Forecastings - Modus Operandi

Methodological note

This paper briefly describes the methodological approach followed in constructing the forecasting models of the references included in the simulation model.

The forecasting activity was divided into the following three consecutive steps:

1. Descriptive analyses.

Each series was analyzed in order to identify and understand any inherent characteristics of the series.

The analytical tools used are as follows:

I. **Seasonal plot**: graph useful in identifying any seasonal (recurring) patterns at the monthly, quarterly and annual levels;

II. **Anomalies Detection plot**: tool used to identify the presence of outliers in the time series;

III. **Time Series Decomposition plot**: graph used to decompose the series into the elementary components of trend, seasonality and error (remainder).

2. Model creation & Validation

Based on the information obtained in the previous step, it was possible to create different models for each reference. In particular, three alternative paths were followed:

I. **Models of classical statistical type and Machine Learning**. More specifically, ETS (Exponential Smoothing), ARIMA (Autoregressive Integrated Moving Average), ARIMA with XGBoost errors (an enhanced version of the previous algorithm), EARTH (Multivariate Adaptive Regression Splines), NNAR (Artificial Neural

Networks), Prophet with XGBoost errors (an algorithm developed by a group of data scientists from Facebook) were created.

II. **Models of Neural Networks**. Set of models involving the creation of more articulated Neural Networks with different numbers of neurons;

III. **H2o AutoML models**. Machine Learning type model classes based on XGBoost, Random Forest, Deep Learning and SVM (Support Vector Machine) algorithms.

The first approach was followed by default for all references; however, in the presence of series with particularly complex trends, the second and third methods were tested. Indeed, it emerged - in the case of series with a rather unpredictable trend - a better fit of Machine Learning type models and in particular of Neural Networks.

For each reference, the statistically best model was selected using Mean Abosolute Error (MAE) and Mean Absolute Percentage Error (MAPE) over the months of testing as error metrics: the lower the value of these two metrics, the higher the accuracy of the model.

3. Forecasting

Forecasting the value of the reference for the next three years was then carried out for each of the estimated models.