macroeconomic imbalances as well as employment and social developments. Auxiliary indicators that allow a better understanding of the risks and help identifying relevant policy measures complement the MIP Scoreboard. The Alert Mechanism Report (AMR) is the starting point of the annual MIP. Based on it, the Commission prepares country-specific in-depth reviews (IDR) for the Member State under review, considering country-specific economic conditions. The MIP concept relies on relevant legislation, runs specific monitoring based on Scoreboard, AMR and IDR and if the case, develops excessive imbalance procedure (EIP). The European Commission uses defined standard indicators related to both internal and external macroeconomic imbalances to assess potential macroeconomic vulnerabilities of EU countries. Internal imbalances can arise from public and private indebtedness, financial and asset market developments, including housing, the evolution of private sector credit flow and the evolution of unemployment. External imbalances arise from the evolution of current account and net investment positions of member states, real effective exchange rates, share of world exports and nominal unit labor costs. This study employs a methodology in which a country-level index serves as the dependent variable, with explanatory variables consisting of three external and four internal imbalance indicators drawn from the Scoreboard. The research question is to find a new tackle to study macro-imbalances imminence and financial crisis risk in view to investigate the MIP headline indicators with the most important impact in the analysis. This paper focuses on the interdependent analysis between the selected indicators and the index CLIFS (Country-Level Index of Financial Stress), to assess the significance of those variables in the crisis risk prediction. In that way, the focus only on those selected MIP indicators allows to compare the outcome of this paper with the literature-based analysis developed by various authors. Although there are similarities in the variable selection, the MIP model used includes macroeconomic indicators as the previous models originally developed, with a major addition of the stress index.

B. Problem statement

Since the latest recent economic and financial crisis has shown many macro-indicators exceeding their indicative thresholds for most EU countries, the analyze moved deeper to country specificities, as well as to the interdependence relation among them, given their abnormal values. The Country-Level Index of Financial Stress (CLIFS) evolution across EU member countries displays very accurate the shock encountered during 2007-2008 interval and the complicated recovery afterwards for some other Central, Eastern and Southern countries of EU (**Error! Reference source not found.**). In terms of volatility, it is evident that following a crisis, levels of financial stress remain relatively low and stable for the most of the countries, while there are other periods of increased stress (such as 2011–2012 and 2020, which may be related to the sovereign debt crisis and the COVID-19 pandemic). According to the latest AMR released in November 2022, many EU countries had made up for the GDP decline experienced during the outbreak of the pandemic and were growing slowly right after. The current outlook reveals a combination of high inflation, uncertainty, linked to the energy crisis, economic activity declining, and economic growth forecast not very optimistic. Financing conditions have visibly tightened, while global conditions show a global slowdown and concerns in emerging markets.

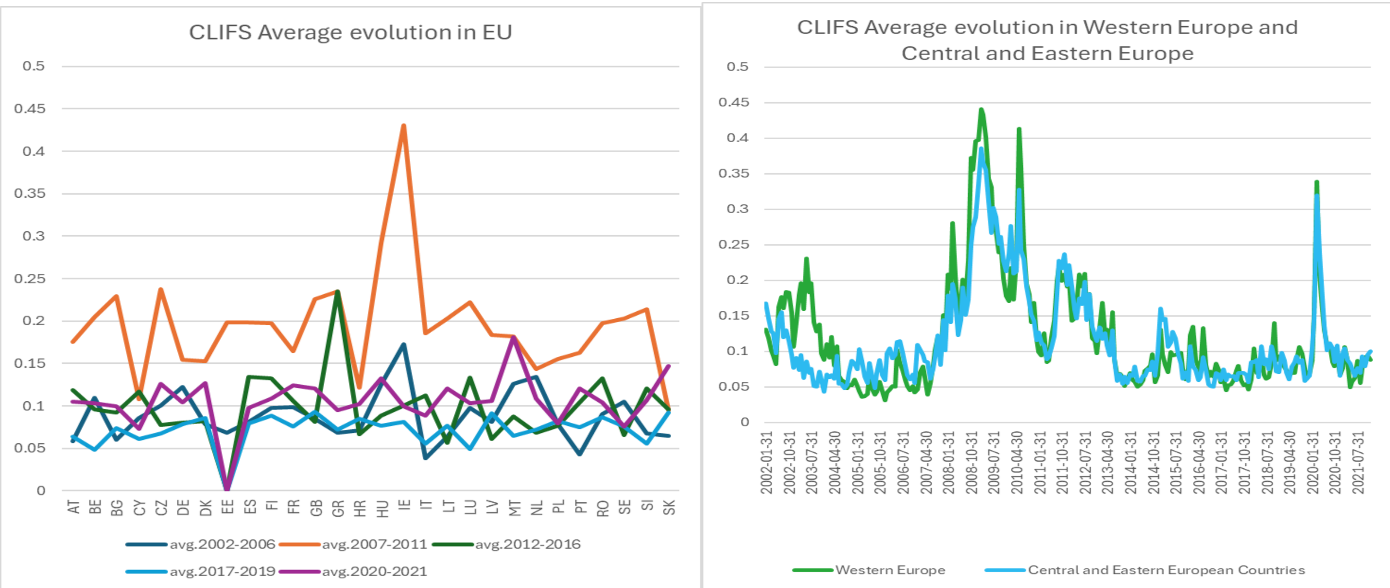

Figure 1 CLIFS evolution within EU, 2002-2021

However, even though there are some studies conducted on MIP topic, there are no specific studies yet related to selected indicators and CLIFS index . Since the topic covers mostly the methodology and headline levels there is an opportunity to conduct a study on this, specifically looking at all EU countries at the period under review. This will be an additional research and knowledge to what already exists on the importance of the before mentioned index and scoreboard indicators.

The paper's remaining sections display as follows. The research directions and results in the existing literature are presented in the following section. Section 3 presents both the data and research methodology. Section 4 provides a summary and discussion of the main findings. The conclusion of Section 5 includes a discussion of the limitations of this research and ideas of future interest.

II. LITERATURE REVIEW

Since subjects such as MIP, Scoreboard and headline indicators are somehow new (dating from the end of 2011), several studies pertaining to this were published, contributing to the general research approach from various perspectives. Macroeconomic imbalances in the EU before the crisis led to the establishment of MIP. The political justification for the MIP is that growing imbalances might have a detrimental impact on other member states due to the spread of crises or the need for expensive interventions, in addition to creating crisis vulnerabilities in individual member states. As a result, when imbalances are observed, demands for improvement are made, and if the reform ideas are deemed insufficient, financial fines may follow. In the years preceding the last financial crisis, concerns were transposed in a study regarding the experience on preventing and controlling such phenomena (Frankel and Wei, 2004). The paper emphasize that the policy decisions made while capital inflows come to an end should have been accelerated swiftly, instead of depleting reserves or drawing short term resources, affecting country's macro indicators. Signals were obvious when short-term liabilities and external foreign debt obligations became high at the same time with foreign direct investment and reserves going down. A three-step way of investigation implied a simple probit analysis to look for the variables capturing the best probability of a currency crisis and secondly, a regression tree analysis to see which variable seem to be more important (technique been used less often than factor analysis), showing robust statistical relationship (Ostrihoň, 2022). The sample consisted of countries from the 1970s to 2004. The third step referred to regression analysis for a cross-section of countries combined with the timing of predicting currency crisis typical profile. What it is worth mentioning, is that the macro-indicators considered consisted of a mix of discipline (inflation, debt, budget deficit), institutional quality (corruption), integration of the finances (the absence of capital controls), currency system (currency pegs), trade to GDP ratio, the nature of inflows (maturity, share of FDI, currency composition), and reserves. The study returned basic conclusions such as:

- inflation is an important predictor of currency crisis linked to high budget deficits.
- the composition variables are relevant for the crisis imminence as for its severity.
- the ratio of short-term debt to reserves is a notably useful indicator.
- the leaders' choices to postpone adjustment caused a severe crisis, and financial constraints are enforced even in the absence of a recession.

Some papers refer to macro-imbalances regarding the relationship between euro and non-euro area countries. Bobeva and Atanasov (2016) tried to measure the risk levels by the range of the deviation from the indicative thresholds, using a new constructed indicator the Integral Macroeconomic Imbalance Indicator, that compared pre- and post-crisis time. The indicator shows specific contraction in the scale of unbalances together with country differences widening, concluding the idea that the two areas will face imbalances regardless the position in or outside. The results also question the AMR capability to see the coming crisis in the context of thresholds macro-frame. The MIP indicators were analyzed also for individual countries as well as for the whole EU. Hence, Torój (2017) focused on the lower limit of -4% on the current account deficit to GDP ratio, using a 2-region, 2-sector New Keynesian DSGE model and Woodford's technique for inequality constraints, with robust results of parameters (in case of Poland and EU). The paper concluded that the method shows lower indicatives on REER and ULC dynamics, even in a conflict with CA constraint, while overall the scoreboard seems to need refining in the future MIP revisions. Dany-Knedlik et al. (2021) studied EC Scoreboard of MIP within the Central and Eastern European countries' (CEEC) experiences with the global financial crisis where the public debt stress that followed have been significantly different from those in other European countries. Accordingly, the adequacy of scoreboard thresholds seems to be set too narrow for this group of countries due to greater and more shifting growth rates across a range of macroeconomic indicators. Firstly, the authors utilized a signals approach to get the best thresholds in the utility function having values between -0.5/+0.5, positive values being useful in crisis prediction. Secondly, the authors use a univariate-probit pooled-panel model, allowing for confidence bands calculation near to the crisis thresholds, with 80% obtained from simple bootstrapping, reporting 10th and 90th percentile of the distribution of optimal thresholds. Thirdly, significant indicators introduced into a multivariate-probit pooled-panel model, to evaluate the early warning system's overall explanatory capacity. Thus, the authors discovered that, among the 32 ideal thresholds that provide positive utility, 19 optimal thresholds are broader than the MIP's thresholds and 10 optimal thresholds are narrower. Only in three instances, they discover that the ideal criteria matched those used in the MIP. The conclusion gave significant indication: current account, international investment position, housing prices, private credit flows, public debt, and financial sector liabilities have all been determined to be significant and helpful in nominal spread crises. Thus, out of the indicators, four have thresholds that are significant and relevant for at least two of the three different types of crises. The authors' work returned that for six indicators (real effective exchange rate, export market share and for four labor market indicators the recommendation would be to drop them on from the scoreboard in case of CEEC. In return, in a working paper S. Erhart (2019) presented the investigation of the lifetime of the 2008-2009 financial turmoil within MIP, in the context of European Systemic Risk Board (ESRB). The study is a novelty since the MIP indicators capacity to early warn test various crisis facts (Balance of payments, Sovereign, Banking and Asset Pricing) and research the MIP Scoreboard's ability to foresee crises by noticing the economy's overheating in advance. The methodology used both signals approach as other empirical studies did before, and the composite indicator envisaged by other authors to evaluate the performance of MIP. The current account was detected as best predicting indicator for all types of crises together with the net international investment position, in line with the rest of the literature found related to the issue as being statistically significant in explaining the turmoil imminence. Since the MIP was conceived to detect and prevent vast and enduring current account deficits that have been experienced by Monetary Union countries, the empirical literature on early detection techniques is now more comprehensive touching the adverse situation issue such as surpluses of current account much above the upper threshold, encountered by countries like Germany and others (Netherlands and Luxembourg). For example, in the case of Germany would be of interest from the imbalances point of view the market share indicator (Gros and Busse, 2013), which was below threshold rather indicating domestic than external growth. The conclusion of the paper stated that from a Keynesian perspective, increased domestic demand in such countries would suggest negligible advantages that wouldn't reduce unemployment in the remainder of the Monetary Union.

Gadea Rivas and Perez-Quiros (2015) used the univariate Markow-Switching models proposed by Hamilton (1989) in a study which tries to look at the evolution of credit indicator as a predictor of crisis, based on a dataset consisting of 39 OECD countries. Contrary to previous literature, these paper findings indicate that credit's role in identifying the economic cycle, and its characteristics is very limited. To summarize the literature review techniques, several additional techniques identification appear relevant. Among them artificial neural networks, present the advantage of complex interaction between the variables (Nag and Mitra, 1999; Peltonen, 2006; Ristolainen, 2018). There are other techniques that are worth mentioning like the value-at-risk models (Bléjer and Schumacher, 1998), restricted VaR models (Krkoska, 2001), Bayesian techniques (Christofides et al., 2016), along with Random Forest (Alessi and Detken, 2018) and Markov-switching approaches, that don't require a predetermined definition of crises (Abiad, 2003; Knedlik and Scheufele, 2008). The MIP was analyzed also using quantile regressions, investigating the connection between the scoreboard indicators and the yield spreads on sovereign bonds. The findings were: two, three, and four quarters ahead of time, MIP indicators can explain the movement of sovereign spreads, but strong non-variant country effects that influence spread evolution and the varying effects that each indicator has on various nations are not captured by the scoreboard.

The overall impact on sovereign spreads has decreased with the introduction of employment metrics (Arahuetes and Gómez, 2018).

III. RESEARCH METHODOLOGY

This part presents the research methodology on Country-Level Index of Financial Stress (CLIFS) used to assess the significance of the macro-indicators in the crisis risk prediction for the period 2007-2021. The theory is broadly connected to the relatively few literature reviews that have been found on this issue and reflects an innovative attempt to find relations between MIP indicators and a financial stress index. Thus, the purpose of the paper is to analyze a new approach for the macroeconomic imbalances imminence and financial crisis risk, with the aim of investigating the MIP headline indicators having the most important influence in their analyze. The specific objectives wish for estimating the relationship among the variables under review, determining the importance of the indicators and issuing statements based on the conclusions drawn within experimentally usage of the CLIFS index.

The section consists of developing a theoretical model built on the significance of the variables in the form of MIP indicators in a turmoil prediction using CLIFS index. The research is in line with previous studies proposed by Breiman (2001) based on Random Forest (RF) method which is a famous machine learning technique involving bagging in the form of bootstrapping and aggregating a high number of trees. The model derived is shown as follows:

CLIFS = f (selected MIP indicators) where

CLIFS is Country-Level Index of Financial Stress

Selected MIP indicators are: *1. Current account balance; 2. General government gross debt; 3. Net international investment position; 4. Private sector credit flow; 5. Private sector debt; 6. Real effective exchange rate; 7. Total financial sector liabilities (“The Macroeconomic Imbalance Procedure (MIP) introduced - Statistics Explained”, n.d.)*.

The study adopted a modified version of a classical regression function upon the need to investigate the importance of the macro-indicators in predicting potential crisis. The model's variables are often stated as a percentage of GDP:

- *Country-Level Index of Financial Stress*, CLIFS – includes six, mainly market-based, financial stress measures that capture three financial market segments: equity markets, bond markets and foreign exchange markets. The importance of the selected indicators has been measured by Country-Level Index of Financial Stress (CLIFS), which it is the dependent variable. Data series are available on the ECB website Data Warehouse.

- *Current account balance*, CAB3 – refer to three-year backward moving average expressed in percent of GDP, based on Eurostat data from Balance of Payments statistics, with the indicative thresholds of +6% and -4%. Frankel and Saravelos (2010) note that the current account balance is one of the most often statistically significant indicators in explaining crisis occurrence based on a thorough literature assessment of 83 publications.
- *General government gross debt*, GGGD – this is the general government debt in percent of GDP, as defined according to Excessive Deficit Procedure (EDP) being the total gross debt at nominal value outstanding at the end of the year, consolidated between and within the sectors of general government. The indicative threshold is 60%. General government debt indicator is included in the scoreboard to provide a more comprehensive picture of Member States' debt levels rather than to track the risks of unsustainable public finances.
- *Net international investment position*, NIIP – according to Eurostat's Balance of Payments statistics, the scoreboard indicator is the net foreign investment position reported as a percentage of GDP, with a default threshold of -35%. The net international investment position (NIIP) compares the various sectors of the domestic economy net financial position (assets minus liabilities) to that of the rest of the world. It gives a comprehensive picture of a nation's net external position and is commonly used in economic research and analysis, with a particular emphasis on crisis risk and a nation's external vulnerability (Frankel and Saravelos, 2010), enabling a stock-flow examination of external situations since it is the stock counterpart to the current account balance. Continually huge current account deficits are the usual cause of very negative NIIPs. In this regard, NIIP has some of the same conceptual difficulties as the current account balances.
- *Private sector credit flow*, PSCF – expressed in percent of GDP, it includes loans and securities other than shares, non-consolidated data. It is the flow counterpart of private sector debt (which is a stock indicator). The indicative threshold of private sector credit is 15%. High loan expansion is empirically linked to a higher incidence of crises (Frankel and Saravelos, 2010), rapidly growing credit is one of the best indicators of financial or banking crises, both in emerging and mature economies.
- *Private sector debt*, PSD – the stock of private sector debt in percent of GDP, defined as the sum of loans and securities other than shares, non-consolidated. The threshold of private sector debt is 160%. The most recent financial crisis served as a reminder that unreasonably high levels of private sector debt enhance sensitivity to economic shocks posing dangers to growth and financial stability.
- *Real effective exchange rate*, REER – the percentage change over three years of the real effective exchange rate (REER) based on consumer price index deflators, with the indicative thresholds of +/–5% and +/–11% for euro-area and non-euro-area countries. This indicator also illustrates why imports are more desirable than domestic manufacturing since it is strongly tied to the terms-of-trade idea. The REER is usually regarded as an early warning signal since it has regularly been demonstrated to be a statistically significant predictor of the occurrence of economic crises (Reinhart et al., 1998).
- *Total financial sector liabilities*, TFSL – The total financial sector liabilities measure the evolution of the sum of all liabilities (currency and deposits, debt securities, loans, equity and investment fund shares/units, insurance, pensions and standardized guarantee schemes, financial derivatives and employee stock options and other accounts payable) of the financial corporations sector. Data are presented in non-consolidated terms, i.e. data consider transactions within the same sector. The MIP indicator is expressed as year over year growth rate, with an indicative threshold of 16.5%.

Data processing and RF estimates were done in Tibco Statistica 13.3. The study makes use of yearly-standardized data from 2002 to 2021 to trace the course of the most recent financial crisis, which began in 2008 and continued through the recovery years and the pandemic years of 2020–2021. Country-Level Index of Financial Stress (CLIFS) measures the financial turmoil likelihood, derived as stated in the data specification, and defines the dependent variable. The descriptive statistics of the variables values across years and countries is presented in Table 1. The analysis of key macro-financial indicators reveals two distinct categories: crisis indicators (CLIFS, PSCF, TFSL) and structural indicators (PSD, GGGD, NIIP). High skewness and excess kurtosis are characteristics of crisis indicators that may point to infrequent but significant financial events. With less kurtosis and a more symmetrical distribution, structural indicators may indicate long-term economic circumstances as opposed to abrupt shocks. When taken as a whole, these disparate distributions demonstrate how both indicator types play complementary functions. Structural indicators may provide depth and context for a more comprehensive economic analysis, while crisis indicators seem to identify imminent threats.

Table 1 Descriptive statistics of the MIP variables

Variable	Mean	Median	Std. Dev.	Skew	Kurt
CLIFS	0.119	0.0908	0.085	2.145055	6.0030
CAB3	-0.808	-0.2500	5.339	-0.530728	0.3282
GGGD	60.503	54.0000	35.963	1.031251	1.3693
NIIP	-34.77	-33.5500	54.017	-0.167316	0.0592
PSCF	6.852	4.6000	11.023	4.649900	47.3076
PSD	136.19	122.2000	68.049	0.942603	0.3976
REER	0.701	0.7000	3.269	0.640714	3.9856
TFSL	1217.346	359.8000	3679.819	5.454520	30.7813

Source: *Authors' own processing*

In order to rule out any potential effects of the conflict crisis in 2022 and its aftermath, we are looking at the period 2002–2021. The sample is based on data available from Eurostat (MIP Scoreboard indicators) and ECB (CLIFS data), enclosing 27 EU member countries (by the end of 2017). The number of observations for each variable varies between 510 and 540 (the lowest number is for CAB3). With respect to CLIFS, the descriptive statistics of the macroeconomic and country variables shows a mean of 0.1, varying between 0.02 and 0.57 depending on the prevalent financial calmness or turbulence. Skewness (2.1) and kurtosis (6.0) indicate positive acceptable asymmetry of the indicator but marking a certain departure from normality. Related to current account balance (CAB3) its mean value recorded -0.8 and varies within -21 and 9.5 across EU. The values distribution of -0.5 and 0.3, respectively, describe a pattern of responses having a normal distribution. In the case of real effective exchange rate (REER), the mean value is 0.7 and the standard deviation 3.2, within a quite large fluctuation interval between -14.5 and 17.7, recording a positive asymmetry of 0.6 (normality) even though too peaked, indicating increases of the indicator across EU member countries related to the trading price competitiveness. The technique employed uses Breiman's (2001) machine learning-based random forest (RF) regression approach, which consists of a forest-like ensemble of decision trees. RF constructs decision trees and combines them to produce a more precise and consistent estimation (or forecast) of the dependent variable based on a collection of independent variables. The technique begins with a root node that contains all observations and divides the data into smaller nodes with an upside-down tree-like structure. This approach combines random feature selection with random sample selection (bagging), where the former requires repeatedly sampling observations with replacement to build the trees. The benefit of RF stems from the simplicity with which it is possible to quantify the relative contribution of each independent variable to the prediction or estimation of the dependent variable. This principle is vital in estimating random forest models since the more variables that are included, the greater the likelihood that the model will be over-fitted. As a result, characteristics that do not contribute value to the prediction are often readily discarded. The out-of-bag sample is 30% of the sample in our case, while the training sample represents 70% of the sample. We calculated 100 trees for CLIFS and set a 5% decrease in training error as the algorithm's stopping condition. Before being included in the algorithm, all variables have been standardized.

In the current research, RF regression uses CLIFS as the dependent variable and seven MIP indicators corresponding to 27-member countries, as the independent variables. The methodology follows a two-step approach, primary based on Random Forest and secondary using Gradient Boosting Trees (GBT). The latter (GBT) provides a prediction model in the form of a collection of decision trees or other weak prediction models, being called a poor learner, usually enabling and outperforming RF. The construction of a gradient-boosted trees model follows the same stage-wise process as previous boosting techniques, but it generalizes other techniques by enabling optimization of any differentiable loss function.

IV. RESULTS

To determine the predictor relevance for CLIFS data, risk estimations and a comparison of actual and projected values, we started by estimating 100 trees for the dependent variable (CLIFS). As with any forecast, there is some uncertainty involved in making one. The risk estimates for the RF algorithms used are determined as residual variances and shown in below. These risk estimates relate to the train vs test samples utilized in the estimation as well as the dependent variable's prediction by the collection of independent factors.

Table 2 Random Forest Risk Estimates from CLIFS Data: Predictive Modeling Results with CLIFS as Response Variable

	Risk estimates (CLIFS Data) / Response: CLIFS	
	Risk Estimate	**Standard error**
Train	0.706	0.105
Test	0.981	0.201

Source: *Authors' own processing*

The risk estimates values for train and test samples of 0.7 and 0.98, respectively reveal the shortcoming linked with the accuracy and the performance of the model, nevertheless the variable rank indicates the most important predictor which is the real effective exchange rate (REER) scoring 1 and secondly, the current account balance (CAB3), with 0.67. Becomes quite straight to explain that both deal with trade competitiveness, the cost of exports and imports. It is worth mentioning the predictors ranking best relate to external imbalances while the next ones show connection to domestic discrepancy (Private Sector Debt – PSD recording 0.63 and Private Sector Credit Flow – PSCF, 0.59). RF performs well with noisy data making it suitable for macroeconomic indicators, particularly relevant in cases like REER which might contain market volatility (Glenn, 2019).

Table 3 Variable Importance Rankings for CLIFS Response Based on Random Forest Model

	Predictor importance (CLIFS Data) / Response: CLIFS	
	Variable Rank	**Importance**
CAB3	67	0.672
GGGD	50	0.503
NIIP	55	0.552
PSCF	60	0.599
PSD	63	0.632
REER	100	1.000
TFSL	52	0.524

Source: *Authors' own processing*

The graphs in **Error! Reference source not found.** demonstrate the key elements mentioned above, while the shape of the residuals suggests the model however performed well, thus the trend line is accurately surrounded by the discrepancies between the expected and actual CLIFS values. Most of the predictions cluster near the red line for lower observed values while some deviations suggest reduced accuracy during extreme stress times. Since the outliers were not removed, they indicate that the model predicts severe stress events accurately.

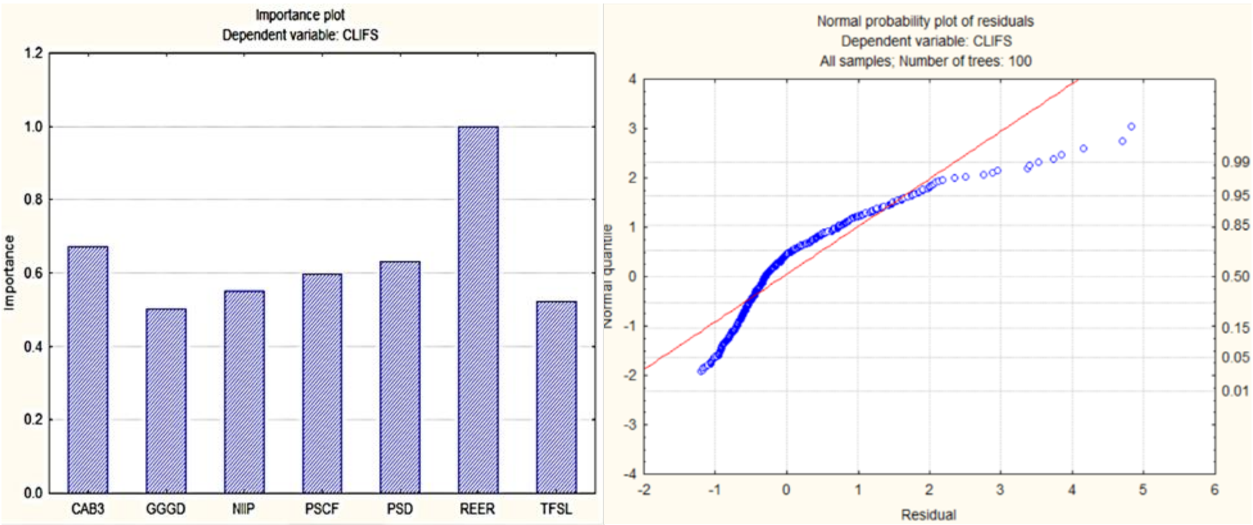

Figure 2 Random Forest Variable Importance Plot for CLIFS Risk Prediction

Source: *Authors' own processing*

Random Forest-Based Analysis of Risk and Importance in Selected Macroeconomic Imbalance Procedure (MIP) Scoreboard Indicators: A New Perspective on CLIFS Relevance

The second stage of the research's approach is to use the Gradient Boosting Trees (GBT) attempting to improve the performance and the accuracy of the model, and to observe the stability of the macro indicators during the second process. We performed then the boosted tree RF, by estimating 188 trees for the dependent variable (CLIFS), expecting to optimize the results and to receive an output with a modified predictor importance ranking and risk estimate. In this case, the outcome was significantly pleasing, confirming initial estimates and assessing the theory behind the ratios.

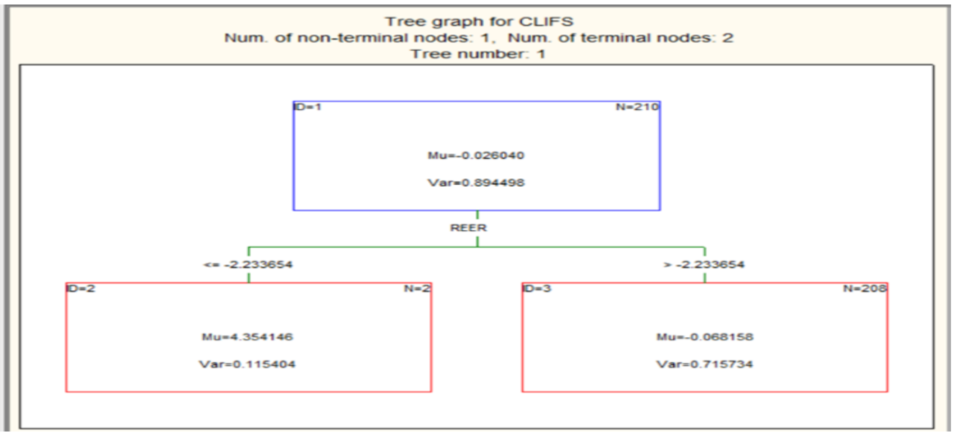

Figure 3 Decision Tree Structure from GBT (Gradient Boosting Trees) **Model Predicting CLIFS Index**

Source: *Authors' own processing*

The tree underlines the first-place importance of the independent variable REER for the dependent variable CLIFS, defining the decision nodes and the statistics behind, including variable variance (**Error! Reference source not found.**).

Table 4 Gradient Boosting Trees (GBT) Risk Estimates from CLIFS Data: Predictive Modeling Results with CLIFS as Response Variable

	Risk estimates (CLIFS Data) / Response: CLIFS	
	Risk Estimate	**Standard error**
Train	0.565	0.071
Test	0.723	0.177

Source: *Authors' own processing*

Table 5 Variable Importance Scores from Gradient Boosting Trees (GBT) Model for CLIFS Prediction

?	**Predictor importance (CLIFS Data) / Response: CLIFS**	
	Variable Rank	**Importance**
CAB3	70	0.704
GGGD	51	0.508
NIIP	74	0.741
PSCF	58	0.577
PSD	73	0.727
REER	100	1.000
TFSL	51	0.508

Source: *Authors' own processing*

In case of GBT, the real effective exchange rate (REER) confirms the score of 1 as in the previous modelling step. Despite recording a better score of 0.7 than previously, the current account balance (CAB3) falls to third place after being surpassed by the Net International Investment Position (NIIP), which ranks second with a score of 0.74 (**Error! Reference source not found.**). The accuracy and the performance of the model display for the train and the test sample 0.56, respectively 0.72, within a much tighter error interval (**Error! Reference source not found.** and 5). The three variables REER, NIIP, and CAB3, which all indicate a country's relationship with the rest of the world in terms of trade competitiveness and external indebtedness, are pertinent to the study's research since they have a significant influence on the CLIFS variable.

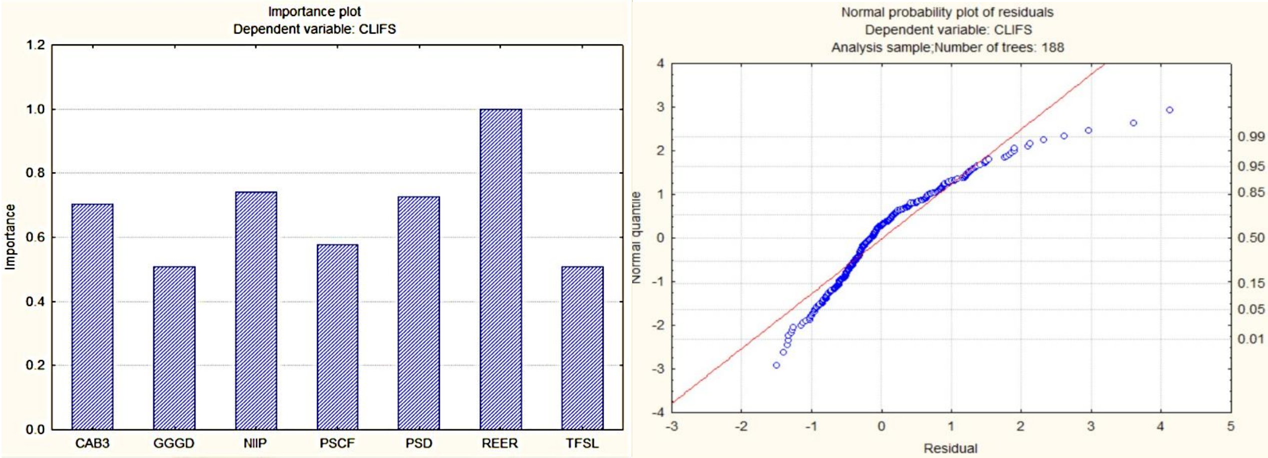

Figure 4 Variable Importance Plot from Gradient Boosting Trees (GBT) Model Predicting CLIFS

Source: *Authors' own processing*

The Figure 5, first graph, illustrates the relationship between observed (X) and predicted (Y) values for the CLIFS index using a Random Forest model with 100 trees. On both axes, the most of the values are concentrate around small values (close to zero). The red line is a representation of the regression between observed and predicted values. The fact taht the model frequently underestimates or overestimates extreme values is indicated by the significant dispersion of points toward the red linse, especially for CLIFS valmeș. The approach tends to predict values that are moderately close, making it challenging to detect values that are very large or very small in the index. In the second graph we see the same relationship between observed (X) and predicted (Y) values for the CLIFS index, using a Gradient Boosting Trees model with 188 trees. Since the points are more widely distributed on axa Y, the red line of recurente appears to follow data trend more precise, and compare to Random Forest, The Boosted Trees model appears to predict observed values more accurately, including extreme values. However, there is still a tendency to underestimate very large values and overestimate very small values, making the general correlation between observed and predicted values better than in the first graph.

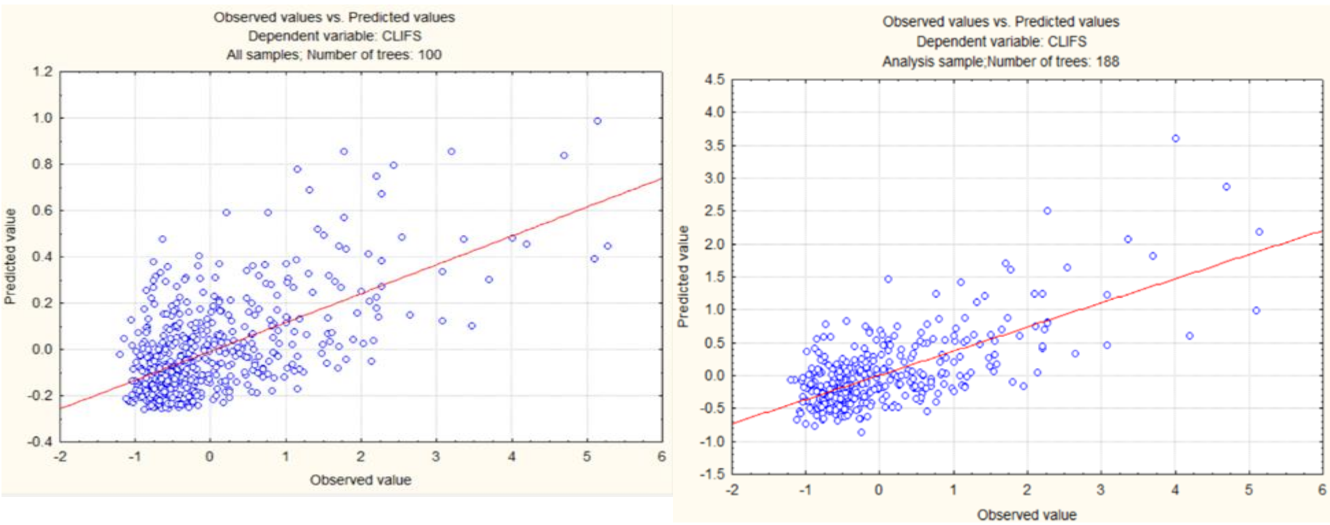

Figure 5 Random Forest (RF) model Plot, Gradient Boosting Trees (GBT) model plot

Source: *Authors' own processing*

The average squared error (Figure 6) calculated for the train and test data in both RF and GBT situations shows values that tend to decline between the range of 0.7 and 0.99 or, even better, that decline dramatically until it reaches the ideal value very close to 0.7. Since there is no ideal MSE value, the model fits data better when the predicted value is lower. In other words, choosing one prediction over another is where the MSE value is best used, as in the case of the current study. The 188-tree model shows some improvement over the 100-tree model, suggesting returns from increased model complexity itself. We see resistance to overfitting the observed vs.predicted CLIFS values shows the prediction pattern even with rare high-stress events as in the current study. RF capabilities relies on strong feature selection which is essential when working with multiple MIP indicators that may have varying levels of importance (Cao, Y., 2024). From a statistical point of view stable indicators like CAB3 and NIIP reflect structural imbalances rather than crisis-specific dynamics.

The errors start to decline quickly and then stabilizes as the number of arbori increases, for test and train data. Moreover, we note reduced overfitting, meaning the difference between the test data error (larger) and the training data error (smaller) is relatively constant, indicating that the model did not experience severe overfitting. In terms of performance, the test data error stays around values of 1, which gives an indication of the model's accuracy for CLIFS prediction.

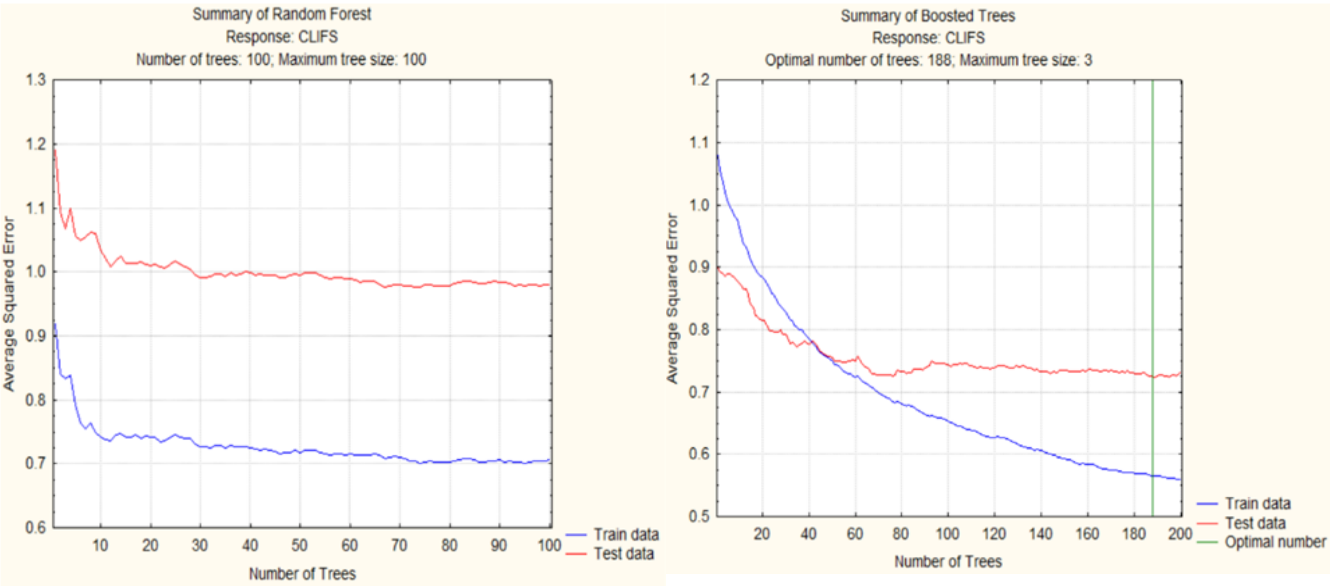

Figure 6 Average squared error values for CLIFS, RF - GBT plots

V. CONCLUSIONS

The analysis covers a considerable range of years, from 2002 to 2021, all the 27 EU member states using machine-learning-based RF regression. Real Effective Exchange Rate (REER), Net International Investment Position (NIIP), and Current Account Balance (CAB3) are the factors that have the greatest impact on the CLIFS (Country-Level Index of Financial Stress). Additionally, Private Sector Debt (PSD) and Private Sector Credit Flow, two macro indicators that indicate both stock and flow, together with the previously stated external imbalances, demonstrate relevance for CLIFS variable and crisis prediction. To conclude, our results contribute to a better understanding of the seven selected MIP indicators in their relation to CLIFS index and to turmoil imminence. These results confirm that the CLIFS Index allows the identification of stress events within specific market segments and in situations where the correlation between sectors intensifies. Thus, the index not only measures the occurrence of financial stress, but also highlights the potential for propagation between sectors in unfavorable conditions (Martínez-Jaramillo *et al.*, 2022).

In our case study Random Forest stabilizes fast the error without significant overfitting, resulting in relatively stable test results. Up to a certain point, boosted trees continue to improve the error of the test data, providing predictive performance superior to Random Forest for the selected dataset. However, there is still a tendency to underestimate very large values and overestimate very small values, making the general correlation between observed and predicted values better than in the RF model. We also may observe that there is a significant dispersion between the observed and predicted values, indicating a positive correlation between them. This suggests that there are additional factors influencing CLIFS that are not fully accounted for by the model and which could be explored in future related research materials. Overall, Boosted Trees have better predictive performance than Random Forest for this set of data, although both models can be improved to better capture extreme or stress values.

Amador and Alves (2020) outlines which macroeconomic variables are the most important for country classification according to MIP categories using machine learning techniques (RF algorithm) and verified that, to categorize countries that require corrective action, particularly through economic adjustment programs, the algorithm highlights the current account balance and the net international investment position as critical factors. This conclusion supports our current study outcome related to MIP and machine learning tools in risk imminence analysis. Casabianca et al. (2022) confirmed that the integration of machine learning models, such as RF or GBT, leads to performance in crisis prediction, outperforming traditional models such as logit regression, in both tests on training and novel datasets. At the same time, the CLIFS index's ability to capture stress in different financial market segments and predict systemic events through improved variable selection and predictive analytics strengthens the robustness of conclusions (Gu et al., 2024).

The current paper's objectives are to: (i) increase understanding of scoreboard headline indicators in connection to the country financial stress index; (ii) carry out effective experimental research; and (iii) offer a review of the literature to other researchers who wish to conduct analyses on the MIP topic. The findings can therefore be used by officials to comprehend the evolution of the data, possible hazards, and the significance of the variables in order to further explore those policies that are important. We see the study on the significance of seven chosen MIP indicators for crisis prediction, which was conducted using a new method like the CLIFS index and Random Forest in conjunction with Gradient Boosting Trees, as a major accomplishment.

Future research on the issue may concentrate on a broader data set and additional macro indicators related to social and employment changes, given that some of them are known to inflate over time, contributing to the precision of the econometric model used for data analysis. Since the paper compares two RF models (100 treesvs.188 trees) and evaluates their predictive performance, it also highlights returns from increasing model complexity itself suggesting data quality are equally important. By demonstrating the limitations of RF and GBT in predicting rare crisis events, the paper lays the groundwork for integrating mixed methods or hybrid approaches to improve accuracy.

REFERENCES

1) Glenn, S. Decision Tree vs Random Forest vs Gradient Boosting Machines: Explained Simply https://www.datasciencecentral.com/decision-tree-vs-random-forest-vs-boosted-trees-explained/
2) Cao, Y. (2024). Prediction and analysis of corporate financial distress based on Random Forest model and GBDT. SHS Web of Conferences, 188, 01001. https://doi.org/10.1051/shsconf/202418801001
3) Abiad, Abdul, (2003), Early Warning Systems: A Survey and a Regime-Switching Approach, No 2003/032, IMF Working Papers, International Monetary Fund, https://EconPapers.repec.org/RePEc:imf:imfwpa:2003/032
4) Alessi, L., Detken, C. (2018), “Identifying excessive credit growth and leverage”, Journal of Financial Stability, Elsevier, Vol. 35, pp. 215–225, doi: 10.1016/J.JFS.2017.06.005.
5) Bobeva, D.N., Atanasov, A. (2016), “Macroeconomic Imbalances in Euro- and Non-Euro Area Member States”, SSRN Electronic Journal, Elsevier BV, doi: 10.2139/SSRN.2818985.
6) Breiman, L. (2001), “Random forests”, Machine Learning, Springer, Vol. 45 No. 1, pp. 5–32, doi: 10.1023/A:1010933404324/METRICS.
7) Casabianca, E.J., Catalano, M., Forni, L., Giarda, E., Passeri, S. (2022), “A machine learning approach to rank the determinants of banking crises over time and across countries”, Journal of International Money and Finance, Pergamon, Vol. 129, p. 102739, doi: 10.1016/J.JIMONFIN.2022.102739.
8) Christofides C., Eicher T., Papageorgiou C.(2016),”Did established Early Warning Signals predict the 2008 crises?, European Economic Review, Volume 81, 2016, Pages 103-114, ISSN 0014-2921, https://doi.org/10.1016/j.euroecorev.2015.04.004, https://www.sciencedirect.com/science/article/pii/S0014292115000604
9) Dany-Knedlik, G., Kämpfe, M., Knedlik, T. (2021), “The appropriateness of the macroeconomic imbalance procedure for Central and Eastern European Countries”, Empirica, Springer, Vol. 48 No. 1, pp. 123–139, doi: 10.1007/S10663-020-09471-9/TABLES/5.
10) Erhart, S. (2019), “Exposition, Climax, Denouement: Life-Cycle Evaluation of the Recent Financial Crisis in the EU by Linking the ESRB Financial Crisis Database to the European Commission’s Macroeconomic Imbalance Procedure Scoreboard”, SSRN Electronic Journal, Elsevier BV, doi: 10.2139/SSRN.3723438.
11) Frankel, J.A., Saravelos, G. (2010), “Are Leading Indicators of Financial Crises Useful for Assessing Country Vulnerability? Evidence from the 2008-09 Global Crisis”, Cambridge, MA, doi: 10.3386/W16047.
12) Frankel, J.A., Wei, S.-J. (2004), “Managing Macroeconomic Crises”, Cambridge, MA, doi: 10.3386/W10907.
13) Gadea Rivas, M.D. and Perez-Quiros, G. (2015), “The Failure to Predict the Great Recession—A View Through the Role of Credit”, Journal of the European Economic Association, Oxford Academic, Vol. 13 No. 3, pp. 534–559, doi: 10.1111/JEEA.12122.
14) Gros, D., Busse, M. (2013), “The Macroeconomic Imbalance Procedure and Germany: When is a Current Account Surplus an ‘Imbalance’?”, 13 November.

15) Gu, X., Mamon, R., Duprey, T. (2024), "Interfacing learning methods for anomaly detection in multi-country financial stress indicators", Knowledge-Based Systems, Elsevier, Vol. 294, p. 111712, doi: 10.1016/J.KNOSYS.2024.111712.
16) Hamilton, J. D., (1989), "A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle." Econometrica, vol. 57, no. 2, pp. 357–84. JSTOR, https://doi.org/10.2307/1912559. Accessed 18 Oct. 2024.
17) Knedlik, T., Scheufele, R., (2008), FORECASTING CURRENCY CRISES: WHICH METHODS SIGNALED THE SOUTH AFRICAN CRISIS OF JUNE 2006?, South African Journal of Economics, 76, issue 3, p. 367-383, https://EconPapers.repec.org/RePEc:bla:sajeco:v:76:y:2008:i:3:p:367-383.
18) Krkoska, L., (2001), Assessing Macroeconomic Vulnerability in Central Europe, Post-Communist Economies, 13, issue 1, p. 41-55, https://EconPapers.repec.org/RePEc:taf:pocoec:v:13:y:2001:i:1:p:41-55
19) Martínez-Jaramillo, S., Montañez-Enríquez, R., Busch, M.O., Ramos-Francia, M., Rodríguez-Martínez, A., Manuel Sánchez-Martínez, J. (2022), "Stress-ridden Finance and Growth Losses: Does Financial Development Break the Link?"
20) Nag, Ashok & Mitra, Amit. (1999), "Neural networks and early warning indicators of currency crisis, Reserve Bank of India Occasional Papers. 20. 183-222.
21) Ostrihoň, F. (2022), "Exploring macroeconomic imbalances through EU Alert Mechanism Reports", European Journal of Political Economy, North-Holland, Vol. 75, p. 102188, doi: 10.1016/J.EJPOLECO.2022.102188.
22) Peltonen, Tuomas A., (2006). "Are emerging market currency crises predictable? A test," Working Paper Series 571, European Central Bank.
23) Reinhart, C. M., Reinhart, V. R. (1998), "Some Lessons for Policy Makers Who Deal with the Mixed Blessing of Capital Inflows". Capital Flows and Financial Crises, edited by Miles Kahler, Ithaca, NY: Cornell University Press, 1998, pp. 93-127. https://doi.org/10.7591/9781501731402-007
24) Ristolainen K., (2018), "Predicting Banking Crises with Artificial Neural Networks: The Role of Nonlinearity and Heterogeneity," Scandinavian Journal of Economics, Wiley Blackwell, vol. 120(1), pages 31-62, January.
25) Schumacher, L., Bléjer, M., (1998), Central Bank Vulnerability and the Credibility of Commitments: A Value-at-Risk Approach to Currency Crises, No 1998/065, IMF Working Papers, International Monetary Fund, https://EconPapers.repec.org/RePEc:imf:imfwpa:1998/065.
26) "The Macroeconomic Imbalance Procedure (MIP) introduced - Statistics Explained". (n.d.). , available at: https://ec.europa.eu/eurostat/statistics-explained/index.php?title=The_Macroeconomic_Imbalance_Procedure_(MIP)_introduced (accessed 11 October 2024).
27) Torój, A. (2017), "Managing external macroeconomic imbalances in the EU: the welfare cost of scoreboard-based constraints", Economic Modelling, North-Holland, Vol. 61, pp. 293–311, doi: 10.1016/J.ECONMOD.2016.10.009.