

7.4

Tool for robust forecasting

This study presents a forecasting equation search tool to support macroeconomic forecasting. The tool provides a systematic procedure for generating, evaluating and selecting forecasting equations. One of its key features is the use of historical forecasting accuracy to guide the search for new equations and to identify equations that are stable over time. It also integrates and automatizes various forecasting steps – such as data cleaning, stationarity testing, stability checks, and economic sign restrictions – into a single framework. The tool generates a large set of candidate equations, which are then filtered and ranked based on stability and forecasting performance. The final selection remains the responsibility of the forecaster. The tool thus facilitates fast construction of stable and interpretable short-term forecasting equations. It is intended both to assist less experienced forecasters and to accelerate the workflow of experienced economists when fast re-specification is required.

1. Introduction

Macroeconomic forecasting at STATEC relies on a large number of relatively simple econometric equations. These equations, used to forecast inflation, economic growth, unemployment, and related variables, must be estimated, monitored and regularly updated in a context where data are volatile and frequently revised between the first and second data releases¹. In such an environment, ensuring the stability and robustness of forecasting equations represents a challenge. This study² presents a tool designed to address these challenges.

2. Motivation

Macroeconomic forecasts are typically based on systems of equations. Each forecasting round involves updating a system of 30 to 100 equations and identities to project a few months, quarters or years ahead.

The use of relatively simple equations is motivated by several operational requirements: equations must be transparent, quickly updateable, and economically interpretable to support communication and policy analysis.

Luxembourg's economy presents however particular forecasting challenges. The small and highly open economy is strongly exposed to external developments, especially in financial services. Data revisions are frequent and sometimes substantial: absolute revisions³ to quarterly real GDP growth have averaged around 0.6 percentage points over 2006-2024, which is larger than the average quarterly growth rate of about 0.5%. In this environment, forecasting equations may become unstable following revisions, coefficients may lose statistical significance, and forecast errors may increase unexpectedly.

This often requires re-specification, which is both time-consuming and difficult to systematize. Moreover, in situations such as during economic shocks there is a need to quickly construct alternative equation specifications.

¹ For a recent example of a simple forecasting equation, see NDC 01-2025, "De nouvelles équations de prévision de l'inflation trimestrielle au Luxembourg".

² This is the first part of a two-part study. The present part describes the new procedure for the forecasting equation selection tool, while the second part, scheduled for NDC 2-26, will showcase its concrete application.

³ Absolute revisions refer to the absolute difference between the first and second releases of real GDP.

The forecasting equation search tool is designed to respond to these challenges by providing:

- A systematic procedure to generate a large set of candidate equations;
- A framework for evaluating their robustness using pseudo-forecasting techniques;⁴
- A filtering process based on econometric, statistical and economic criteria;
- A user-friendly interface that can be used by both experienced and less experienced forecasters.

Furthermore, compared with other complex equation-search procedures (such as LASSO, GETS, LARS, RIDGE, ElasticNet)⁵, this procedure is specifically designed for forecasting. Instead of relying on how well the model matches past data, as the aforementioned procedures do, it also uses historical forecasting performance to guide the search for new forecasting equations, which helps find equations that are more stable over time. It also allows existing equations to be revised and updated quickly if their performance begins to deteriorate. Finally, unlike complex procedures that return only a single forecasting equation, it generates a broad set of candidate equations, allowing the researcher to choose an alternative if the initial specification does not make economic sense.

3. General Structure of the Procedure

The procedure consists of three main stages:

1. Data preparation, including cleaning, seasonal adjustment and transformation;
2. Equation set creation, involving the generation and estimation of a large number of candidate equations;
3. Equation selection, based on a sequence of filtering criteria and performance metrics.

This structure ensures that the final selected equations are both statistically sound and operationally useful.

3.1 Data Preparation

Before the data can be analyzed and used to produce forecasts, it must first be adjusted. This involves removing seasonal patterns, treating outliers that could impair forecast accuracy and applying stationarity transformations.

3.1.1 Data adjustment and statistical data preparation

The data preparation process begins with the import of raw data and the elimination of variables that are unsuitable for econometric analysis: non-numeric variables, variables with little or no variation, and those containing exclusively missing observations are removed. We also remove variables with too many missing values at the end or within the sample, as well as variables with a short data history, since these are likely to be poor predictors. Duplicate or highly similar series are also identified and discarded. These steps are implemented through automated routines complemented by visual checks, ensuring a consistent and reliable dataset prior to statistical processing. The remaining series undergo statistical preparation, including seasonal adjustment, outlier treatment and stationarity transformations.

a) Seasonal adjustment

The first step of statistical preparation of the dataset involves seasonal adjustment. Seasonal patterns are removed from the data, as they represent repetitive and predictable within-year fluctuations that can be reintroduced into the forecasts if necessary. Eliminating them at an early stage allows focussing solely on the trend and cyclical components, which typically improves analysis and forecasting performance. Seasonal adjustment is performed using the X-13ARIMA-SEATS⁶ method implemented in the R package seas, with diagnostic plots reviewed visually to ensure accuracy. Figure A shows an example of automatic seasonal pattern detection for four time series: in the first three, no seasonal component is identified, as indicated by the overlap of the raw and adjusted series, whereas in the fourth one a seasonal pattern is detected and removed, leading to a visible divergence between the original and seasonally adjusted data.

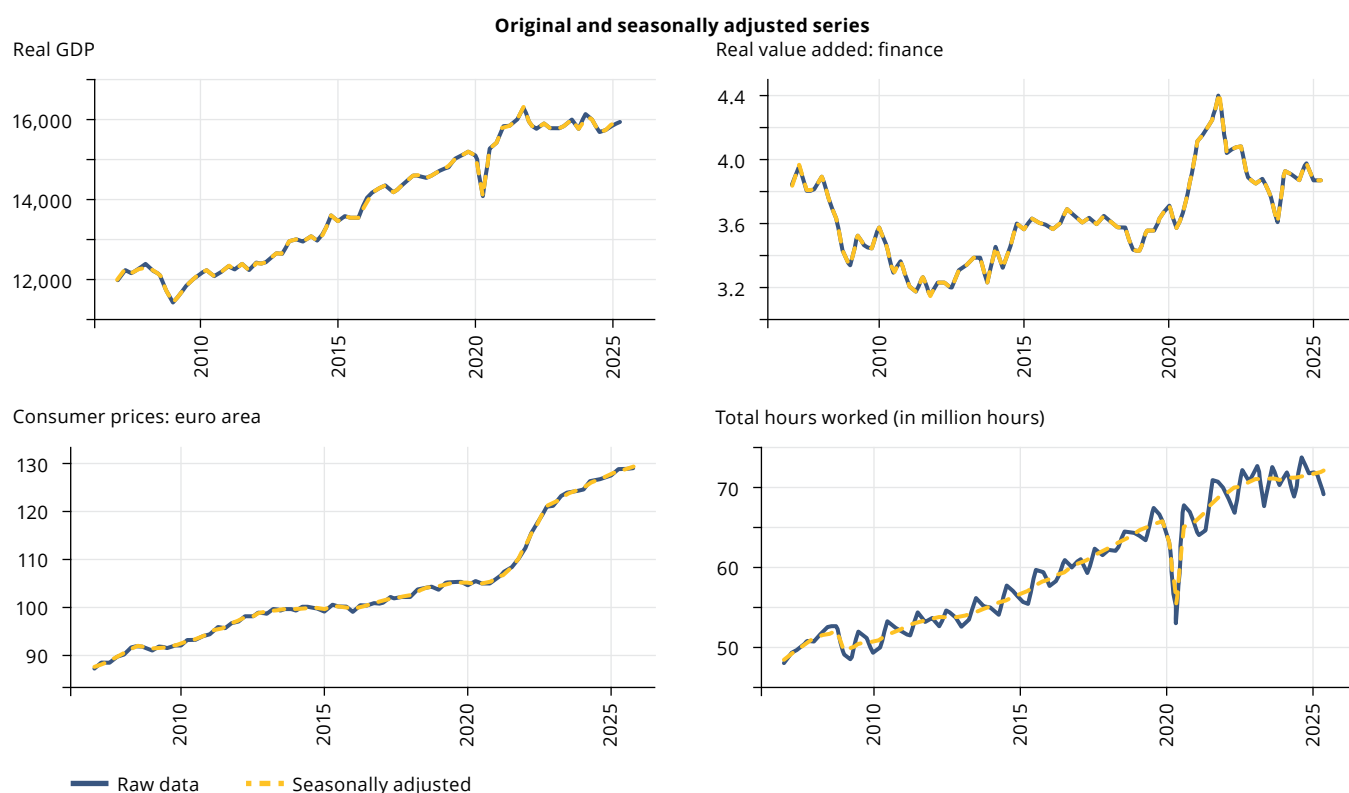
⁴ In pseudo-forecasting techniques, we simulate forecasting conditions using historical data. Put simply, we use data up to a given point in time, estimate the equation as if future data were unknown, predict the next observation and compare it with the realized value.

⁵ These methods are regularized or "algorithmic" variable-selection techniques and are described in more detail in Hastie et al. (2009).

⁶ X-13 ARIMA-SEATS is a statistical procedure developed by the U.S. Census Bureau for seasonal adjustment and time-series decomposition. It uses ARIMA modeling to extend the series and the SEATS method to separate observed data into trend-cycle, seasonal, and irregular components.

Figure A

Example of a deseasonalization procedure



Source: STATEC

Note: Only total hours worked contain seasonal effects.

b) Stationarity transformations

Following seasonal adjustment, the series are tested for stationarity. In simple terms, a stationary time series does not exhibit persistent trends,⁷ whereas non-stationary data typically display evolving levels over time. When trends are present, it is standard practice to transform the data – most commonly into growth rates or first differences, which transform the data to achieve stationarity. Ensuring stationarity is essential because the validity of econometric inference, including equation diagnostics, hypothesis testing and confidence intervals for forecasts, relies on the assumption that the data entering the equation possess stable statistical properties.

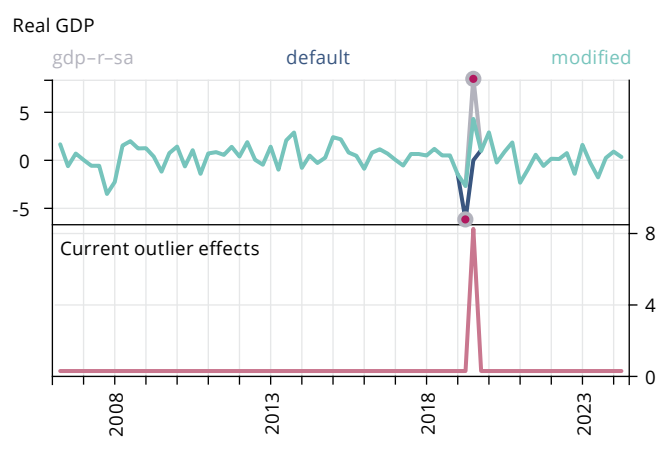
To this end, the stationarity testing procedure has been fully automated. Its practical objective is to determine whether a series can be modelled in levels or requires transformation, such as differencing or logarithmic differencing, to stabilize its mean and variance. The assessment is based on Augmented Dickey-Fuller (ADF) tests conducted according to a sequential methodology. This procedure involves iterating between two types of tests: the first identifies the appropriate deterministic components – such as a constant or a time trend – while the second evaluates whether the series is stationary. A well-known challenge arises because the specification of deterministic components depends on the stationarity of the series, while the stationarity test itself requires the correct specification of these components. To resolve this circularity, a sequential testing strategy, as described in Enders (2015), is applied and implemented programmatically. This approach ensures a consistent, transparent and reproducible determination of the appropriate transformation for each time series.

⁷ More precisely, a weakly stationary data exhibits time-invariant mean and variance and an autocorrelation function that does not change over time but only depends on the lag.

c) Outlier treatment

Outliers (extreme observations) can strongly affect estimation, equation validation and forecast accuracy.⁸ Figure B displays Luxembourg's real GDP growth during the COVID-19 period, when the collapse and rebound were so large that an equation fitted on raw data would risk over-reacting to this unusual episode and performing poorly in more typical periods (see the two red circles in Figure 2, which indicate collapse and rebound in real GDP growth associated with the COVID-19 crisis).

Figure B
Example of attenuation of outliers



Source: STATEC

Note: The upper panel of the figure shows the quarterly growth rate of Luxembourg's real GDP (grey-blue line). The lower panel displays the detected outlier periods. The green line in the upper panel shows the adjusted real GDP growth rate, with outliers attenuated.

Rather than deleting such observations, we reduce their values. This approach recognizes that extreme events may contain useful information about the economy, but prevents their magnitude from dominating estimation. As a result, the green line in Figure B represents Luxembourg's real GDP growth rate with COVID-19 collapse and rebound attenuated. Attenuation offers a practical compromise: it reduces the disproportionate influence of rare episodes while retaining part of their informational content.

3.1.2 Output

The output of the data preparation stage is a dataset cleaned in three respects: problematic variables are removed, retained series are seasonally adjusted and transformed when needed to achieve approximate stationarity, and influential outliers are identified and attenuated. These processed data are then passed to the next stage, where candidate forecasting equations are generated.

3.2 Equation Set Creation

The part of the procedure that generates a set of candidate forecasting equations is designed to handle two practical cases:

1. A large pool of potential predictors, and
2. A smaller, pre-selected set of candidate predictors.

Because the second case is nested within the first, we focus here on the more general and more demanding case in which the forecaster begins with a large number of potential predictors (for example, 200 predictors or 800 if up to 4 lags are included).

The central problem is that many predictors may be plausible, but only a few can be included in any one equation, and the best combination is unknown ex ante. The objective is therefore not to identify a single preferred equation immediately, but to generate a candidate set (for example, 200,000 potential equations) that is sufficiently broad, forecast-oriented, and interpretable to support subsequent selection.

A useful equation-generation procedure should satisfy three requirements:

1. **Forecasting power:** the equations should generate accurate predictions;
2. **In-sample adequacy:** the equations should provide a reasonable description of historical data;
3. **Diversity:** the procedure should explore a broad range of predictor combinations rather than concentrating too heavily on a narrow subset of variables and should return a set of equations rather than a single "best" equation.

The third requirement is particularly important. Automatic selection procedures, such as those mentioned in Section 2, typically return a single equation chosen solely on statistical criteria. However, if the variables included in the single equation cannot be economically justified,⁹ the resulting equation is not credible and, in effect, no usable equation is obtained. Such procedures overlook alternative equations that are economically meaningful and perform well in forecasting. By contrast, this approach returns a set of equations, allowing us to choose among alternative equations that forecast equally well.

⁸ Outliers are identified using an iterative ARIMA-based procedure in the spirit of Chen and Liu (1993).

⁹ If a variable enters the model with an incorrect sign, for example, if an increase in prices leads to an increase in real consumption.

An exhaustive search over all admissible predictor combinations is not feasible (for example, with 800 potential predictors and a maximum of 5 predictors per equation the total number of equations is 2.7 trillion). The procedure therefore proceeds in two steps. First, the predictor set is reduced to a manageable subset. Second, candidate equations are generated from this reduced set through a stage-wise search that combines in-sample and pseudo-out-of-sample criteria.

a) Step 1: Predictor pre-selection

The first step reduces the initial set of candidate predictors to a smaller pool that is sufficiently rich to contain promising forecasting information, but small enough to permit systematic equation search. The first step is hence a screening device, not a final selection rule.

To that end predictors are retained using several complementary procedures: simple correlation screening, partial correlation screening conditional on an autoregressive benchmark, and algorithmic methods such as Least Absolute Shrinkage and Selection Operator (Lasso), least angle regression, and stepwise selection. These procedures are combined because no single method is sufficiently reliable on its own. Correlation-based measures are transparent and fast to compute but do not produce a complete equation, whereas algorithmic procedures typically return only one preferred specification which includes predictors that can be difficult to justify economically or can be unstable over time.

It should be noted that these pre-selection procedures rely on full-sample criteria and may therefore favor equations that fit a particular sample well but will generalize poorly and will produce less accurate forecasts as data are updated and revised.

b) Step 2: Stage-wise equation search

After pre-selection, candidate equations are generated by a stage-wise search. The objective is to obtain equations that combine good historical forecasting performance, adequate in-sample fit, and diversity in predictor composition. Forecasting performance matters because the equations are intended for operational use. In-sample adequacy is relevant because equations that fit historical data plausibly are easier to interpret and less likely to reflect chance correlations. This is especially important in large-scale searches, where some equations will inevitably perform well over a short evaluation window purely by chance, even though they are economically implausible or unstable.¹⁰ Requiring both forecasting performance and in-sample adequacy reduces this risk.

The search begins with one or two exhaustive estimation stages in which the algorithm estimates all possible equations containing a small number of predictors (e.g. up to two predictors). We call these stages initial estimation stages. Exhaustive estimation in initial stages is deliberate. It includes equations with predictors that perform poorly initially or on their own, but that may still improve an equation once combined with another predictor in later stages, as described in the next paragraph.

Beyond the initial stages, exhaustive estimation becomes computationally infeasible because the number of possible predictor combinations increases exponentially. For this reason, the procedure switches to guided expansion. At each guided expansion stage, equations from the previous stage are ranked by a composite suitability criterion that combines in-sample fit with pseudo-out-of-sample forecasting accuracy. This is the key contribution of the procedure. Because it uses pseudo-out-of-sample performance to guide equation search, it selects equations that are more stable over time and more relevant for actual forecasting.¹¹

At each guided expansion stage, only a pre-defined fixed share of the best-performing equations is retained for further expansion (e.g. 50% of equations will be expanded). For each retained equation, predictors not yet included are evaluated through their correlation with the equation residuals. The residuals represent the unexplained part of the forecasted variable, and a new predictor should help predict this component. Predictors with a strong correlation with the residuals are therefore natural candidates for inclusion in the next guided expansion stage.

Guided expansions end when the maximum allowed number of predictors per forecasting equation is reached, leaving a large set of candidate forecasting equations that must be filtered and selected, as described in the next section.

¹⁰ For example, from a set of 200,000 equations, it is easy to find “nonsensical” equations that nevertheless predict the last 20 data points (equivalent to five years of quarterly data) very well.

¹¹ That is, the procedure ensures robust performance not only on the full sample but also on sub-samples or when new data are added.

Box**Summary of the selection procedure****Define:**

- Maximum number of predictors per equation (for example, $(k = 5)$).
- Maximum equation size for exhaustive estimation (for example, maximum two predictors per equation).
- Pseudo-out-of-sample evaluation window (for example, the last 40% of the sample).

Stages 1–2: exhaustive search over simple equations

- Estimate all possible equations using all combinations of up to i predictors (for example, $i = 2 < k$).
- For each equation, compute:
 - full-sample Akaike Information Criterion (AIC)¹², and
 - pseudo-out-of-sample Root Mean Squared Forecast Error (RMSFE)

Later stages: guided expansion

- Rank equations from the previous stage using a composite criterion based on AIC and RMSFE.
- Retain only a chosen share of the best-performing equations (for example, the best 50%).
- For each retained equation, compute the correlation between its residuals and predictors not yet included in that equation.
- Expand equations by adding predictors with the highest residual correlation, subject to a cap on the total number of equations estimated in the new stage.
- Optionally exclude equations with high multicollinearity or other undesirable properties.

Outcome

- A large and diverse set of candidate forecasting equations that balances forecasting performance, in-sample adequacy, and computational feasibility.

Overall, the proposed procedure is intended to strike a balance between broad coverage, good forecasting performance, and practical feasibility.

3.3 Equation Filtering and Selection

The equation search procedure generates a very large number of candidate forecasting equations, often reaching several tens or even hundreds of thousands of specifications. While this exhaustive exploration is useful to ensure that potentially relevant equations are not overlooked, it also implies that a structured and rigorous selection process is required to identify a manageable and reliable subset of equations.

The objective of the equation filtering and selection stage is therefore to progressively reduce the initial set of candidate equations by applying a sequence of complementary filters. Overall, the equation filtering and selecting stage transforms a large and heterogeneous set of candidate equations into a small number of robust, interpretable and operational forecasting equations. It constitutes a central component of the procedure, linking the automated equation search to practical forecasting needs.

a) Econometric, Stability and Economic (ESE) Filtering Strategy

The ESE process follows a sequential approach. Starting from the full set of estimated equations, successive criteria are applied to eliminate inadequate specifications. Each successive step reduces the set of remaining equations while increasing the average quality and forecasting suitability of the equations that remain.

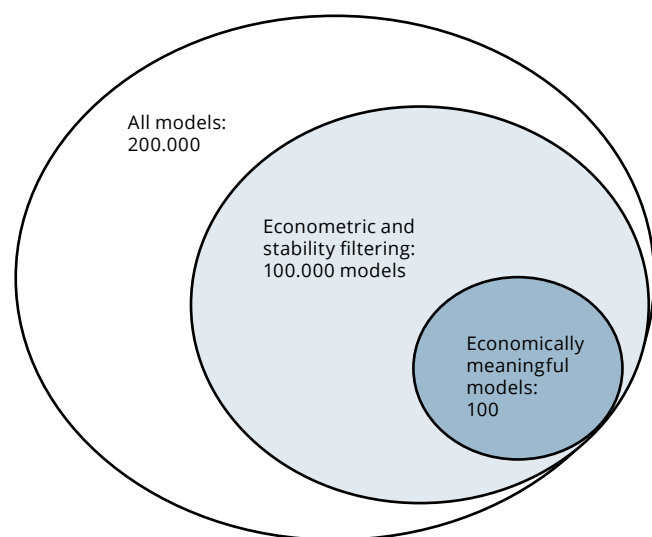
¹² The Akaike Information Criterion (AIC) is one of the most widely used model selection criteria in time series analysis, balancing how well the model fits the data against model complexity. The Root Mean Squared Forecast Error (RMSFE) is a summary measure of a model's forecast accuracy.

In practice, filtering typically reduces the initial set of equations to a subset ranging from around 100 to 2,000 equations. This subset is sufficiently small to allow for detailed inspection, while still preserving a diversity of specifications.

The ESE filtering is organized around three main pillars (see Figure C):

- 1) Econometric adequacy,
- 2) Parameter stability,
- 3) Economic consistency.

Figure C
Stylized diagram of filtering steps



Source: STATEC

Econometric Filtering

The first stage of the selection process focuses on standard econometric diagnostics. The aim is to ensure that the retained equations satisfy basic statistical requirements.

A key criterion concerns the statistical significance of coefficients. Equations in which explanatory variables are not statistically significant at conventional levels (typically 5%) are discarded. This requirement helps to avoid over-parameterized specifications and ensures that retained regressors contribute meaningfully to the explanation of the dependent variable.

In addition, residual diagnostics play an important role. In particular, equations exhibiting autocorrelation in the residuals are excluded, as this indicates that part of the dynamics of the dependent variable is not properly captured. Standard tests (e.g. autocorrelation tests over several lags) are used for this purpose. Other diagnostic checks, such as heteroscedasticity or normality of residuals, may also be considered. While these are not always binding constraints, they provide useful information on the reliability of inference and the overall quality of the equation.

Stability Filtering

The second stage focuses on the stability of equation parameters over time, which is a crucial aspect in a real-time forecasting context. In practice, even equations that perform well in-sample may prove unreliable if their coefficients are sensitive to changes in the estimation sample. Stability filtering is hence a key element of the procedure, as it directly addresses one of the main motivations of the tool, namely the sensitivity of forecasting equations to data revisions.

To address this, the procedure relies on recursive estimation techniques. Equation coefficients are re-estimated over expanding or rolling samples, and their statistical significance is monitored over time. A typical criterion requires that coefficients remain significant (e.g. at the 10% level) over a sufficiently large and recent portion of the sample, such as the last 30%.

Equations that exhibit strong parameter instability – such as coefficients that frequently change sign or lose significance – are excluded. This step ensures that the selected equations are robust to small changes in the data and are therefore more likely to perform well in real-time forecasting exercises.

Economic Filtering

The third and last stage introduces economic criteria into the selection process. While econometric diagnostics are essential, they are not sufficient to guarantee a reasonable economic narrative of an equation.

The procedure therefore allows the user to impose theoretical or narrative restrictions on the equations. These restrictions typically concern the expected sign or the presence (or absence) of certain variables. For example, an increase in interest rates would generally be expected to have a negative effect on future economic activity, while an increase in external demand should have a positive effect on current activity. Equations that violate such restrictions can be excluded from the candidate set.

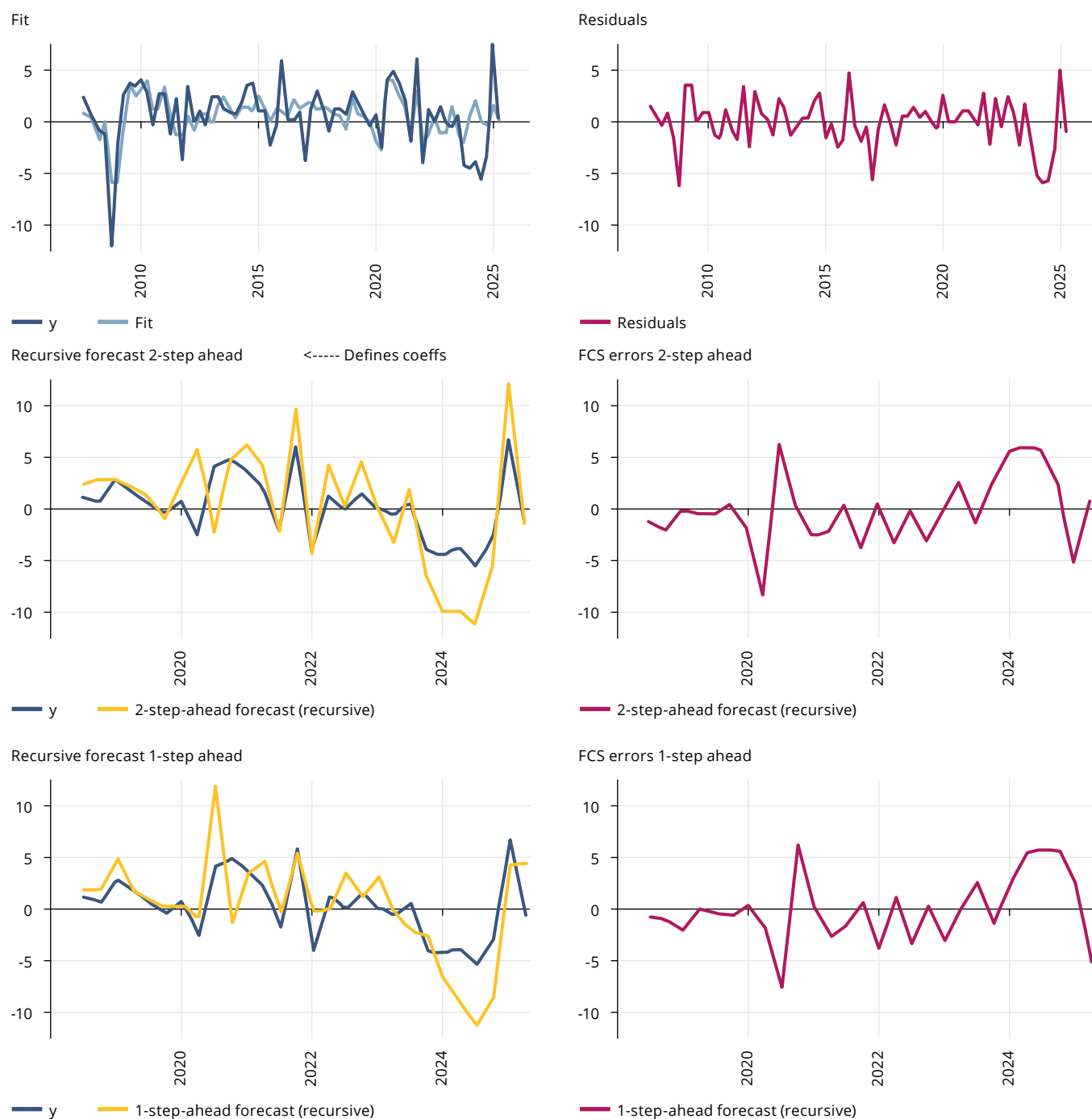
Figure D
Illustration of equation-selection objects



Source: STATEC

Figure E

Illustration of “real-time” forecast accuracy for Luxembourg real export growth



Source: STATEC

This step ensures that the retained specifications are consistent with well-established economic relationships and can be more easily interpreted and communicated.

At the same time, the filtering remains flexible: the user can decide whether to impose strict restrictions or to allow for a broader set of equations, depending on the context and the purpose of the analysis.

b) Final Ranking and Selection

After applying the successive filters, the remaining set of equations is typically reduced to a manageable number of (say) one hundred equations. These equations are then ranked according to a combination of performance indicators. A central criterion is the RMSFE obtained from pseudo-forecasting exercises. This measure provides an indication of the equation's predictive accuracy in a setting that mimics real-time forecasting.

In addition, in-sample fit measures, such as the information criterion or the coefficient of determination, may be considered. Particular attention can also be given to recent forecast performance, for example over the last two years, in order to capture the equation's ability to adapt to the most recent economic conditions.

The final step involves a manual review of the best performing equations, typically focusing on a small subset (e.g. the top 10 specifications). Detailed reports are available for each equation (illustrated in [Figure D](#)), including coefficient estimates, diagnostic tests, stability indicators and measures of variable contributions.

The detailed report also includes figures illustrating the equation's historical forecasting performance, as shown in the bottom two panels of [Figure E](#). These panels correspond to one-quarter-ahead and two-quarters-ahead real-time forecasts. In each panel, the dark blue line represents the observed values of Luxembourg's real exports, while the yellow lines show the corresponding real-time one- and two-quarter-ahead forecasts. These figures provide a visual assessment of forecast reliability and accuracy, which will be discussed in more detail in the next Note de conjoncture (NDC 2-2026), to be released in December of this year.

The final equation selection is therefore not purely mechanical but combines quantitative criteria with expert judgement. This ensures that the chosen specification is not only statistically robust but also economically meaningful and suitable for communication purposes.

Greater emphasis on the results of the equation selection, along with a more comprehensive and detailed application of the proposed tool, will be provided in a follow-up study. This second part will extend the present analysis by providing deeper empirical insights and practical implementation aspects, thereby highlighting the robustness and usability of the approach. It is scheduled for publication in the next Note de conjoncture (NDC 2-2026), to be released in December of this year.

Conclusion

This study presents a systematic and operational tool for the selection of robust short-term forecasting equations, developed to support macroeconomic forecasting activities at STATEC. In a context characterized by frequent data revisions, continuous structural changes and the need for timely and transparent forecasts, the proposed procedure offers a structured and reproducible approach to equation development. By integrating automated data cleaning, rigorous statistical preparation, large-scale generation of equations and a comprehensive filtering framework, the tool enhances the efficiency, reliability and consistency of the forecasting process.

The methodology combines established econometric best practices – such as seasonal adjustment, stationarity testing, outlier treatment and pseudo-out-of-sample evaluation – within a coherent and user-friendly framework. The stage-wise search procedure generates a diverse set of candidate equations, while the sequential filtering process ensures that the final equations meet strict econometric standards, demonstrate parameter stability and remain consistent with economic theory. Importantly, the approach preserves the central role of expert judgement, enabling economists to select equations that are not only statistically sound but also economically interpretable and suitable for communication and policy analysis.

Beyond improving forecast accuracy and robustness, the tool significantly reduces the need for manual re-specification and facilitates the rapid construction of alternative equations in response to economic shocks or data revisions. It thereby represents a valuable complement to the existing forecasting toolkit, strengthening both the transparency and the resilience of short-term macroeconomic analysis.

Future developments will focus on expanding diagnostic capabilities, refining outlier detection procedures, enhancing stability assessments and deepening the economic interpretation of equation results. As these enhancements are implemented and validated through continued practical application, the forecasting equation search tool is expected to play an increasingly important role in supporting evidence-based economic analysis and decision-making.

Overall, the procedure provides a pragmatic and efficient bridge between econometric rigour and operational forecasting needs, contributing to more robust, transparent and timely macroeconomic forecasts for Luxembourg.

References

Chen, C., & Liu, L.-M. (1993). Joint estimation of equation parameters and outlier effects in time series. *Journal of the American Statistical Association*, 88(421), 284–297.

Cochrane, J. H. (1991). A critique of the application of unit root tests. *Journal of Economic Dynamics and Control*, 15(2), 275–284.

Enders, W., 2015. *Applied econometric time series* fourth edition. New York (US): University of Alabama, 30.

Hastie, T., Tibshirani, R. and Friedman, J., 2009. *The elements of statistical learning*.

NDC 01-2025, “De nouvelles équations de prévision de l’inflation trimestrielle au Luxembourg

Stock, J. H. (1991). Confidence intervals for the largest autoregressive root in US macroeconomic time series. *Journal of Monetary Economics*, 28(3), 435–459.