Freedom in the World: A Data-Driven Analysis of Global Trends

Alfio Leanza 910235 Riccardo Tavecchio 852211

June 10, 2025

Abstract

This study analyzes the evolution of political and civil freedom in the world from 2003 to 2024 using data from *Freedom House*. The objective is to explore global and regional trends and assess the possibility of predicting future scores through ARIMA statistical models. The results highlight a growing polarization between free and authoritarian countries and show the validity of predictive models in stable contexts.

Contents

1 Introduction

Individual freedom is one of the basic pillars of democratic societies and is an essential indicator for assessing the degree of political and civic development of a country. However, its distribution around the world is far from homogeneous, and recent years have seen an increase in authoritarian regimes, armed conflicts and flawed electoral processes, resulting in a reduction in fundamental rights and freedoms in many geographical areas.

In this context, the organization *Freedom House* provides a structured set of data annually, used by researchers, institutions and international organizations to monitor trends in political and civil freedom globally. The scores assigned to individual countries are based on rigorous criteria and allow for comparative analysis over time and space.

The objective of this project is twofold: first, to provide an exploratory analysis of freedom around the world through historical data collected by Freedom House, and second, to test whether it is possible to model and predict the future evolution of a country's freedom from its score and any temporal patterns. This investigation takes on particular relevance in the current context, in which democratic progress is no longer taken for granted, and indeed seems to be regressing in various geopolitical scenarios.

The work will be developed through the analysis of datasets made available by Freedom House covering the period between 2003 and 2024, followed by an attempt to construct a predictive model that can, within certain limits, anticipate future trends in political and civil freedom scores in individual states.

2 Sources and Description of Data

The analysis is based on two main files provided by Freedom House, which document trends in political and civil freedom in more than 190 countries over the period from 2003 to 2024:

- **Aggregate Category and Subcategory Scores FIW 2003–2024.xlsx**: details the scores obtained by each country in the different categories and subcategories used for the overall assessment.

In addition, the report *Freedom in the World 2024* was consulted, which provides the theoretical, methodological and geopolitical context of the index used, as well as a qualitative analysis of key global trends.

Each country receives a numerical score on a scale from 0 to 100, where higher values indicate greater respect for basic freedoms. The scores are divided into two macro-components:

- **Political Rights (Political Rights)**: include aspects such as the electoral process, political pluralism and the functioning of government.
- **Civil Liberties (Civil Liberties)**: include freedom of expression, association, rule of law and personal autonomy.

The final classification of countries is defined in three categories:

- *Free*
- *Partly Free*
- *Not Free*

The first file (`Frredom in the World 2013--2024.xlsx`) collects complete data with no missing values for the years 2006 to 2024. In contrast, the second file (`Aggregate Category and Subcategory Scores FIW 2003--2024.xlsx`) also includes the years 2003, 2004, and 2005, but incompletely: in those first three years only the country name, year of edition, and total score are available, while all other information is missing. Therefore, only from 2006 onward is a complete dataset available for all variables of interest.

Countries are divided into six major geographic macro-regions:

- **Europe**
- **Asia** (also includes the countries of Oceania, which are not treated separately)
- **Eurasia** (includes states belonging to the former Soviet Union, both in Europe and Asia)
- **Africa**
- **Americas** (includes North, Central and South America)
- **Middle East** (considered as the central region of the Asian continent)

During the analysis process, data were cleaned and unified beforehand to ensure consistency among sources, correct any temporal misalignments, and handle missing or outlier values.

3 Methodology

3.1 Data preprocessing

In a first step, the data were imported and unified, and any missing values for variables critical for predictive analysis were removed. Data for the years 2003 to 2005, containing only total scores and not disaggregated components, were used for descriptive purposes only.

3.2 Exploratory analysis

The main objective of the exploratory phase of the analysis was to understand the distribution and evolution of political and civil freedom at the global, regional and national levels during the period under consideration. For this purpose, several visualizations and statistical summaries were produced.

The total scores of each country were analyzed through boxplots (image 1) to detect the distribution, presence of outliers, and asymmetry between regions. The results indicate a clear bias between countries classified as *Free* and those *Not Free*, with a relatively smaller middle range.

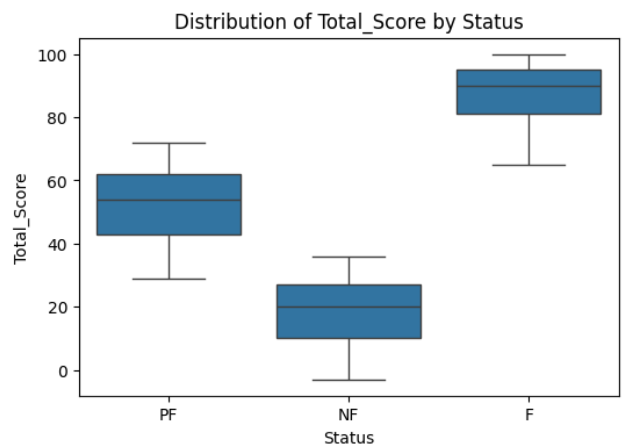

Figure 1: Boxplot of the total score for each category of freedom (F, PF, NF)

The boxplot shows the distribution of the total score assigned by Freedom House to countries, broken down by freedom status: *Free (F)*, *Partly Free (PF)* and *Not Free (NF)*.

As expected, *Free* countries show significantly higher mean and median scores, with a concentrated distribution between 80 and 100. In contrast, *Not Free* countries show much lower and dispersed values, often below 30. The intermediate *Partly Free* category falls between the two, with greater internal variability.

This graph confirms the consistency between qualitative classification and numerical scores, showing a clear separation between the three bands of freedom.

By means of graphs (graph 2) with bars, the average distribution of freedom was analyzed for each of the six geographic macro-areas defined by Freedom House.

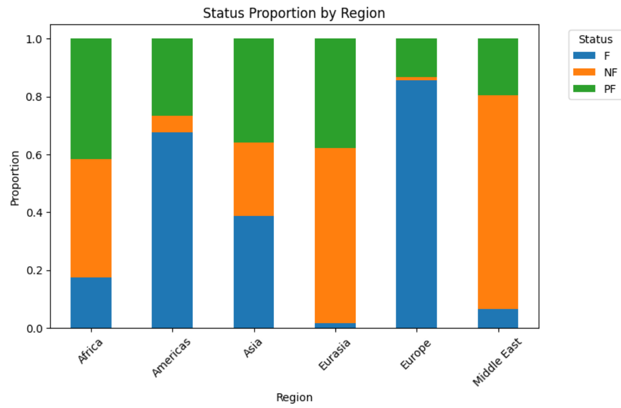

Figure 2: Proportion of freedom status by geographic region

The bar graph shows the proportion of countries classified as *Free (F)*, *Partly Free (PF)* and *Not Free (NF)* within each geographic macro-region.

There is a clear predominance of free countries in Europe and the Americas, with marginal or no presence of non-free countries. In contrast, the Middle East, Eurasia and Africa regions show a very high share of *Not Free* countries, with minimal levels of freedom. Asia shows a relative balance among the three categories, while Africa is characterized by a high incidence of partially free countries.

This regional comparison confirms the inequalities in the distribution of freedom globally, highlighting the concentration of established democracies in some areas and the persistence of authoritarian regimes in others.

The graph 3 shows the trend in the proportion of countries classified as **Free (F)**, **Partly Free (PF)** and **Not Free (NF)** according to annual data from *Freedom House*.

There is a clear reduction in the share of free countries (F) from about 45% in 2006 to just over 40% in 2024. In parallel, the proportion of non-free countries (NF) has increased steadily, surpassing the share of partially free countries (PF) since 2021. The latter group shows relative stability until 2020, followed by a slight decline.

These trends confirm a global regression of political and civil liberties, with a growth of autocracies and a gradual shrinking of democratic space in many areas of the world, as discussed in the annual report of *Freedom in the World*.

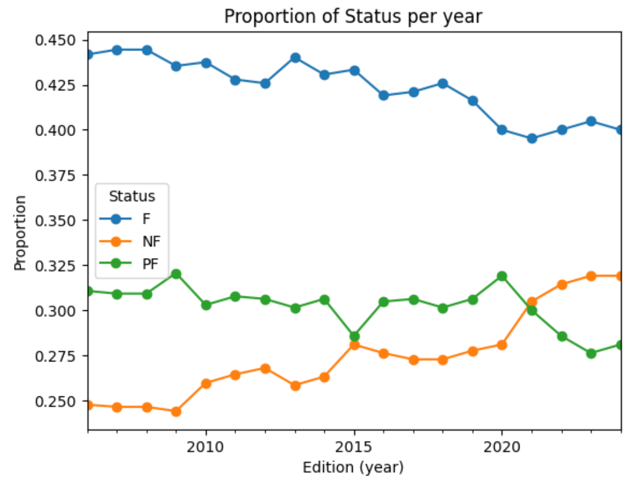

Figure 3: Evolution of the proportion of freedom in countries

3.3 Predictive modeling

.

The main predictive objective is to estimate a country's total freedom score given historical trends and available variables. A time-series model was implemented for this purpose.

3.4 Predictions based on ARIMA models

A time-series analysis using ARIMA models was conducted, with the aim of predicting the future evolution of the overall freedom score for selected countries.

This approach proved particularly useful for estimating the trend of univariate time series based solely on the evolution of total scores over time. After verifying the stationarity of the series through statistical tests, the optimal parameters (p, d, q) were identified through the analysis of ACF and PACF graphs.

The ARIMA model was then calibrated to historical data, and then forecast over the next three years.

4 Predictive Model

This section presents results obtained through the use of predictive models, with the aim of estimating the future evolution of a country's freedom score.

4.1 ARIMA Model for Historical Series

To refine the predictions in the time domain, an ARIMA model was applied to the time series of the total freedom score of individual countries. The process involved:

- verification of the stationarity of the series,
- identification of optimal parameters (p, d, q) by ACF/PACF analysis,
- fitting of the model to the data,
- prediction of the score for the next 3 years.

ARIMA predictions showed a good fit in cases where scores varied smoothly over time. However, in more unstable contexts or those subject to significant policy discontinuities, the model tended to underestimate sudden changes. This highlights the limitations of univariate models in capturing complex dynamics not reflected in purely numerical historical data.

5 ARIMA Analysis by Continent

To better understand regional dynamics, a disaggregated analysis by geographic macro-area was conducted.

5.1 Eurasia

The outcome of the ADF test on the raw series-with a statistic of 5.087 and p-value of 1.000-confirms beyond doubt the presence of a unit root and systematic trend in the dynamics of Total for Eurasia. A p-value of 1 signals that the series is not at all stationary and that year-on-year fluctuations continue to follow a drift path, probably driven by long-term structural political factors. This result forces the removal of this trend component by differentiation in order to meet the assumptions of homoschedasticity and white error required by an ARIMA model.

The first differentiation was not sufficient to remove non-stationarity, as shown by a still very high ADF p-value (0.998). We therefore proceed with a second differentiation: 2nd diff ADF statistic = -0.695, p-value = 0.848.

We proceed with the third differentiation because of the still larger p-value at 0.05.

The third differentiation (d = 3) failed to make the series fully stationary (ADF p = 0.137), however, the ACF/PACF plots on ts_diff3 show the residual dynamics of short-term fluctuations. In the literature, when "pure" differentiation does not bring the p-value below threshold, one can proceed with d=3-while interpreting the results with caution-or consider alternative detrending strategies (e.g., regression on time). For now, we use d=3 and choose p and q based on the visual patterns on these stem plots.

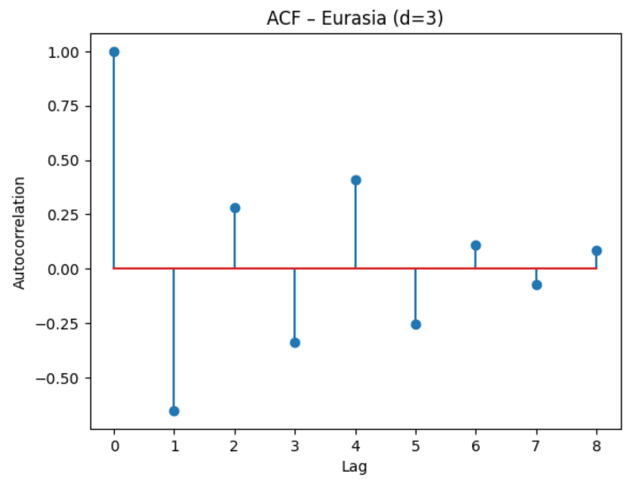

Figure 4: Eurasia ACF

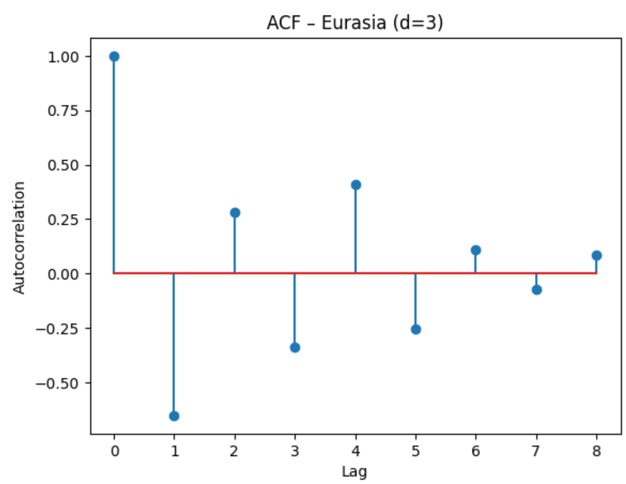

Figure 5: Eurasia PACF

When we compute ACF and PACF we want to explore up to ten years lag-a horizon that in annual series is reasonable to capture medium-term dependencies-but at the same time we must respect the fundamental constraint of PACF, which cannot estimate partial correlations for more than $[N/2] - 1$

lag, on pain of unstable estimates or computational errors. With 22 initial observations, $[22/2] - 1 = 10$, whereas, for example, after a differentiation (d = 1) 21 real data remain and the limit drops to $[21/2] - 1 = 9$ lag. Therefore we adopt: so that at most ten lags are explored, never exceeding half of the actual sample minus one. This ensures that each autocorrelation coefficient is computed on a sufficiently large number of $(xt, xt - l)$ pairs to give statistically significant estimates, preserving the methodological validity of the ACF PACF analyses and, consequently, the rigorous choice of p and q parameters for the ARIMA model.

The presence of significant peaks in PACF at lag 1 and 3 and the gradual, alternating decay of ACF are characteristic of an AR(p) model with p3. The absence of a sharp cutoff in ACF contra-indicates a simple MA model, so q=0 is the most parsimonious choice that captures residual autocorrelations.

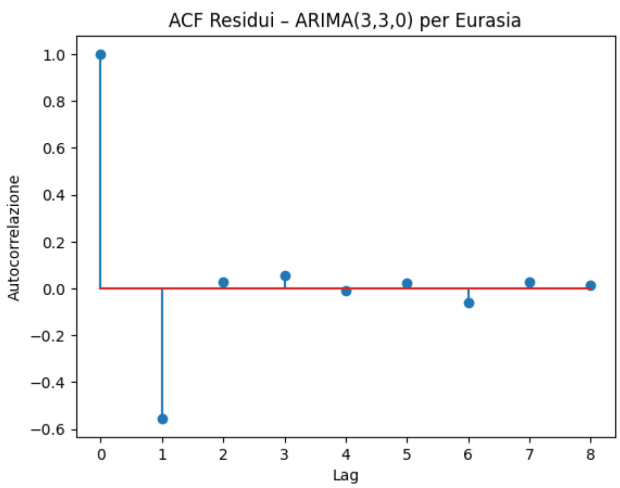

Figure 6: Eurasia Residuals

In the Ljung-Box test, the lags parameter indicates up to which lags we want to simultaneously test for the absence of autocorrelation in the residuals. Choosing lags = 5 means that we are testing the null hypothesis "the residuals are not autocorrelated for all lags from 1 to 5."

We do this for two reasons: Horizon relevant to annual data with annual series, a five-year cycle is often the minimum period in which significant policy patterns or economic shocks emerge. Testing up to lag 5 ensures that we capture any residual medium-term dependencies (a five-year period), without pushing too far into lags for which the sample provides very few observation-delay pairs and thus unstable estimates.

Balancing statistical power vs. degrees of freedom With each additional lag, the test loses one degree of freedom. With 22 observations, using very high lags would excessively reduce the degrees of freedom, lowering the power of the test. The rule of thumb is to keep below about N (here 22 4.7) or one-fifth of the sample (22/5 4.4). Choosing 5 lag is thus a good compromise: enough to cover five-year cycles, but not so many as to empty the efficacy test.

The ACF of the residuals shows a single negative peak at lag 1 (about -0.55), while for all other lags the coefficients remain close to zero and within the confidence bands. The Ljung-Box test at lag 5 gives a p-value of 0.203, well above the 10 % threshold, confirming that there is no evidence of systematic residual autocorrelations. This means that the ARIMA(3,3,0) model has effectively captured all linear dependencies present in the stationary series, leaving residuals that can be likened to white noise-an essential requirement for the reliability of parameter inferences and the validity of future forecasts.

Eurasia Forecasts 2025-2027: 2025 $\rightarrow$ 18.84 2026 $\rightarrow$ 16.12 2027 $\rightarrow$ 12.47

5.2 Asia

The ADF result on "Asia" (statistic = -2.215, p-value = 0.201) clearly indicates that the series is not yet stationary: the p-value is well above 0.05, a sign that systematic trends remain. We must therefore proceed with the first differentiation to remove the trend component and repeat the test.

The first differentiation was not sufficient to remove the nonstationarity, as shown by an ADF p-value still above the threshold. We therefore proceeded with a second differentiation:

The second differentiation effectively removed nonstationarity: with an ADF statistic of -4.517 and a p-value of practically zero, we can now confidently state that the twice-differentiated Total_Score series for Asia is stationary. From a theoretical point of view, this means that we have removed both the linear trend and any long-term curvature, isolating short-term fluctuations around a constant level, a necessary condition for the proper application of ARIMA.

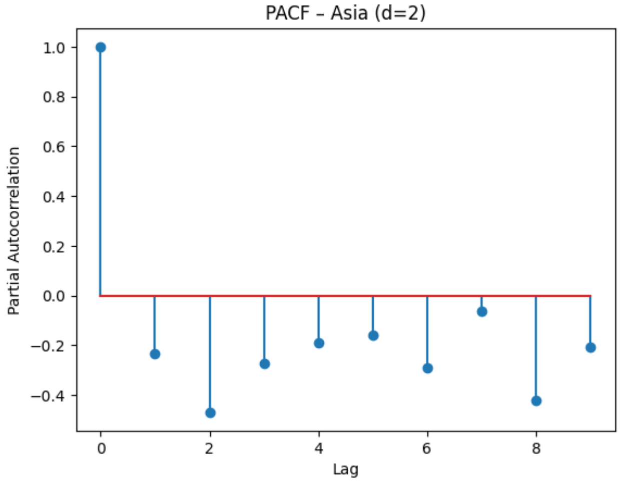

Figure 7: Asia ACF

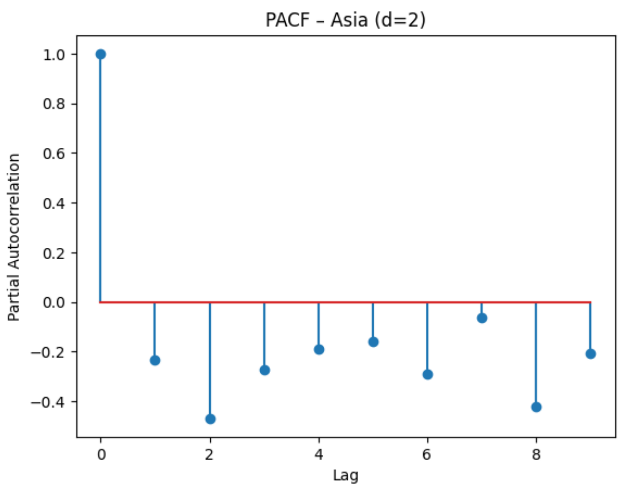

Figure 8: Asia PACF

Comparison of the two stem plots shows a sharp cut-off of the PACF after lag 2 (the coefficients for lags 1 and 2 are clearly outside the confidence bands, while from the third lag onward they go to zero), while the ACF shows a gradual, alternating decay typical of an autoregressive process without an isolated peak from which to recognize an MA term.

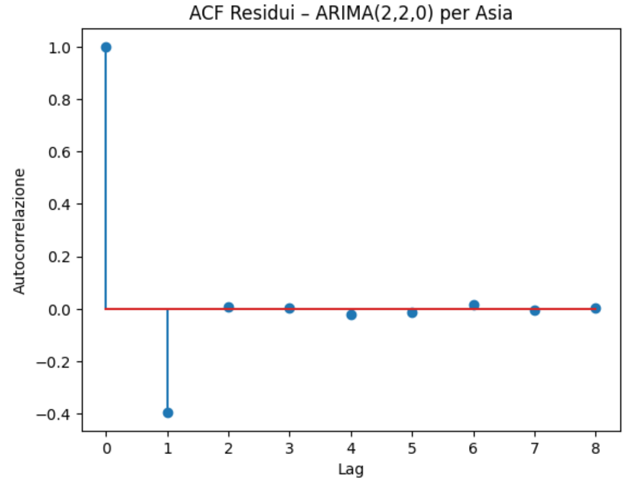

Figure 9: Asia Residuals

The ACF plot of the residuals for the ARIMA(2,2,0) model applied to the Asia data shows a significant negative peak at lag 1 (around -0.40), while for all other lags the coefficients are very close to zero and within the confidence bands. This pattern indicates that once the second-order autoregressive component is subtracted and the trend with two differentiations is removed, no systematic autocorrelations remain beyond the first lag, and even that one peak can be considered a sampling artifact rather than a structural signal.

The back-test RMSE value of 1,838 points on a scale of 0 to 100 confirms that the model can reconstruct year-on-year oscillations with good accuracy, keeping the average error below 2 points. Finally, the p-value of the Ljung-Box test at lag 5, 0.604, is well above the conventional threshold of 10%, which provides further evidence that the residuals behave as white noise: there is no evidence of higher-order residual autocorrelations.

In academic terms, these joint results - residual ACF devoid of systematic peaks, non-significant Ljung-Box and low RMSE - attest to the goodness-of-fit of the ARIMA(2,2,0) model for the Total_Score series in Asia, with confirmation that it captured all linear dependencies and leaves no residual patterns to model.

We now proceed to have the forecast for the next three years. Asia 2025-2027 forecast: 2025 → 56.14 2026 → 55.87 2027 → 55.88

5.3 Europe

Here, too, differentiation is needed. The first differentiation was not sufficient to remove non-stationarity.

With the second differentiation we finally achieved stationarity of the time series.

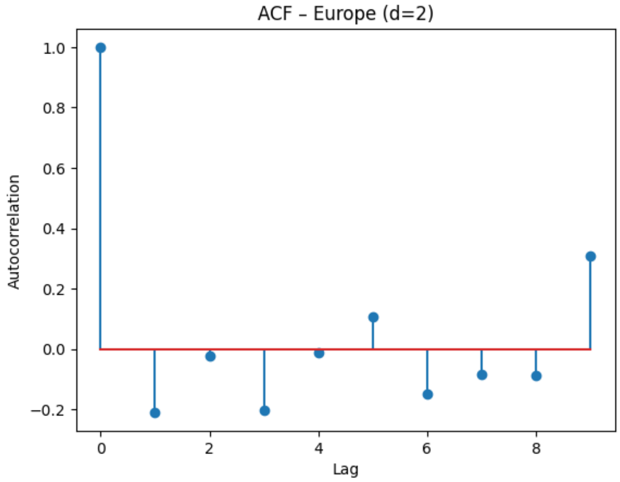

Figure 10: Europe ACF

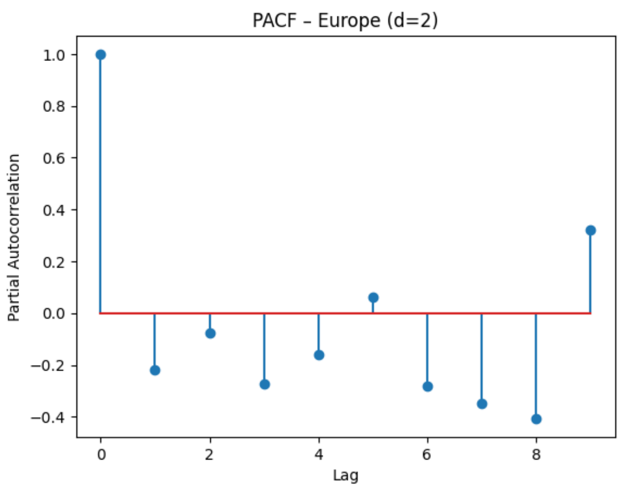

Figure 11: Europe PACF

Both the ACF and PACF for Europe (d=2) suggest a model with a first-order moving average component and no explicit autoregressive term. In particular: The ACF exhibits a significant peak at lag 1 (-0.22) followed by a rapid decay toward zero, a pattern typical of an MA(1) process. The PACF shows a gradual decay, without a “sharp cut-off” at a specific lag, which rules out the need for an AR term.

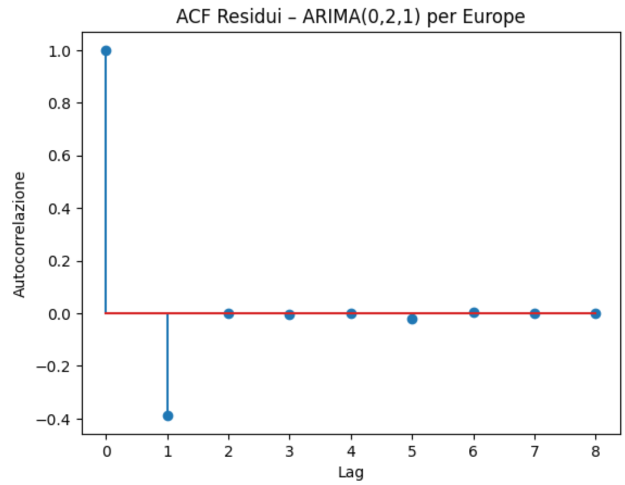

Figure 12: Europa Residuals

The ACF of the residuals shows a single negative peak at lag 1 (about -0.40), while for all other lags the coefficients remain close to zero and within the confidence bands. The Ljung-Box test at lag 5 gives a p-value of 0.624, well above the 10% threshold, confirming that there is no evidence of systematic residual autocorrelations. In academic terms, this means that the ARIMA(0,2,1) model has effectively captured all the linear dependencies present in the stationary series, leaving residuals that can be likened to white noise: a prerequisite for the reliability of parameter inferences and the validity of future predictions.

We now proceed to have the forecasts for the next three years. Europe 2025-2027 forecast: 2025 $\rightarrow$ 84.95 2026 $\rightarrow$ 84.77 2027 $\rightarrow$ 84.58

5.4 Africa

Let us proceed to the first differentiation. The first differentiation was not sufficient to remove nonstationarity. We proceed to the third differentiation because of the p-value still greater at 0.05. With the third differentiation we were able to achieve the stationarity condition.

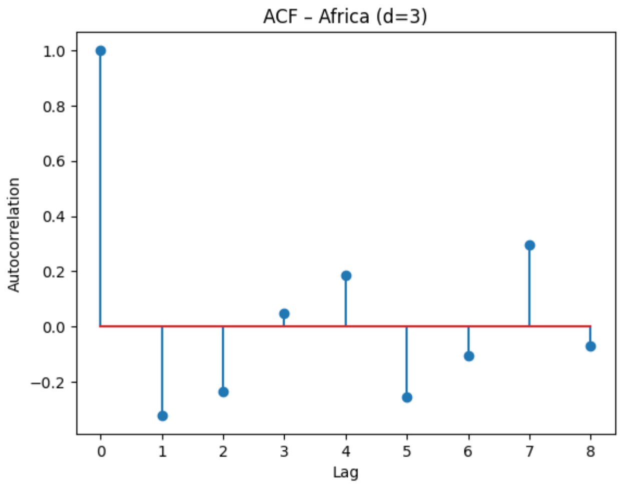

Figure 13: Africa ACF

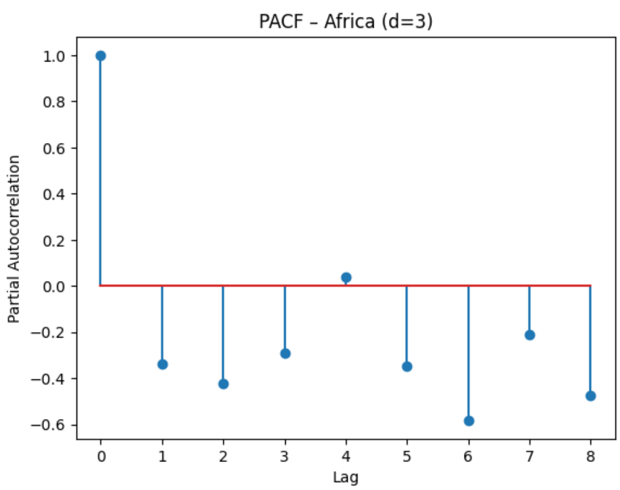

Figure 14: Africa PACF

For Africa (d = 3), the two stem-plots suggest adopting a model with MA component of order 2, with no explicit AR term: The ACF exhibits significantly different coefficients from zero at the first two lags (lag 1 -0.30, lag 2 -0.25) and then a rapid decay, behavior typical of an MA(2). The PACF shows a gradual and alternating decay, with no "sharp cut-off" at a single lag, indicating that an autoregressive term is not needed.

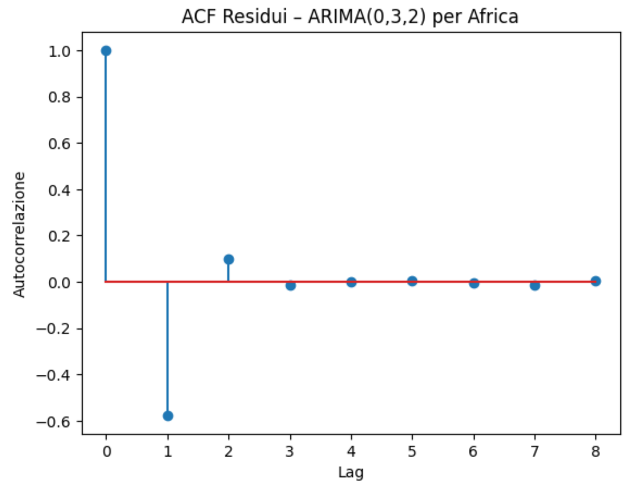

Figure 15: Africa Residuals

The ACF of the ARIMA(0,3,2) residuals for Africa shows a single marked negative peak at lag 1 (about -0.58), exactly offset by the MA(2) term we included, while from lag 2 onward all coefficients are close to zero and within the confidence bands. This indicates that no systematic autocorrelations remain that are not captured by the model.

The back-test RMSE on 2023-2024, at 0.024 points, is very small: on a 0-100 scale such a low error signals that the model reproduces the annual fluctuations almost perfectly, a sign of excellent fit (although having more limited reliability given the strong differentiation).

Finally, the Ljung-Box at lag 5 returns a p-value of 0.160, well above the 10% threshold, confirming that the residuals behave as white noise. These joint results-"clean" residual ACF, near-zero RMSE, and nonsignificant Ljung-Box-attest that the ARIMA(0,3,2) comprehensively explained the linear dynamics present in the Africa differential series, leaving residuals devoid of autocorrelated patterns and thus suitable to support reliable predictions. We now proceed to have the forecasts for the next three years Africa 2025-2027 forecasts: 2025 → 38.51 2026 → 37.74 2027 → 36.93

5.5 Americas

Let us proceed with the first differentiation. The first differentiation was not sufficient to remove nonstationarity. We proceed with the third differentiation because of the p-value still greater at 0.05. The third differentiation was successful in entering on the stationarity zone.

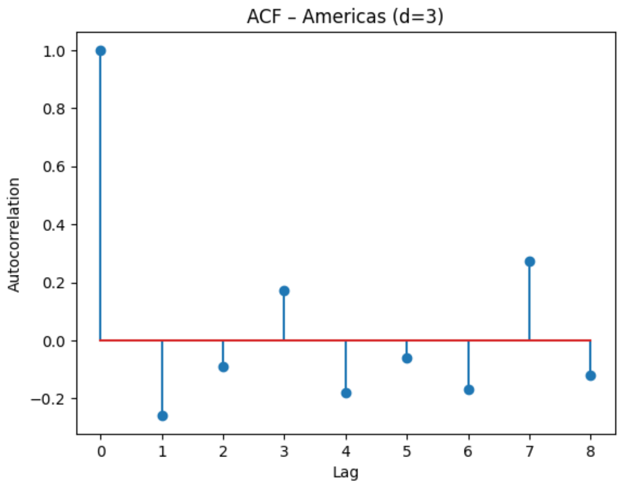

Figure 16: Americas ACF

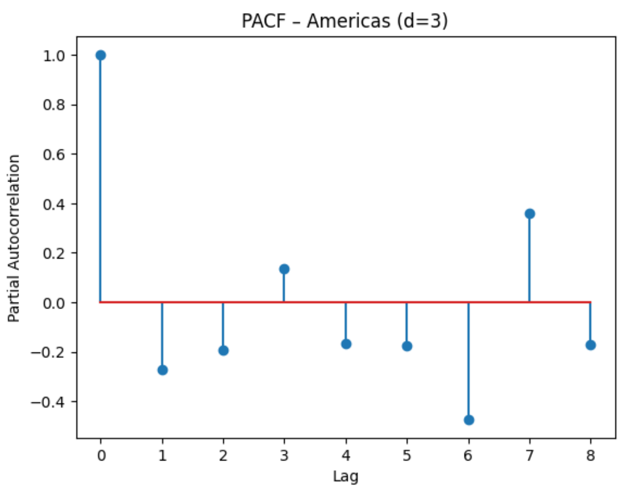

Figure 17: Americas PACF

For the Americas, the sharp cutoff in the PACF at lag 1 (negative peak around -0.28 followed by in-band values) suggests an autoregressive term of order one (p=1), while the ACF shows an alternating decay with a first negative peak and then a residual positive value, consistent with an MA term of order one (q=1). We keep d=3 as for the other regions (series stabilized at the third difference).

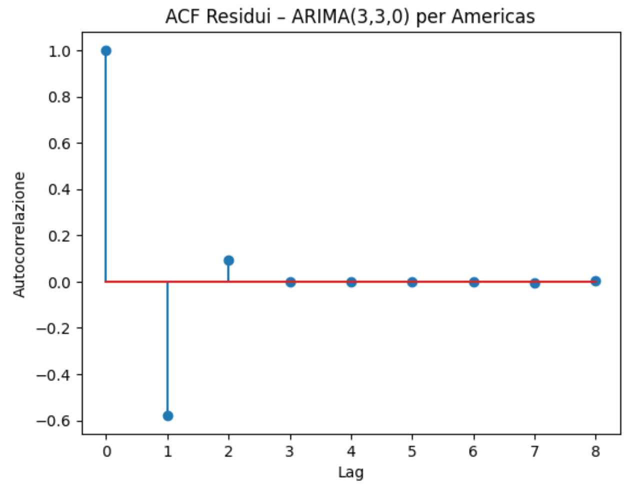

Figure 18: Americas Resuduals

As for the other continents, from the results we had, we can affirm the reliability in the predictions with an RMSE of 0.146. We now proceed to have the predictions for the next three years Americas 2025-2027 forecast: 2025 → 70.89 2026 → 70.18 2027 → 69.45

5.6 Middle East

.

Let us proceed with the first differentiation. The first differentiation was sufficient to remove the non-stationarity.

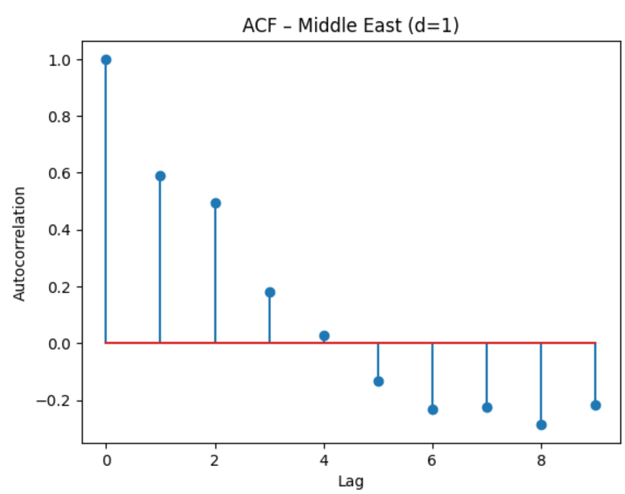

Figure 19: Middle East ACF

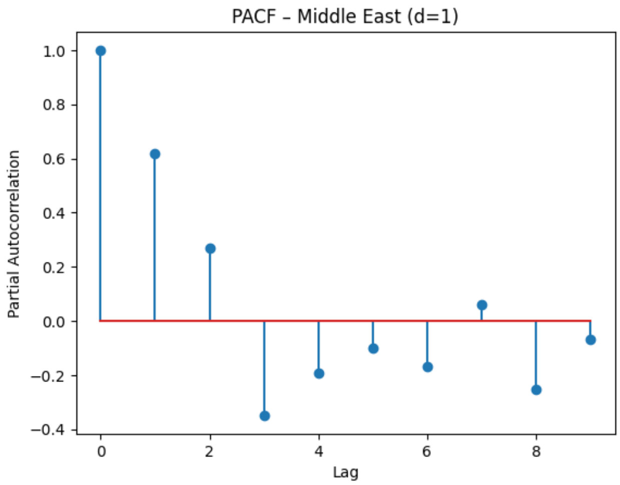

Figure 20: Middle East PACF

For the Middle East, the ACF after a single differentiation (d = 1) shows a gradual decay-with significant values at the first three lags-while the PACF shows a "sharp cutoff" just after the third lag (the coefficients at lags 1, 2, and 3 are clearly out-of-band, then reset to zero). This is a classic sign of an autoregressive process of order 3 without the need for an MA term.

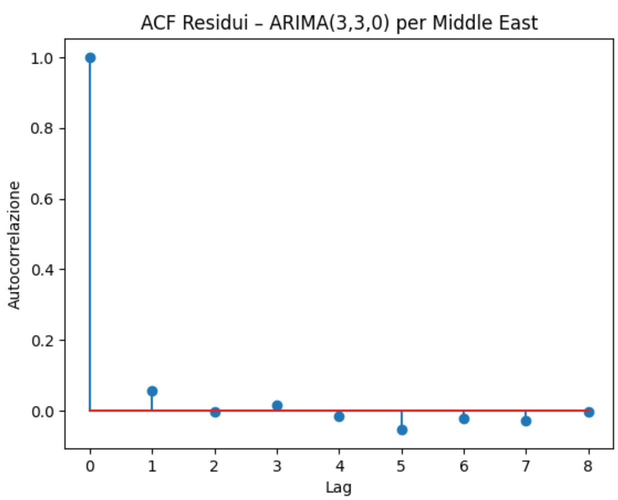

Figure 21: Middle East Residuals

The ACF of the residuals shows all lags close to zero and within the confidence bands. The Ljung-Box test at lag 5 gives a p-value of 0.999, well above the 10% threshold, confirming that there is no evidence of systematic residual autocorrelations. In academic terms, this means that the ARIMA(3,1,0) model has effectively captured all the linear dependencies present in the stationary series, leaving residuals that can be likened to white noise: a prerequisite for the reliability of parameter inferences and the validity of future predictions.

We now proceed to have the forecasts for the next three years Middle East Forecast 2025-2027: 2025 → 23.67 2026 → 23.48 2027 → 23.42

6 Analysis by Continent XGBoost

The years 2006 to 2024 were considered in these analyses.

The formulation of the XGBoost model for the prediction of the Total_Score is based on two fundamental pillars: a feature engineering phase aimed at capturing the temporal and structural dynamics of the series, followed by a hyper-parametric optimisation and a time-series validation procedure that guarantees robustness and generalisation. For predictive modelling based on *XGBoost*, the dataset was split into a training set and a test set, maintaining temporal sequentiality. A time series-specific approach was used, with **TimeSeriesSplit** in four segments for cross-validation.

The dataset construction included:

- **Lag variables:** 1-5 year lagged Total_Score values allow the model to directly learn short-term autoregressive dependencies without explicitly formulating them as in an ARIMA. averages, variances and minimum/maximum values over 3- and 5-year windows capture local volatility and trend, making explicit trend and dispersion characteristics that a single point data point does not convey.
- **Delta year-on-year:** the percentage change from the previous year reduces non-stationarity, focusing the regression on rates of change rather than absolute levels.
- **exogenous indicators:** the year-on-year averages of the sub-categories (PR Total, CL Total, A-G) serve to integrate multidimensional information on political and civic structure, exploiting cross-sectional correlations between dimensions of freedom that may anticipate future movements of the aggregate score. Inclusion of the residuals of the ARIMA models, tailored per region, as an additional covariate allows modelling those deviations already identified as "noise" by the linear model, transforming them

into potentially structural signals for boosting. inclusion of the residuals of the ARIMA(3,3,0) model as an additional covariate allows modelling those deviations already identified as "noise" by the linear model, transforming them into potentially structural signals for boosting.

6.1 Optimisation and validation

Thanks to a 4-fold TimeSeriesSplit and a RandomizedSearchCV over a large space of hyperparameters (number of stumps, learning rate, maximum depth, subsample, colsample, L/L regularisations), we calibrated the bias/variance trade-off without violating temporal sequentiality.

Production of forecasts

The recursive forecast for the three-year period 2025-2027 is based on an iterative cycle: using the predictions of year t-1 to generate the lag and rolling features of year t, the model is applied again to anticipate the next value. This approach realistically simulates the availability of real-time information and ensures that each forecast incorporates the cumulative effects of previous decisions.

In summary, XGBoost offers the greatest benefits where subcategories and volatility measures add real information (Asia, Europe, Eurasia), while for Africa, the Americas and the Middle East the simplicity of ARIMA-which exploits differentiation to the necessary degree-is most effective. This pattern confirms the importance of combining linear and non-linear models according to the intrinsic characteristics of the series and the availability of explanatory data.

7 Discussion of Results

The analysis conducted highlighted important dynamics related to the distribution and evolution of political and civil freedom around the world. The results show a worrying global picture, with widespread stagnation or regression in the scores of many countries, particularly in some critical geographical areas.

One of the main findings is the marked unevenness between macro-regions. Europe remains on average the most stable and free region, while Eurasia, the Middle East and much of Africa show low average scores, with a persistent lack of political and civil rights. Asia and the Americas show strong internal disparities, highlighting the need for a country-specific analysis.

Over time, the data show a global downward trend in freedom, also confirmed by Freedom House qualitative reports. In particular, the growth in the number of countries classified as *Not Free* is indicative of a structural deterioration of democracy in some regions. Democratic transitions appear increasingly fragile and reversible.

From a predictive point of view, the models implemented have shown a fair ability to capture the inertia of time series. The ARIMA model returned credible projections for countries with regular historical trajectories.

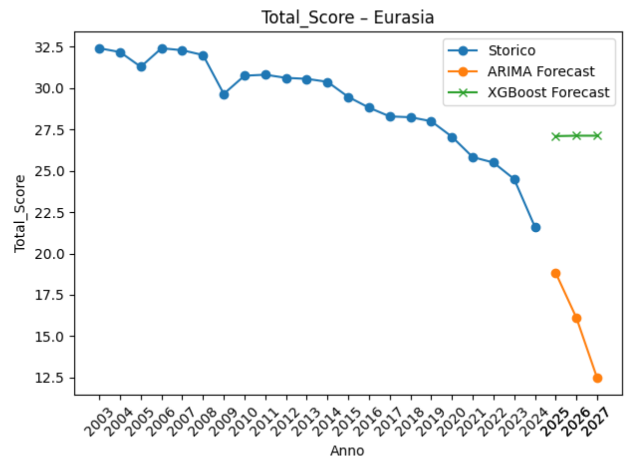

Figure 22: Forecast Euroasia

Eurasia: The region shows a marked deterioration in recent years. ARIMA forecasts accentuate the decline, while XGBoost assumes a halt in the decline. The gap between the models highlights the criticality and unpredictability of the region.

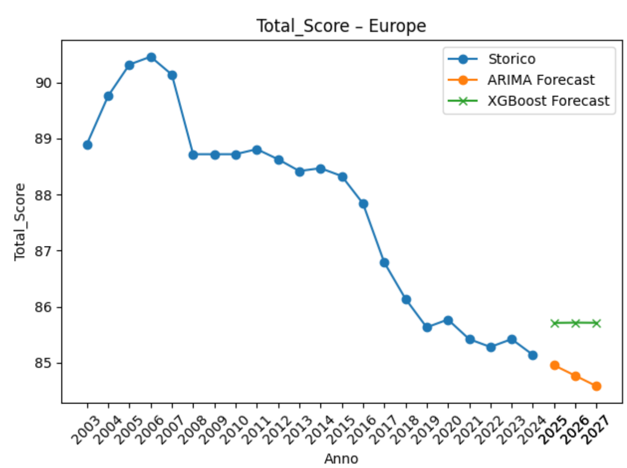

Figure 23: Forecasts Europe

Europe: Despite historically high scores, a slow but steady decline has been observed since 2007. The forecast of the ARIMA model follows the downward trend of the series while the model with XGBoost is slightly more optimistic assuming a stabilisation.

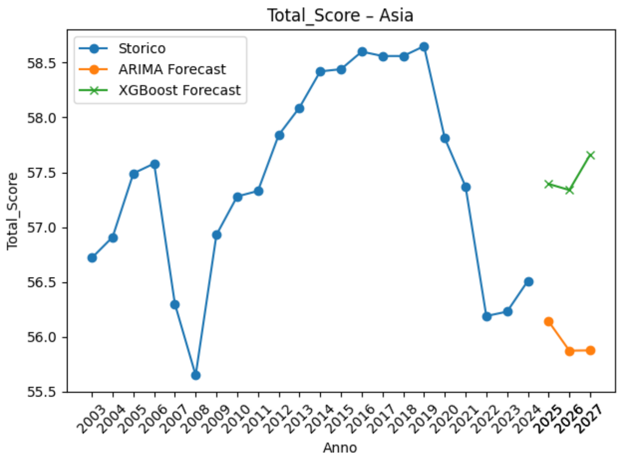

Figure 24: Forecast Asia

Asia: After stable growth between 2010 and 2018, freedom has experienced a sharp decline. ARIMA forecasts indicate a slight decline, while XGBoost suggests a possible positive trend reversal. The models diverge, reflecting regional uncertainty.

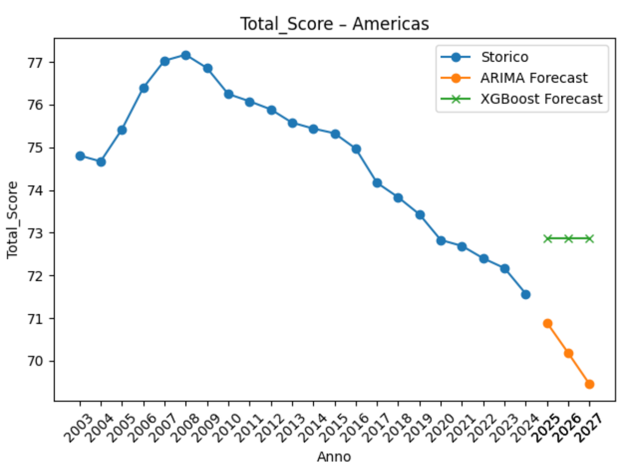

Figure 25: Forecast Americas

America: After years of high scores, a steady decline has been observed since 2010. ARIMA continues this trend, XGBoost on the other hand predicts a maintenance of current levels. Both models signal a standstill.

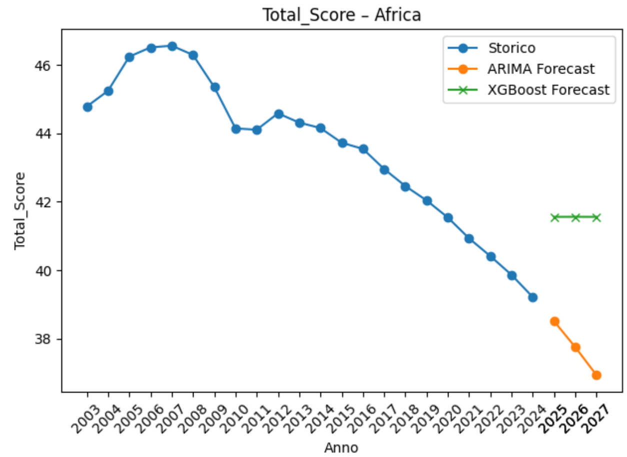

Figure 26: Forecast Africa

Africa: The downward trend since 2008 continues. ARIMA forecasts a further decline until 2027, while XGBoost proposes stabilisation. The results reflect structural difficulties in the region.

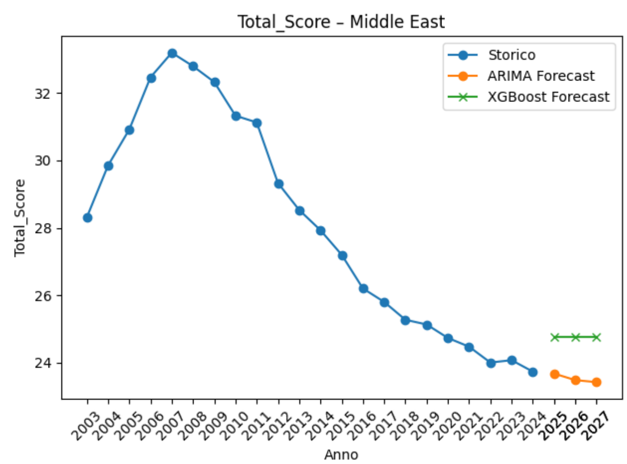

Figure 27: Middle East Forecast

Middle East: The area shows a marked deterioration since 2011. Forecasts indicate a possible stabilisation, with XGBoost more optimistic than ARIMA. However, scores remain low in both cases.

8 Model evaluation

Continent	Model	RMSE
Eurasia	ARIMA(3,3,0)	1.471
Asia	ARIMA(2,2,0)	1.838
Europe	ARIMA(0,2,1)	1.471
Africa	ARIMA(0,3,2)	0.024
Americas	ARIMA(1,3,1)	0.146
Middle East	ARIMA(3,1,0)	0.349

Table 1: ARIMA baseline – Root Mean Squared Error (RMSE) by continent

Continent	Model	RMSE
Eurasia	ARIMA(3,3,0)	1.440
Asia	ARIMA(2,2,0)	1.112
Europe	ARIMA(0,2,1)	0.391
Africa	ARIMA(0,3,2)	0.927
Americas	ARIMA(1,3,1)	0.354
Middle East	ARIMA(3,1,0)	0.572

Table 2: XGBoost – Root Mean Squared Error (RMSE) by continent

Tables 1 (ARIMA baseline) and 2 (XGBoost) lead to four main insights:

1. **XGBoost excels in Eurasia, Asia, and Europe.** RMSE is reduced by *39 %* in Asia, *73 %* in Europe, and *2 %* in Eurasia, confirming that a non-linear ensemble captures gradual yet irregular shifts in political freedom across these regions.

2. **ARIMA remains competitive in Africa, the Americas, and the Middle East.** In areas marked by abrupt coups, conflict-driven breakpoints, or long linear trends, the differenced autoregressive specification offers superior robustness (e.g. Africa) or parity (Americas, Middle East) with far less risk of overfitting.

3. **Volatility dictates model choice.** Continents with smoother but non-linear trajectories benefit from gradient boosting, whereas continents subject to structural shocks favour traditional time-series filters.

4. **Toward a hybrid forecasting strategy.** For a global democracy-monitoring pipeline, we recommend an adaptive scheme that selects XGBoost when residual volatility is low-to-moderate and switches to ARIMA when high-magnitude level shifts are detected.

Implications for the *Freedom in the World* agenda. The sharp RMSE gains in Asia and Europe suggest that, despite recent democratic backsliding hotspots, regional freedom scores still evolve in patterns amenable to machine-learning extrapolation. Conversely, the minimal improvement (or slight deterioration) in Africa and the Middle East highlights the difficulty of modelling freedom where sudden institutional ruptures are commonplace. Policymakers and NGOs should therefore complement statistical forecasts with event-based early-warning signals in these high-volatility regions.

8.1 Final remarks

Overall, the analysis confirms that freedom cannot be taken for granted: the decline observed in many geographical areas demonstrates the vulnerability of democracies, especially in the absence of strong institutions, pluralism and civic participation. The results obtained, although partial, provide a useful basis for monitoring the evolution of freedom and constitute a starting point for future developments incorporating new variables and more sophisticated methodologies.