Forecasting the Consumer Price Index of Ecuador Using Classical and Advanced Time Series Models

Chapter *in* Communications in Computer and Information Science · October 2023

DOI: 10.1007/978-3-031-45438-7_9

3 authors:

Juan Riofrio
Universidad Yachay Tech
8 PUBLICATIONS **29** CITATIONS

Saba Infante
Yachay Tech University, Ecuador and Universidad de Carabobo, Venezuela
72 PUBLICATIONS **204** CITATIONS

Aracelis Hernández
Universidad de Carabobo, Facultad de Ciencias y Tecnología, Venezuela
18 PUBLICATIONS **49** CITATIONS

Forecasting the Consumer Price Index of Ecuador using Classical and Advanced Time Series Models

Juan Riofrio[1][0000−0001−5666−4947], Saba Infante[1][0000−0001−8883−2730], and Aracelis Hernández[1][0000−0002−6094−4826]

School of Mathematical and Computational Sciences, Yachay Tech University, Urcuquí, Ecuador
{jriofrio, sinfante, ahernandez}@yachaytech.edu.ec

Abstract. In this article, a set of tools are used to forecast and estimate time series models that characterize the patterns generated by Consumer Price Index (CPI) in Ecuador. Estimating this macroeconomic indicator is important because it allows designing public policies, issuing rules that govern the market, controlling inflation, investment risks, in addition to other relevant considerations. To achieve this goal, a variety of models were evaluated including: Support Vector Regression (SVR), Particle filter (PF), SARIMA, Fast Fourier Transform (FFT), Theta Method, among other models. The results obtained in this empirical study demonstrate two promising methods for the prediction of the Ecuadorian CPI. The technique implemented with FFT showed accuracy and forecast robustness, the model allows to identify and analyze the frequency components within a time series and provides valuable information on the dominant patterns, cycles and seasonality present in the data. FFT model obtained a Mean Absolute Error (MAE) of 0.7347 and 0.1340 in the testing and validation steps respectively. The 4Theta model also produced good results, this method allows capturing both the level and the trend of the components of a time series. The 4Theta method achieved a MAE of 0.3453 in the testing step and a validation MAE of 0.8678. This study is the first to employ PF for Ecuador's CPI forecast. Despite suboptimal results, it represents an important initial step in using this method for inflation prediction. To enhance accuracy, a refined movement equation and improved multi-step forecasting with PF are being developed for future work.

Keywords: Consumer Price Index · Time Series Forecasting · Forecasting Models · Darts Module · Inflation.

1 Introduction

The consumer price index (CPI) is a measure of the average change over a time horizon in the prices paid by consumers for a common basket of assets and services. This index attempts to quantify and measure the average cost of life in the countries, estimating the purchasing power in terms of a monetary unit. It is also used to measure inflation or deflation and largely explain the behavior of market dynamics. Accurately estimating the next inflation rate is of great interest to the authorities in charge of designing state public policies and those in charge of dictating the rules that govern the market. For example, Central Banks predict future inflation trends in order to justify the setting of interest rates and control the variation of inflation within a confidence interval that includes the target set [12,4,18]. It is also important to predict future inflation for tax authorities because it allows them to adjust social security payments and tax revenues [4]. In the business sector, it is of interest to estimate future inflation, because it helps them to forecast price dynamics and mitigate investment risks. In the public and private debt sectors, interest payments are governed by the behavior of inflation [1,31].

To estimate the CPI in Ecuador we propose to use some classic time series models, Support Vector Regression, a sequential prediction algorithm, together with some machine learning methods implemented in the Darts Library. Support Vector Regression (SVR), first introduced in [7] in the field of statistical learning theory and structural risk minimization, has proven useful in a variety of prediction problems and classifications. SVR was developed as a regression variation of Support Vector Machine (SVM) method and has a reliable performance in predicting or forecasting time series data. The method tries to find the best hyperplane (dividing line) between classes, it also can address difficulties such as non-linearity, local minimum, and high dimension where the ARIMA model fails. In this cases, SVR overcome overfitting so it has a good performance and it has advantages in optimizing the pattern recognition system with good generalization and accuracy results [26].

Sequential Monte Carlo (SMC) algorithms are techniques used to simulate samples sequentially from distributions that evolve over time, are characterized by being highly flexible and widely applicable in many fields of science. In particular, in this paper we will discuss Particulate Filters (PF) that belongs to the SMC algorithms. The PFs were developed in the decade of the (1990) by [15,19,9,8,14,28], with the purpose of approximating arbitrary distributions, possibly multi-modal and in high-dimensional spaces. The method consists of generating a set of weighed samples called particles to approximate the posterior density of the unknown states. The prediction step involves a Markovian recursion that is updated iteratively based on the states predicted in past times.

The contribution of this work consists in predicting the consumer price index in Ecuador from macroeconomic and financial variables. To achieve the objective, a variety of classical time series models, a state-space model fitted by the particle filter, and machine learning models were estimated. The proposed models will be evaluated using some goodness-of-fit measures and the best forecasting models will be selected. Overall, this study contributes to the field of economic forecasting in Ecuador by providing empirical evidence on the performance of several methods in predicting the CPI. The findings suggest that these methods can serve as valuable tools for policymakers, economists, and businesses in making informed decisions, managing inflation risks, and formulating effective monetary and fiscal policies to ensure stability and economic growth in Ecuador.

The remaining of this paper is structured as follows: Section 2 extend the background of this work and explains similar works to the one proposed in this study. Section 3 describes the metrics, data set and overviews the proposed methods used in this study. Section 4 gathers the main results obtained from the experiments. Section 5 explores the significance of the work's results. Finally, the concluding remarks are drawn in Section 6.

2 Related Work

Inflation forecasting is a challenge of great importance and difficult to achieve. It has been for many decades an area of active research in academia, fiscal institutions, financial entities, central banks, and industry. The main lines of research include structural macroeconomic modeling, forecasting of macroeconomic variables, and the study of monetary policies. Some important works in this area include the work of [33] show that statistics such as average inflation rate, conditional volatility, and persistence levels are shifting in time. They concluded inflation is a non-stationary process, which further limits the amount of relevant historical data points. In [35], the authors compared traditional time series models with machine learning methods for inflation forecasts in the US between the years (1984) and (2014).In the study they concluded that the machine learning model prevails over time series models for forecasting core personal consumption expenditure inflation, and that time series models are better for forecasting core consumer price index inflation. The research developed in [23] compared inflation forecasting with several machine learning models such as lasso regression, random forests, and deep neural networks. The covariates considered were: availability of cash, availability of credit, online prices, housing prices, consumption data, exchange rates and interest rates, they focused on predicting the disaggregated indices that make up the CPI. [4] proposed a hierarchical architecture based on recurrent neural networks to predict the partial disaggregated CPI inflation components. They developed a hierarchical recurrent neural network model that uses information from the higher levels of the CPI hierarchy to improve the predictions at the more volatile lower levels, and concluded that the proposed methodology allows the use of additional forecasting measures to estimate sectoral price changes in some specific components.

In [27], the authors carried out a comparative study of predictive models that include Neural Networks based on a Long Short-Term Memory (LSTM) architecture, Support Vector Regression, SARIMA and Exponential Smoothing, to estimate the consumer price index in Ecuador in a horizon of (12) months. As a result, they obtain that the best model predictive is Support Vector Regression using a polynomial kernel. In [29], the researchers dealt with the CPI forecast problem in Ecuador by using different architectures of an LSTM network, and a convolutional network architecture (CNN). They conclude that combining the CNN network with an LSTM obtains the best forecast for the CPI time series of Ecuador. The method presented in [5] forecasted some time series of the macroeconomy of Ecuador using state-space models combined with Kalman and smoothed Kalman filters. The time series analyzed were Gross Domestic Product (GDP), GDP rate, CPI, industrial production index and active interest rate. In all the cases studied, the estimates obtained by the proposed models reflect the real behavior of the Ecuadorian economy.

3 Materials and Methods

3.1 Particle Filtering

The particle filter (PF) considers the estimation problem of sequentially generated observations, using a transition equation that describes the distribution of a hidden Markov process denoted by $\{x_t, \quad t \in \mathbb{N}\}$, called a vector of latent states (unobserved states), and an observation equation describing the plausibility of measured data at discrete times denoted by $\{y_t, \quad t \in \mathbb{N}\}$. The model is defined in terms of the probability densities:

$$\begin{aligned} x_t &= f(x_{t-1}) + u_t \quad u_t \sim N(0, \sigma_u^2) \quad \text{(state evolution density)} \\ y_t &= h(x_t) + v_t \quad v_t \sim N(0, \sigma_v^2) \quad \text{(observation density)} \end{aligned} \tag{1}$$

where $x_t \in \mathbb{R}^{n_x}$: represent the unobserved states of the system, $y_t \in \mathbb{R}^{n_y}$: represent observations over time t, $f(.), h(.)$: represent the nonlinear functions of the states and of the observations, $u_t \in \mathbb{R}^{n_u}, v_t \in \mathbb{R}^{n_v}$:represents white noise processes. The interest lies in estimating the unknown states $\mathbf{x_{1:t}} = \{x_1, \ldots x_t\}$ from the measurements $\mathbf{y_{1:t}} = \{y_1, \ldots y_t\}$. The joint distribution of the states and observations can be obtained directly by the probability chain rule directly by the probability chain rule:

$$p\left(x_{1:t}, y_{1:t}\right) = f\left(x_1\right) \left(\prod_{k=2}^{t} f\left(x_k | x_{k-1}\right)\right) \left(\prod_{k=1}^{t} h\left(y_k | x_k\right)\right)$$

where $f\left(x_1\right)$ is the distribution of the initial state.
In order that this inference can be performed on high-dimensional models, and with nonlinear structures, many approximation techniques have been proposed; In particular, it is proposed to use the particle filtering methods, the filtering density is approximated with an empirical distribution formed from point masses, or particles. Suppose that we have at time $t-1$ weighted particles

$$\left\{\mathbf{x^{(i)}_{1:t-1}}, \omega_{t-1}^{(i)}, i = 1, \ldots, N\right\} \tag{2}$$

drawn from the smoothing density $p(x_{t-1}|\mathbf{y_{1:t-1}})$, We can consider this an empirical approximation for the density made up of point masses,

$$p_N\left(x_{t-1}|\mathbf{y_{1:t-1}}\right) \approx \sum_{i=1}^{N} \omega_t^{(i)} \delta_{x_{t-1}^{(i)}}(x_{t-1}), \quad \sum_{i=1}^{N} \omega_t^{(i)} = 1, \quad \omega_t^{(i)} \geq 0 \tag{3}$$

where $\delta_{x_{t-1}^{(i)}}(x_{t-1})$ denotes the Dirac-delta function.

To update the smoothing density from time $t-1$ to time t, factorize it as

$$p_N\left(x_t|\mathbf{y_{1:t}}\right) = p_N\left(x_{t-1}|\mathbf{y_{1:t-1}}\right)\frac{h\left(y_t|x_t\right)f\left(x_t|x_{t-1}\right)}{p\left(y_t|\mathbf{y_{1:t-1}}\right)} \tag{4}$$

A new state is then generated randomly from an importance distribution, $q(x_t|x_{t-1}, y_t)$, and appended to the corresponding trajectory, x_{t-1}. The importance weight is updated to:

$$\omega_t^{(i)} = \frac{h(y_t|x_t^{(i)})f(x_t^{(i)}|x_{t-1}^{(i)})}{q(x_t^{(i)}|x_{t-1}^{(i)}, y_t)}\omega_{t-1}^{(i)} \tag{5}$$

where:

$$\omega_{t-1}^{(i)} = \frac{p(x_{t-1}^{(i)}|\mathbf{y_{1:t-1}})}{q(x_{t-1}^{(i)}|\mathbf{y_{1:t-1}})} \tag{6}$$

then

$$p_N(x_t|\mathbf{y_{1:t}}) \approx \sum_{i=1}^{N}\omega_t^{(i)}\delta_{x_t^{(i)}}\left(x_t\right) \tag{7}$$

Given at time $t-1$, $N \in \mathbb{N}$ random samples $\{\mathbf{x}_{1:t-1}^{(i)}\}$ distributed approximately according to $p\left(x_{t-1}|\mathbf{y_{1:t-1}}\right)$, the Monte Carlo filter proceeds as follows at time t:

Step 1: Sequential Importance Sampling

- Generate N i.i.d. samples $\{\tilde{x}_t^{(i)}, i = 1, \ldots, N\}$ from the proposal density $q(x)$:

$$\tilde{x}_t^{(i)} \sim q(x_t|\mathbf{x_{1:t-1}^{(i)}}, \mathbf{y_{1:t}}) = f(x_t|\tilde{x}_{t-1}^{(i)}) + u_t^{(i)} \quad , \quad u_t^{(i)} \sim N(0, \sigma_u^2)$$

  and set $\mathbf{\tilde{x}_{1:t}^{(i)}} = \{\mathbf{x_{1:t-1}^{(i)}}, \tilde{x}_t^{(i)}\}$.
- For $i = 1, ..., N$, evaluate the importance weights up to a normalizing constant

$$\omega_t^{(i)} \propto \frac{h\left(y_t|\mathbf{y_{1:t-1}}, \mathbf{\tilde{x}_{1:t}^{(i)}}\right)f\left(\tilde{x}_t^{(i)}|\tilde{x}_{t-1}^{(i)}\right)}{q(x_t|\mathbf{x_{1:t-1}^{(i)}}, \mathbf{y_{1:t}})}$$

- For $i = 1, ..., N$, normalize the importance weights:

$$\tilde{\omega}_t^{(i)} = \frac{\omega_t^{(i)}}{\sum_{j=1}^{N}\omega_t^{(j)}} \quad , \quad \sum_{i=1}^{N}\tilde{\omega}_t^{(i)} = 1$$

- Evaluate $\hat{N}_{eff} = \frac{1}{\sum_{i=1}^{N}[\tilde{\omega}_t^{(i)}]^2}$

Step 2 Resampling

– If $\hat{N}_{eff} \geq N_{thres}$,

$$x_{1:t}^{(i)} = \tilde{x}_{1:t}^{(i)}, \quad for \quad i = 1, \ldots, N$$

otherwise

– For $i = 1, ..., N$, sample an index $j(i)$ distributed according to the discrete distribution with N elements satisfying

$$p\left(j(i) = l\right) = \tilde{\omega}_t^{(l)}, \quad for \quad l = 1, \ldots, N$$

– For $i = 1, ..., N$, $\mathbf{x_{1:t}^{(i)}} = \tilde{\mathbf{x}}_{\mathbf{1:t}}^{j(i)}$ and $\tilde{w}_t^{(i)} = \frac{1}{N}$.

3.2 Particle Filter Forecaster.

PF uses the value of the last observation y_t to predict the next value. Then, it has to be adapted to forecast several steps ahead from the last observation. We implemented a Particle Filter Forecaster (PFF) that uses the last observation to predict the next value. But as it does not have a new observation to update the particles weights, it uses the predicted value as an observation (it assumes $y_t = x_t$) to update the weights and proceed to predict the next value. This process repeats as many times as values needed for the forecast horizon. It is advisable to not forecast too long into the future. It is important to mention that normally PF uses sensors measurements or a movement equation to predict the next value. Equation (8) presents the movement equation used by our PFF to predict the CPI of Ecuador for the training step, and Equation (9) shows the movement equation used by our PFF to forecast the CPI of Ecuador for the testing and validation steps, both of them represents the term $f(x_{t-1})$ from the state evolution density in Equation 1.

$$f(x_{t-1}) = \sqrt{\mu}\eta x_{t-1} \tag{8}$$

$$f(x_{t-1]}) = \sqrt{\eta}\mu x_{t-1} \tag{9}$$

where μ is last month rate of change $\mu = y_{t-1}/y_{t-2}$ and η is the rate of change in the same month but from last year, specifically $\eta = y_{t-12}/y_{t-13}$. The reason behind the use of this two different equations resides in the availability of true values. In both cases, η in computed from actual values, but in Equation (9), μ is calculated from predicted values. Then, it was decided to give less importance to μ in Equation (9) by obtaining the square root of it.

3.3 Support Vector Regression

The Support Vector Regression (SVR) algorithm was formulated by [37] and is a machine-learning algorithm that is perhaps the most elegant of all kernel-learning methods. The SVR consist of a small subset of data points extracted by the learning algorithm from the training sample itself. Support Vector Machines

(SVM) tackle binary classification problems by converting them into convex optimization problems [36]. The optimization task involves identifying the hyperplane that maximally separates classes while accurately classifying as many training points as possible. Support vectors are used to represent this optimal hyperplane. The SVM sparse solution and its ability to perform a good generalization make it suitable for regression problems. In the context of SVR, the concept of an ϵ-insensitive region is introduced around the function, referred to as the ϵ-tube. This tube reshapes the optimization problem to identify the best-fitting tube for the continuous-valued function while considering the trade-off between model complexity and prediction error. To formulate SVR as an optimization problem, a convex ϵ-insensitive loss function is defined, and the flattest tube enclosing most training instances is sought. Consequently, a multiobjective function is constructed by combining the loss function and the geometric properties of the tube. The convex optimization problem, which has a unique solution, is then solved using appropriate numerical optimization algorithms. The hyperplane is expressed in terms of support vectors, which are training samples lying outside the tube boundary. Similar to SVM, the support vectors in SVR heavily influence the tube's shape [3]. SVR models have recently been used to handle problems such as nonlinear, local minimum, and high dimension. This model can even guarantee higher accuracy for long-term predictions compared to other computational approaches in many applications. We used the last 12 monthly records as input to predict the CPI value of the next month. Then, we encounter the same problem as PF, it can only predict one step ahead. To overcome this problem, we again use the model predictions as input values to predict the next step. This increase the prediction error as it uses predictions that already have an error to predict the next value, but it gives the model the ability to forecast several steps into the future.

3.4 Darts Module

Darts is a Python library developed by [16], which aims to forecast and detect anomalies in univariate and multivariate time series on large data sets. The methodology allows: curve fitting, regression, classification, forecasting, clustering, anomaly detection and reinforcement learning. The library contains a variety of classic time series models, Machine learning and global models, such as: ARIMA, AutoARIMA [32], Baseline Models, Block Recurrent Neural Networks [30], CatBoost model, Croston method, D-Linear, Exponential Smoothing (ES), Fast Fourier Transform (FFT), Kalman Filter Forecaster (KFF), LightGBM Model, Linear Regression model, Neural Basis Expansion Analysis for Interpretable Time Series forecasting (N-BEATS) [24,25], N-HiTS, N-Linear, Prophet [34], Random Forest, Regression ensemble model, Regression Model, Recurrent Neural Networks, Temporal Convolutional Network, Temporal Fusion Transformer (TFT) [20], Theta Method [2], Transformer Model, VARIMA and XGBoost Model. The library also allows to perform retrospective analysis of fitted models, to combine predictions and to simultaneously analyze many model structures at the same time.

FFT method [13] is a useful tool to decompose any deterministic or non-deterministic signal into its frequencies, from which information can be extracted. The FFT provides an excellent mechanism for calculating the Discrete Fourier Transform (DFT) of a time series, that is, discrete data samples [22]. It is widely used in applied mathematics, engineering, and computer science; in particular, in the theory of signal processing systems in high dimensions. The algorithm is fast and of low computational complexity and it allows finding the solution to complex spectral problems [21].

The Theta method was proposed in [2], with the objective of extracting more information from the data, which allows accurate, robust, and reliable univariate time series modeling and forecasting. The Theta model leading to the creation of a Theta line $z(\theta)$, is achieved as the solution of Equation 10:

$$\nabla^2 z_t(\theta) = \theta \nabla^2 y_t = \theta(y_t - 2Y_{t-1} + y_{t+2}), \quad t = 3, 4, \ldots, n, \tag{10}$$

where $y_1, \ldots, y_n$ is the original time series, and ∇ is the difference operator (i.e. $\nabla y_t = y_t - y_{t-1}$).
The methodology decomposes the original data into two or more lines, called Theta lines. Two first points of the Theta line, for $t = 1$ and $t = 2$, can be obtained by minimizing $\sum_{t=1}^{n} [y_t - z_t(\theta)]^2$. An analytical solution of Equation 10 is Equation 11 [17]:

$$z_t(\theta) = \theta y_t + (1 - \theta)(a + bt), \quad \theta \in (0, 1), \quad t = 1, \ldots, n \tag{11}$$

where a and b are constants determined by minimization the sum of squared differences $\sum_{t=1}^{n} [y_t - z_t(\theta)]^2$. The resulting Theta line expressed by Equation 11 is a linear regression model applied to the data directly [10].
The forecasting procedure when using Theta method is carried out in the following steps [11]:

- **Step [1]**. Deseasonalization: Firstly, the time series is tested for statistical significant seasonal behaviour. If it expresses a seasonal component, it is deseasonalised using typically classical multiplicative decomposition.
- **Step [2]**. Decomposition: The time series is decomposed into two Theta lines, $z(0)$ and $z(2)$.
- **Step [3]**. Extrapolation: $z(0)$ is extrapolated as a normal linear regression line, while $z(2)$ is extrapolated using simple exponential smoothing.
- **Step [4]**. Combination: The forecast is generated from the extrapolated $z(0)$ and $z(2)$ lines by their combination with equal weights.
- **Step [5]**. Reseasonalisation: The forecast is reseasonalised if the original time series was identified as seasonal in step 1.

The method performed well, particularly for monthly series and for microeconomic data [2].

3.5 Data set

The data set used during this research was the CPI of Ecuador. The CPI is a nationwide monthly measure that tracks fluctuations in the overall price level of goods and services (inflation) consumed by households in nine Ecuadorian cities. It represents the final consumption patterns of households across different income levels (high, medium, and low) residing in urban areas of the country. Each month, around 25350 price data points are gathered for the 359 items that constitute the CPI Basket of goods and services. As the Ecuadorian economy suffered a major crisis since 1998, the CPI values increased abnormally for a period. We decided to take only the CPI values since the Ecuadorian economy stabilized. Specifically, we only used the CPI values from January 2001 to May 2023. The data set contains 269 monthly records, the base (CPI=100) was adjusted in 2014. The minimum and maximum values in the data set are 49.47 and 110.77 respectively. The mean of the data set is 87.78 and the standard deviation is 17.66. During experimentation, the data set was divided in 3: train, test and validation. The training data set consists of 252 records, from January 2001 to December 2021. The principal objective of this research is to predict the CPI 12 months into the future, then the testing data set has 12 records corresponding to the year 2022. Finally, the validation data set has only 5 records, from January to May 2023 as this are the values that have already been released. The data set can be found at https://github.com/ColdRiver93/CPI-Forecasting. Figure 1 shows the graph of the Ecuadorian CPI divided into the 3 data sets.

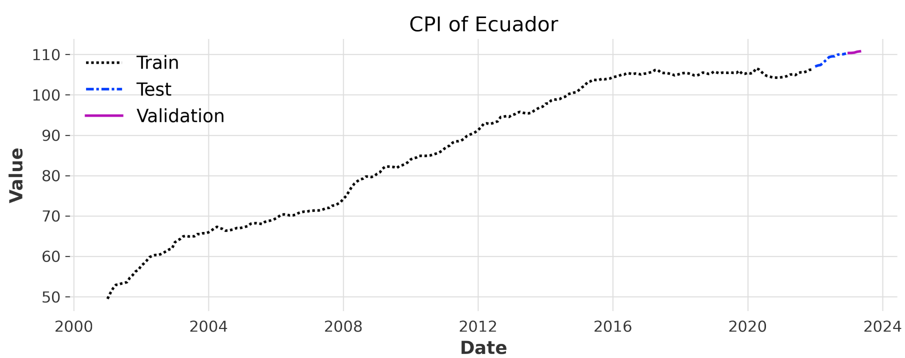

Fig. 1. CPI of Ecuador divided into train, test and validation data sets.

3.6 Metrics

In the literature there is a variety of metrics used to measure model performance, which are based on the difference between the true value and the estimated value $(y - \hat{y})$, or between the squared difference $(y - \hat{y})^2$. These metrics are related to the loss functions in the norms L_1 and L_2 that minimize the error when all differences are summed. These measures put emphasize errors; due to the use of an L_2 norm, predictions that are farther away from the actual values are penalized to a greater extent compared to closer predictions.

1. **The MSE (Mean Squared Error), and RMSE (Root Mean Squared Error)**, often referred to as quadratic loss or L_2 loss is a standard metrics used in model evaluation. For a sample of n observations (y_i) and n corresponding model predictions $\hat{y}_i$, the MSE is given:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2, \quad and \quad RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} = \sqrt{MSE}$$

   The square root does not affect the relative ranks of the models, but produces a metric with the same units of y, which is convenient for estimating the standard error under normally distributed errors. The RMSE has been used as a standard statistical metric to measure model performance in research studies,[6].
2. **The Mean Absolute Error (MAE)**, measures the average of the sum of absolute differences between observation values and predicted values. The MAE is another useful measure widely used in model evaluation. Then, MAE is defined as:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|$$

   Similarly to MSE and RMSE, this method can only have positive values. This is so that is can avoid the cancellation of positive and negative values. On the other hand, there is no error penalization, and therefore this method is not sensitive to outliers.
3. **The mean absolute percentage error (MAPE)**, is one of the most popular measures of the forecast accuracy due to its advantages of scale-independency and interpretability. MAPE is the average of absolute percentage errors (APE). Let y_i At and $\hat{y}_i$ denote the actual and forecast values at data point i, respectively. Then, MAPE is defined as:

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right|$$

   where n is the number of data points.

4 Results

The main results of the experimentation phase are presented in this section. In Section 3 we explained the division of the data set into 3 smaller data sets for the training, testing and validation steps. We gathered the metrics of the testing and validation steps to compare the different models. We performed a grid search to get the best hyperparameters of each method explained in Section 3, and choose the model with the best MAE in the testing step. After selecting the best model, we forecasted the CPI for each month of 2023 and compared the first five months with the validation data set.

Normally, PF is used to estimate the value x_t from x_{t-1}. Then, this method is good if we want to predict 1 step into the future. During this research, the task was to forecast the next 12 values from the last actual value obtained. Therefore, we performed some adjustments to the method so it can forecast the next 12 values. We present the results of both implementations to compare their performance. As expected, the performance of the PF is better than the PFF as it uses the actual recent values to update the weights. PFF trades accuracy for the ability to forecast several months into the future. PF obtained a MAE of 0.2906 in the testing step, compared to the MAE of 1.9181 obtained by the PFF. During the validation step, PF and PFF obtained a MAE of 0.1019 and 0.3599, respectively. This results proved that PF is a good method to predict one step into the future, but some work should be done to apply PF to the forecasting problem (PFF).

SVR method also has 2 different implementations. The first implementation uses the actual CPI values as input to the model, but it only predicts one step into the future. The other implementation, uses its own predictions as input to predict several steps into the future (i.e. 12 months into the future). We named this second implementation as SVR Forecaster (SVRF). Here it is also expected that SVR has better error scores than SVRF, but SVRF is able to forecast the CPI for the complete year. SVR scored a testing and validation MAE of 0.3537 and 0.0889 respectively, showing a great performance when the task is to predict one step into the future. On the other hand, SVRF scored a higher testing MAE of 2.0972 and an a slightly higher validation MAE of 0.1054. It is important to mention that both models were the same (same hyperparameters), the difference in both implementations were the input given to the model in the forecasting step. The best model found had a polynomial second degree kernel with a kernel coefficient equal to 15 and an independent term equal to 15. The model used an epsilon value of 0.001 and a regularization parameter of 0.01.

In macroeconomic indicators' forecasting, usually KFF is used instead of PFF. Darts module has an implementation of KFF that we used to forecast the next 12 months. After the grid search, we found that the best size of the Kalman filter state vector is 1 for this particular problem. We can compare the KFF with the PFF as they performed the same task. The KFF outperformed the PFF in the testing step obtaining a MAE of 1.1821, but during validation the proposed PFF had better results than KFF that scored a MAE of 0.8831. We used other methods implemented in the Darts module, not only the KFF.

We also used the ES method presented by Holt and Winters from Darts module to forecast the next 12 months. The hyperparameters of the best model found were a multiplicative type of trend component that is damped, an additive type of seasonal component and three periods in a complete seasonal cycle. This model obtained a MAE of 0.7212 in the testing step and a MAE of 3.4545 in the validation step. This values indicates that the model was overfitted to the test data set during the grid search, and had a poor performance during validation.

ARIMA method, is broadly used for time series forecasting. Darts module also presents an implementation of SARIMA, as well as an AutoARIMA implementation. AutoARIMA automatically search for the best ARIMA model that looses the less information about the time series. AutoARIMA produced bad results scoring a MAE of 2.1585 and 3.1185 in the testing and validation steps respectively. The best SARIMA model found had the parameters (1,2,1) and for the seasonal component (4,2,4) with 3 periods per season. This model got a MAE of 0.2618 in the testing step and during validation it obtained a MAE of 1.6064. SARIMA obtained the best score in the testing step, but during validation this model did not behave like expected. This model was overfitted to the testing data and did not forecasted correctly for the validation data set.

The 4Theta method included in Darts module obtained one of the best results in the testing step. Even though, SARIMA was the model with the best testing score, as explained before, it was overfitted. The best 4Theta model was the model with a value of the theta parameter equal to 24, a seasonality period of 24, an additive type of model combining the Theta lines, a multiplicative type of seasonality, and an exponential type of trend to fit. The best model of this method obtained a testing MAE of 0.3453 and a validation MAE of 0.8678.

We considered FFT model to be the best model to predict the CPI of Ecuador for 2023 together with the 4Theta model. This model got a testing MAE of 0.7347, and a validation MAE of 0.1340, the best validation score of all the models. Although, this model did not obtained the best testing score, according to the validation score it was the best model to predict the Ecuadorian CPI for the next months of 2023. The best model used one frequency for forecasting, and a polynomial kind of detrending was applied before performing DFT, with a second degree polynomial used for detrending.

We also used an Artificial Neural Network (ANN) method called Neural Basis Expansion Analysis Time Series Forecasting (N-BEATS). This method implemented in the Darts module, provide us with a powerful forecasting tool. In this method we tuned several hyperparameters to create the ANN architecture. The best ANN architecture found using grid search had six stacks that make up the whole model with 12 blocks making up every stack. It was formed of six fully connected layers preceding the final forking layers in each block of every stack, made up of 12 neurons each. The expansion coefficients was set to six. We used the Softplus as the activation function of the encoder/decoder intermediate layer. The length of the input sequence fed to the model was 12 records, representing an entire year of monthly records. The length of the output forecast of the model was one, meaning we use the 12 previous months to forecast the next month. The testing and validation MAE obtained by the model were 1.3110 and 1.2793 respectively.

The complete results of all the methods are presented in Table 1, the t-MAE, t-MAPE and t-RMSE columns show the testing errors, while the t-MAE, t-MAPE and t-RMSE columns present the validation errors. It is important to remember that PF and SVR had better results because they used actual values to predict only one step into the future.

Table 1. Testing and validation MAE, MAPE, RMSE for each method

Method	t-MAE	t-MAPE	t-RMSE	v-MAE	v-MAPE	v-RMSE
PF	0.2906	0.002676	0.3832	0.1019	0.000922	0.1161
PFF	1.9181	0.017537	2.0747	0.3599	0.003254	0.4178
SVR	0.3537	0.003252	0.4479	0.0889	0.000805	0.0971
SVRF	2.0972	0.019172	2.2754	0.1054	0.000954	0.1145
KFF	1.1821	0.010815	1.2676	0.8831	0.007992	0.8994
ES	0.7212	0.006574	1.0032	3.4545	0.031248	3.5001
AutoARIMA	2.1585	0.019734	2.3369	3.1185	0.028215	3.1193
SARIMA	0.2618	0.002395	0.3541	1.6064	0.014526	1.7247
4Theta	0.3453	0.003176	0.4228	0.8678	0.007845	1.0098
FFT	0.7347	0.006817	1.0391	0.1340	0.001213	0.1432
N-BEATS	1.3110	0.011989	1.4175	1.2793	0.011576	1.2852

5 Discussion

In this section the results presented in Section 4 are analyzed. Also, the limitations encountered during this work are discussed and the ground for future work is determined. First, lets start by defining the best methods according to the results. The best methods in the testing step were SARIMA, 4Theta and ES and FFT in that order. While, the best methods in the validation step were SVRF, FFT, PFF and 4Theta in that order. Then, we consider FFT and 4Theta methods to be the best methods to forecast the CPI of Ecuador one year into the future because of their good performance during both, testing and validation. Table 2 shows the forecasted CPI values for the remaining months of 2023 performed by FFT and 4Theta methods. While, Figure 2 visually presents the forecasted CPI values for 2023.

Table 2. Forecasted CPI values for the remaining months of 2023 using FFT and 4Theta methods.

Method	Jun	Jul	Aug	Sep	Oct	Nov	Dec
FFT	110.97	111.07	111.17	111.26	111.34	111.43	111.51
4Theta	112.26	112.4	112.53	113.22	113.59	113.96	114.33

In Figure 2, it can be seen that FFT forecast is more conservative, predicting little inflation for the rest of the year. While, 4Theta forecast predicts a high inflation during the last months of 2023. The best case scenario for Ecuadorians is the predictions made by the FFT model, as it predicts prices are slightly increasing, but not too much.

The CPI is a macroeconomic indicator that reflects the inflation of a country. Inflation in Ecuador greatly increased since the first months of 2022. This brings complications during the forecasting task performed in this study. The problem is that the testing data set contains the values of 2022 when Ecuador faced the

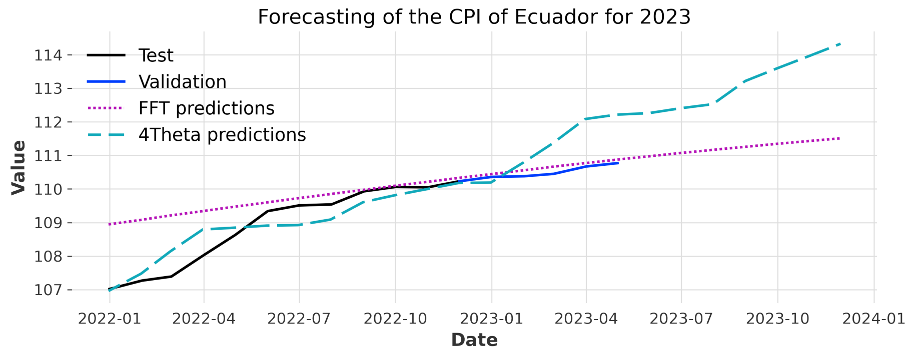

Fig. 2. Forecasted CPI of Ecuador for 2023 using FFT and 4Theta methods.

greater inflation increase since 2015. Then, some models could not predict this high inflation as the last months of the training data set present little or no inflation. Also, as we tested the models with the values showing a great inflation increase, some models predicted the same high inflation increase for the next year. This problem makes the models prone to overfitting, thus a different train-test-validation data set split should be considered.

We believe the N-BEATS model can generate better results. Due to the high number of hyperparameters to tune and the computational cost of the algorithm, we did not tested enough models until we find an optimal result. A wider grid search should be performed to find a model that better forecast the CPI of Ecuador. Even better, an optimization algorithm can be used to find the optimal ANN architecture for the used data set. The Darts module implementation of the N-BEATS method is a powerful tool that should be better exploited.

It has been seen that PF works well for this kind of problem. The proposed PF moves the particles according to the inflation variation of the last month and last year, so it predicts better when we have a constant inflation increase or decrease over the months. On the other hand, the proposed PF does not behave well when the values are continuously increasing and decreasing over time (e.g. FOREX data). On other words, the proposed method works better with smooth time series with trend. As we can see in Figure 1, this data set mostly complies with this characteristics. Also PF adjust the particle weights according to the last actual observation, giving this method the ability to correct the measurements once a new value is received. PF showed to be a good method when we need to predict only one step ahead.

The accuracy of the PF method depends on the movement equation to predict the movement of the particles from time $t-1$ to t. We used Equation 8 and Equation 9 to predict the movement of the particles trying to predict the actual value, but this equation should be reformulated in future work to better predict the CPI of Ecuador. The implementation proposed in this research is a good approximation, but for better results a new study should be done concerning that matter.

This work is the first one to use PF in forecasting the CPI of Ecuador. Although the results were not the best, it was a good first step using this method to predict the Ecuadorian inflation. As said before, a better movement equation should be developed to improve the predictions. Also, a best method for the PFF to forecast several steps into the future is being developed for future work. For example, using FFT predictions as actual values to update the particles weights when we forecast 12 months ahead. Finally, this work is completely reproducible, the data set can be found at https://github.com/ColdRiver93/CPI-Forecasting along with the Python implementation used during this research.

6 Conclusions

This work proposed a methodology to forecast the CPI of Ecuador for the remaining months of 2023. Several methods were used to complete this task, ranging from classical models like SARIMA to more recent methods as N-BEATS. We also used other methods during this research, in specific we used PF, SVR, KFF, ES, 4Theta, and FFT methods. We searched for the best method to forecast the CPI of Ecuador one year into the future. The study aimed to provide insights into the accuracy and effectiveness of these forecasting techniques in capturing the complex dynamics of CPI fluctuations in the Ecuadorian economy.

The findings of this research indicate that two methods have demonstrated promising results in predicting the CPI. FFT showed strong forecasting accuracy and robustness, with its capability to identify and analyze the frequency components within a time series and to provide valuable information about the dominant patterns, cycles, and seasonality present in the data. The 4Theta method, which combines exponential smoothing and linear regression to capture both the level and trend components of a time series, also produced accurate predictions.

However, it is worth noting that forecasting accuracy can be influenced by various factors, such as the length and quality of the available historical data, the choice of model parameters, and the specific characteristics of the CPI series. Therefore, future research should explore additional approaches, such as ensemble models or hybrid methodologies, to further enhance forecasting accuracy and robustness.

Acknowledgements This work was carried out under the project Statistical methods for modeling data generated by complex systems with registration MATH23-04, Yachay Tech University. The authors would like to acknowledge the help received from “Laboratorio de la Facultad de Ingeniería de Sistemas” from Escuela Politécnica Nacional. Without the computational resources provided by the laboratory, this research would not have been possible.

References

1. Almosova, A., Andresen, N.: Nonlinear inflation forecasting with recurrent neural networks. Journal of Forecasting **42**(2), 240–259 (2023)
2. Assimakopoulos, V., Nikolopoulos, K.: The theta model: a decomposition approach to forecasting. International journal of forecasting **16**(4), 521–530 (2000)
3. Awad, M., Khanna, R., Awad, M., Khanna, R.: Support vector regression. Efficient learning machines: Theories, concepts, and applications for engineers and system designers pp. 67–80 (2015)
4. Barkan, O., Benchimol, J., Caspi, I., Cohen, E., Hammer, A., Koenigstein, N.: Forecasting cpi inflation components with hierarchical recurrent neural networks. International Journal of Forecasting **39**(3), 1145–1162 (2023)
5. Bautista Vega, H., Infante, S., Amaro, I.R.: Estimation of the state space models: An application in macroeconomic series of ecuador. In: Conference on Information and Communication Technologies of Ecuador. pp. 31–45. Springer (2021)
6. Chai, T., Draxler, R.R.: Root mean square error (rmse) or mean absolute error (mae)?–arguments against avoiding rmse in the literature. Geoscientific model development **7**(3), 1247–1250 (2014)
7. Cortes, C., Vapnik, V.: Support-vector networks. Machine learning **20**, 273–297 (1995)
8. Doucet, A., De Freitas, N., Gordon, N.: An introduction to sequential monte carlo methods. Sequential Monte Carlo methods in practice pp. 3–14 (2001)
9. Doucet, A., Godsill, S., Andrieu, C.: On sequential monte carlo sampling methods for bayesian filtering. Statistics and computing **10**, 197–208 (2000)
10. Dudek, G.: Short-term load forecasting using theta method. In: E3S Web of Conferences. vol. 84, p. 01004. EDP Sciences (2019)
11. Fiorucci, J.A., Pellegrini, T.R., Louzada, F., Petropoulos, F., Koehler, A.B.: Models for optimising the theta method and their relationship to state space models. International Journal of Forecasting **32**(4), 1151–1161 (2016)
12. Friedman, M.: The lag in effect of monetary policy. Journal of Political Economy **69**(5), 447–466 (1961)
13. Ghaderpour, E., Pagiatakis, S.D., Hassan, Q.K.: A survey on change detection and time series analysis with applications. Applied Sciences **11**(13), 6141 (2021)
14. Godsill, S.J., Doucet, A., West, M.: Monte carlo smoothing for nonlinear time series. Journal of the american statistical association **99**(465), 156–168 (2004)
15. Gordon, N.J., Salmond, D.J., Smith, A.F.: Novel approach to nonlinear/non-gaussian bayesian state estimation. IEE proceedings F (radar and signal processing) **140**(2), 107–113 (1993)
16. Herzen, J., Lässig, F., Piazzetta, S.G., Neuer, T., Tafti, L., Raille, G., Van Pottelbergh, T., Pasieka, M., Skrodzki, A., Huguenin, N., et al.: Darts: User-friendly modern machine learning for time series. The Journal of Machine Learning Research **23**(1), 5442–5447 (2022)
17. Hyndman, R.J., Billah, B.: Unmasking the theta method. International Journal of Forecasting **19**(2), 287–290 (2003)
18. Ida, D.: Sectoral inflation persistence and optimal monetary policy. Journal of Macroeconomics **65**, 103215 (2020)
19. Kitagawa, G.: Monte carlo filter and smoother for non-gaussian nonlinear state space models. Journal of computational and graphical statistics **5**(1), 1–25 (1996)
20. Lim, B., Arık, S.Ö., Loeff, N., Pfister, T.: Temporal fusion transformers for interpretable multi-horizon time series forecasting. International Journal of Forecasting **37**(4), 1748–1764 (2021)

21. Liu, C., Chen, K., Zhang, J., Wang, Y., Wang, H.: Using fft to reduce the computational complexity of sub-nyquist sampling based wideband spectrum sensing. In: Journal of Physics: Conference Series. vol. 1237, p. 022004. IOP Publishing (2019)
22. Liu, W., Liao, Q., Qiao, F., Xia, W., Wang, C., Lombardi, F.: Approximate designs for fast fourier transform (fft) with application to speech recognition. IEEE Transactions on Circuits and Systems I: Regular Papers **66**(12), 4727–4739 (2019)
23. Medeiros, M.C., Vasconcelos, G.F., Veiga, Á., Zilberman, E.: Forecasting inflation in a data-rich environment: the benefits of machine learning methods. Journal of Business & Economic Statistics **39**(1), 98–119 (2021)
24. Oreshkin, B.N., Carpov, D., Chapados, N., Bengio, Y.: N-beats: Neural basis expansion analysis for interpretable time series forecasting. arXiv preprint arXiv:1905.10437 (2019)
25. Oreshkin, B.N., Carpov, D., Chapados, N., Bengio, Y.: Meta-learning framework with applications to zero-shot time-series forecasting. Proceedings of the AAAI Conference on Artificial Intelligence **35**(10), 9242–9250 (2021)
26. Ramageri, B.M., et al.: Data mining techniques and applications. Indian journal of computer science and engineering **1**(4), 301–305 (2010)
27. Riofrío, J., Chang, O., Revelo-Fuelagán, E., Peluffo-Ordóñez, D.H.: Forecasting the consumer price index (cpi) of ecuador: A comparative study of predictive models. International Journal on Advanced Science, Engineering and Information Technology **10**(3), 1078–1084 (2020)
28. Ristic, B., Arulampalam, S., Gordon, N.: Beyond the Kalman filter: Particle filters for tracking applications. Artech house (2003)
29. Rosado, R., Abreu, A.J., Arencibia, J.C., Gonzalez, H., Hernandez, Y.: Consumer price index forecasting based on univariate time series and a deep neural network. In: International Workshop on Artificial Intelligence and Pattern Recognition. pp. 33–42. Springer (2021)
30. Salinas, D., Flunkert, V., Gasthaus, J., Januschowski, T.: Deepar: Probabilistic forecasting with autoregressive recurrent networks. International Journal of Forecasting **36**(3), 1181–1191 (2020)
31. Shibitov, D., Mamedli, M.: Forecasting russian cpi with data vintages and machine learning techniques. Tech. rep., Bank of Russia (2021)
32. Smith, T.G., et al.: pmdarima: Arima estimators for python. Retrieved from **309** (2017)
33. Stock, J.H., Watson, M.W.: Why has us inflation become harder to forecast? Journal of Money, Credit and banking **39**, 3–33 (2007)
34. Taylor, S.J., Letham, B.: Forecasting at scale. The American Statistician **72**(1), 37–45 (2018)
35. Ülke, V., Sahin, A., Subasi, A.: A comparison of time series and machine learning models for inflation forecasting: empirical evidence from the usa. Neural Computing and Applications **30**, 1519–1527 (2018)
36. Vapnik, V.: The support vector method of function estimation. In: Nonlinear modeling: Advanced black-box techniques, pp. 55–85. Springer (1998)
37. Vapnik, V.N.: A note on one class of perceptrons. Automat. Rem. Control **25**, 821–837 (1964)