WAVELET-BASED MULTISCALE APPROACH IN HIGH-DIMENSIONAL FORECASTING MODELS

ALDRYN QUISPE

Abstract. This paper proposes a multiscale forecasting framework that combines wavelet-based decompositions with high-dimensional estimation techniques. By decomposing inflation into frequency components and estimating separate models for each component using frequency-matched predictors, the approach captures heterogeneous dynamics across short- and long-run horizons. An optimization step selects, for each forecast horizon, the combination of components and models that minimizes out-of-sample forecasting error. We apply the methodology to forecast Brazilian inflation and evaluate its performance against standard benchmarks, including random walk and autoregressive models, survey-based expectations, and high-dimensional forecasting methods. The results show that the proposed method delivers substantial gains at short horizons and generally improves upon survey and high-dimensional benchmarks, although gains narrow at longer horizons. The component selection reveals that different estimation strategies specialize across frequency bands. Overall, the findings suggest that exploiting frequency-specific information through wavelet decompositions, combined with high-dimensional methods, provides a promising and interpretable tool for inflation forecasting.
Keywords: Wavelets; Multiscale modeling; Frequency-specific forecasting; High-dimensional forecasting; Inflation forecasting.

1. INTRODUCTION

Forecasting time series accurately is a fundamental task in many fields, including finance, macroeconomics, energy demand, and environmental monitoring. However, real-world time series often exhibit complex dynamics such as nonstationarity, structural breaks, seasonality, and patterns that evolve across multiple temporal scales. Traditional forecasting models, such as autoregressive and moving average specifications, typically rely on the raw series or a limited number of lagged variables as predictors. Although these models can capture simple temporal dependencies, they may fail to fully exploit richer structures embedded in the data, particularly when relevant information is distributed across different frequencies or time horizons.

The forecasting of macroeconomic and financial variables has become even more challenging with the growing availability of large datasets. In data-rich environments, researchers and practitioners can exploit a large number of potential predictors to improve predictive performance. This setting has motivated the development of high-dimensional forecasting methods that rely on regularization and variable selection techniques. Among these, the LASSO has become particularly popular due to its ability to simultaneously perform parameter estimation and variable selection while controlling model complexity. Empirical evidence shows that high-dimensional models combined with appropriate shrinkage and model selection procedures can significantly improve forecasting accuracy in macroeconomic applications (Medeiros and Vasconcelos, 2016).

At the same time, another strand of the literature has emphasized the importance of modeling the multiscale nature of economic and financial time series. Many economic variables evolve simultaneously at different time horizons, reflecting short-term fluctuations, medium-term business cycle dynamics, and long-run trends. Wavelet analysis provides a flexible framework for capturing these features by decomposing a time series into components associated with different frequency bands while preserving temporal localization. This property allows researchers to analyze and model dynamics operating at distinct time scales, making wavelets particularly suitable for forecasting applications. Crowley (2007) provides a comprehensive survey of wavelet methods, highlighting their applicability to business cycle analysis, filtering, and forecasting. Early applications of wavelet-based methods to economic forecasting have shown that wavelet-based multi-resolution analysis can improve forecasting performance by isolating heterogeneous dynamics that may be obscured in standard time-domain representations (Fernandez, 2008; Rua, 2011, 2017).

More recent contributions have provided further evidence that frequency-specific information is central to forecasting accuracy. Faria and Verona (2018) show that decomposing predictors into frequency-specific components via wavelet multiresolution analysis and forecasting each component separately yields substantial out-of-sample gains in equity premium prediction, with improvements driven by the ability to isolate informative frequency bands from noisy components. Verona (2026) extends this logic by proposing a sum-of-the-cycles method that decomposes inflation into frequency-specific components via wavelet analysis and models each component separately. Applied to U.S. inflation, the key insight is that predictors are heterogeneous in their frequency content, and aligning each predictor with the inflation component it best captures extracts substantially more predictive information from the data.

Other recent contributions combine wavelet decomposition of predictors with machine learning techniques, demonstrating that the relative importance of different frequency bands varies over the evaluation period, indicating that the predictive content of short-term and long-term dynamics is time-varying (Risse, 2019; Alexandridis et al., 2024). Sengupta et al. (2025) further demonstrate the value of wavelet-based decompositions for inflation forecasting in emerging economies, applying a filtered ensemble wavelet neural network to CPI inflation in BRIC countries and showing that wavelet-transformed predictors combined with exogenous uncertainty measures improve forecast performance relative to standard approaches. These findings indicate that wavelet-based representations provide a mechanism for uncovering predictive structure that is distributed across frequencies and that standard time-domain models fail to exploit.

Although these contributions have demonstrated the value of frequency-specific information for prediction across different asset classes and macroeconomic variables, the systematic combination of wavelet-based predictor reparametrizations with sparse high-dimensional estimation methods in macroeconomic forecasting has received comparatively less attention. While wavelet techniques provide a rich set of components capturing dynamics at different frequencies, selecting the most informative components for prediction in high-dimensional settings remains a methodological challenge.

This paper proposes a forecasting framework that combines multiscale wavelet representations with sparse predictive modeling. The proposed approach extends the sum-of-the-cycles logic to a high-dimensional Brazilian inflation forecasting environment. Following Verona (2026), inflation is decomposed into frequency components and separate high-dimensional models are estimated for each component. Specifically, each series is decomposed into wavelet coefficients at multiple scales, which are then reconstructed to form new time series representing fluctuations at different frequencies. These derived series are matched with the corresponding inflation component and used as predictors in high-dimensional estimation methods, allowing for data-driven selection of the most relevant frequency components for prediction. By combining the multiscale representation provided by wavelets with high-dimensional methods that exploit sparsity and dimensionality reduction, the proposed framework captures heterogeneous dynamics across frequencies, enhances interpretability, and mitigates overfitting in data-rich environments.

To assess the empirical relevance of the proposed methodology, we apply it to the problem of forecasting Brazilian inflation, a setting that provides a natural testing ground for high-dimensional and multiscale predictive methods. Inflation dynamics in emerging economies such as Brazil are known to be complex, reflecting the interaction of monetary policy, exchange rate movements, fiscal conditions, and expectations formation. These dynamics often evolve across multiple time horizons, with short-run fluctuations driven by transitory shocks and long-run behavior shaped by structural and policy-related factors. As a result, accurately forecasting inflation requires models capable of capturing heterogeneous temporal patterns as well as exploiting information from a large set of macroeconomic indicators.

The empirical analysis builds on the data-rich environment considered by Garcia et al. (2017), which combines a large panel of macroeconomic variables with survey-based expectations from professional forecasters. This dataset is particularly suitable for evaluating the proposed framework, as it contains both realized economic indicators and forward-looking measures of expectations, which are known to play a central role in inflation dynamics, especially at short horizons. At the same time, the relatively large number of predictors poses a challenge for traditional forecasting models, making it necessary to adopt methods that can effectively handle high dimensionality and perform variable selection.

Within this context, the proposed wavelet-based multiscale approach provides a natural extension of the standard forecasting framework. By decomposing macroeconomic variables into components associated with different frequency bands, the method allows the forecasting model to exploit information contained in short-term fluctuations, medium-term cycles, and long-term trends. When combined with high-dimensional predictive techniques such as LASSO, targeted factor models, and complete subset regression, the framework can select the most informative variables and frequency components in a data-driven manner.

The empirical application therefore serves a dual purpose. First, it evaluates whether forecasting each frequency component of inflation separately, and recombining the resulting forecasts, improves accuracy relative to models that target the aggregate inflation series directly. Second, it provides evidence on how different frequencies contribute to predictive performance across forecast horizons, shedding light on the role of multiscale dynamics in inflation forecasting. This setting allows us to assess the practical gains from integrating frequency-specific modeling with modern high-dimensional forecasting methods in a realistic and policy-relevant environment.

The contributions of this paper are threefold. First, it extends the frequency-specific forecasting approach to a high-dimensional setting, formalizing a systematic procedure for decomposing macroeconomic predictors into frequency components and matching them with the corresponding inflation component. Second, it demonstrates the complementarity between wavelet-based frequency decomposition and high-dimensional estimation methods that exploit sparsity and dimensionality reduction, highlighting how different specifications contribute to predictive performance across forecast horizons. Third, it provides an empirical evaluation of the proposed methodology, extending the evidence on frequency-based forecasting gains to a data-rich emerging economy setting.

The remainder of the paper is organized as follows. Section 2 presents the proposed methodology, including the wavelet decomposition procedure and the frequency-specific forecasting framework. Section 3 describes the empirical application, the dataset, and the forecasting performance evaluation. Finally, Section 4 concludes and discusses possible directions for future research.

2. A WAVELET-BASED MULTISCALE APPROACH

Wavelets are time-frequency filtering and decomposition tools that separate the dynamics of a time series into components associated with distinct frequency bands while retaining temporal localization. Wavelet analysis relies on functions generated by translation and dilation operations, which allow a signal to be represented across different resolution scales. In contrast to standard Fourier analysis, wavelets provide a multiresolution representation that isolates the dynamics of a time series into components associated with specific frequency bands while preserving temporal localization, useful when the behavior of the series is not homogeneous over time (Percival and Walden, 2000; Gençay et al., 2002).

One of the key advantages of wavelet representations is their ability to effectively analyze non-stationary signals. Many economic and financial time series exhibit evolving statistical properties, such as time-varying volatility, structural breaks, and transient cycles. Because wavelets are localized in both time and frequency, they are able to capture short-lived features and abrupt changes that may be obscured in global spectral representations. This property makes wavelets particularly suitable for analyzing signals whose frequency content varies over time, allowing the detection of localized oscillations and time-varying periodicities (Percival and Walden, 2000; Gençay et al., 2002).

Wavelet decompositions also provide a powerful framework for representing and reconstructing complex signals. Through multiresolution analysis, a signal can be expressed as the sum of components operating at different scales, allowing for flexible reconstruction, filtering, and denoising procedures. These properties have been widely used in signal processing and in the numerical representation of stochastic processes. In particular, wavelet bases provide sparse and efficient representations of stochastic fields, which has motivated their application in the approximation and analysis of stochastic differential equations (SDEs) and stochastic partial differential equations (SPDEs), where multiscale structures naturally arise (Mallat, 1998; Meyer, 1992).

Another important development concerns the relationship between wavelets and sparse signal representations, which underlies modern compressed sensing techniques. Many natural signals admit sparse representations in wavelet bases, meaning that most of the information of the signal can be captured by a relatively small number of significant coefficients. This sparsity property enables efficient signal compression, denoising, and recovery from incomplete observations, forming the basis for compressed sensing methods in high-dimensional signal processing and statistics (Donoho, 1995; Candès and Wakin, 2008).

These properties have also motivated the use of wavelets in forecasting applications. By decomposing a time series into multiple frequency components, wavelet analysis allows predictive models to exploit information embedded in different temporal scales. Short-term fluctuations, medium-term cycles, and long-term trends can be modeled separately and subsequently combined to improve predictive performance. As a result, wavelet-based representations provide a flexible framework for incorporating multiscale information into forecasting models, particularly in environments characterized by complex dynamics and large predictor sets (Gençay et al., 2002).

Formally, a wavelet basis is generated from a mother wavelet $\psi(\cdot)$ and the scaling function $\phi(\cdot)$. The family of functions can be written as

$$\psi_{j,k}(t) = 2^{-j/2}\psi(2^{-j}t - k) \qquad \phi_{j,k}(t) = 2^{-j/2}\phi(2^{-j}t - k) \tag{1}$$

where j controls the scale and k the time location. Under this representation, the scaling function captures the smooth, low-frequency component of the signal, whereas the wavelet describes its local fluctuation across different resolution levels.

The Discrete Wavelet Transform (DWT) arises from the dyadic discretization of the scale and location parameters and allows a finite discrete series to be rewritten in terms of detail and smooth coefficients. Under this representation, the series is projected onto an orthonormal family of discretized wavelet functions, so that the coefficients $d_{j,k}$ measure the contribution of the detail components at scale j and location k, while $a_{J,k}$ denotes the smoothing coefficient at the coarsest scale. This decomposition can be viewed as the outcome of applying a high-pass wavelet filter $h = \{h_l\}_{l=0}^{L-1}$ and a low-pass scaling filter $g = \{g_l\}_{l=0}^{L-1}$, where the coefficients are obtained by

$$d_{j,k} = \sum_{l=0}^{L-1} h_l\, x_{2k-l}, \qquad a_{j,k} = \sum_{l=0}^{L-1} g_l\, x_{2k-l}. \tag{2}$$

The index $2k$ reveals the dyadic downsampling intrinsic to the DWT. At each scale, only every other observation enters the computation, halving the number of coefficients (Percival and Walden, 2000). While this makes the DWT computationally efficient, it introduces sensitivity to sample alignment that becomes particularly relevant in rolling window estimation, where the transform is recomputed at each forecasting origin. As a result, at scale j the decomposition retains $T/2^j$ coefficients, and the smooth component at the coarsest level J is described by $T/2^J$ coefficients.

The Maximal Overlap Discrete Wavelet Transform (MODWT) addresses two limitations of the DWT that are relevant here by removing the downsampling step. As introduced by Percival and Walden (2000), the MODWT is defined through rescaled versions of the wavelet and scaling filters, $\tilde{h}_l = h_l/\sqrt{2}$ and $\tilde{g}_l = g_l/\sqrt{2}$. For $j = 1, \ldots, J$ and $t = 0, \ldots, T-1$, the wavelet and scaling coefficients are given recursively by

$$\tilde{W}_{j,t} = \sum_{l=0}^{L-1} \tilde{h}_l\, \tilde{V}_{j-1,t-2^{j-1}l} \qquad \tilde{V}_{j,t} = \sum_{l=0}^{L-1} \tilde{g}_l\, \tilde{V}_{j-1,t-2^{j-1}l} \tag{3}$$

where $\tilde{V}_{0,t} = x_t$. Since no downsampling is applied, both $\tilde{W}_{j,t}$ and $\tilde{V}_{j,t}$ retain the same length T as the original series at every scale j. The term $2^{j-1}l$ in the summation index shows that, as j increases, the coefficients combine observations further apart in time, capturing increasingly persistent dynamics. As discussed in Percival and Walden (2000), a key property of the MODWT is *shift-invariance*. In contrast to the DWT, the transform is not sensitive to the starting point of the sample, which makes it particularly suitable for rolling window estimation.

The MODWT coefficients can be used to decompose and reconstruct the series into additive multiresolution components associated with different resolution levels. In multiresolution analysis (MRA), the original series is written as the sum of a smooth component and a set of detail components. Specifically, any series x_t can be reconstructed as

$$x_t = \sum_{j=1}^{J} D_{j,t} + S_{J,t} \tag{4}$$

where $D_{j,t}$ denotes the detail component at scale j, and $S_{J,t}$ denotes the smooth component capturing the low-frequency content not accounted for by the detail levels. Each component is reconstructed from the MODWT coefficients via the inverse transform, preserving the same length T as the original series. The detail component $D_{j,t}$ isolates variations occurring over 2^{j-1} to 2^j periods, so that finer scales correspond to short-run fluctuations and coarser scales to more persistent dynamics.

In the empirical application considered in this paper, the multiresolution decomposition is applied to both the inflation series and the macroeconomic variables in the dataset, following the frequency-specific forecasting framework described in the next section. We use the high-dimensional dataset proposed by Garcia et al. (2017), and the wavelet decomposition is applied only to the macroeconomic variables in the dataset. We consider the MODWT with a Daubechies[1] wavelet filter of length 4 and $J = 1$. For each series x_t we obtain two components: $D_{1,t}$, the detail component capturing short-run fluctuations, and $S_{1,t}$, the smooth component capturing the low-frequency content. The same decomposition is applied to the inflation target, yielding a detail component $\pi_t^{D_1}$ and a smooth component $\pi_t^{S_1}$, which are forecast separately and subsequently recombined.

2.1. **Wavelet-Based Multiscale Predictive Framework.** This paper proposes a forecasting framework that exploits the multiscale structure of both inflation and its predictors. The central idea, following Verona (2026), is that rather than targeting aggregate inflation directly, the series is first decomposed into frequency components. Each component is then forecast separately using models that incorporate the matching frequency components of the macroeconomic predictors, and the resulting component-level forecasts are recombined into an aggregate inflation forecast.

Let $\mathbf{x}_t^M = (x_{1t}, \ldots, x_{N_M t})'$ denote the set of macroeconomic predictors observed at time t. For each variable x_{it}, we apply a maximal overlap discrete wavelet transform (MODWT) to obtain a multiresolution representation composed of detail and smooth components at different scales. Formally, the decomposition can be written as

$$x_{it} = S_{J,t}^{(i)} + \sum_{l=1}^{J} D_{l,t}^{(i)}, \tag{5}$$

where $D_{l,t}^{(i)}$ denotes the detail component at scale l, capturing fluctuations within a specific frequency band, and $S_{J,t}^{(i)}$ represents the smooth component associated with the lowest-frequency variation of the series. The parameter J denotes the maximum decomposition level.

The wavelet coefficients are then reconstructed to generate a set of time series associated with different frequency bands. The same multiresolution decomposition is applied to the

[1] Daubechies wavelets have compact support, and a filter length of four provides a parsimonious decomposition that is commonly regarded as suitable for macroeconomic data and relatively short samples. In addition, compared with the Haar filter, it yields smoother detail and scaling components and a more refined separation across frequency bands. See Daubechies (1992) and Percival and Walden (2000).

inflation target. Let π_t denote inflation at time t. The multiresolution decomposition yields the additive representation

$$\pi_t = \sum_{j=1}^{J} \pi_t^{D_j} + \pi_t^{S_J}, \tag{6}$$

where $\pi_t^{D_j}$ denotes the detail component of inflation at scale j and $\pi_t^{S_J}$ denotes the smooth component. By construction, these components sum exactly to the original series.

For each inflation component $\pi_t^{D_j}$, the corresponding predictor vector is constructed by matching the frequency of the macroeconomic predictors to that of the inflation component being forecast. Let $\mathbf{w}_t^{D_j}$ collect the detail components $D_{j,t}^{(i)}$ for all macroeconomic variables $i = 1, \ldots, N_M$, and $\mathbf{w}_t^{S_J}$ for the smooth component. The component-specific predictor vector is then defined as

$$\mathbf{z}_t^{D_j} = \left((\mathbf{w}_t^{D_j})', (\mathbf{x}_t^E)'\right)', \tag{7}$$

where $\mathbf{x}_t^E$ denotes the set of expectation variables included in levels. This design ensures that each inflation component is modeled using macroeconomic information operating at the same frequency, while expectation variables remain unchanged. This choice isolates the contribution of the multiscale transformation to the macroeconomic predictors while preserving the original informational content of expert forecasts.

The component-specific predictor vector $\mathbf{z}_t^{D_j}$ retains the high-dimensional structure of the original dataset, as each macroeconomic variable contributes one frequency-specific component per inflation band. To effectively exploit this expanded information set, we combine the wavelet representation with forecasting methods designed for high-dimensional environments. In particular, we employ variable selection and dimension reduction techniques, capable of identifying the most informative predictors among the large set of multiscale features.

The forecasting models considered in this paper include LASSO-type regularization methods, targeted factor models, and complete subset regressions with predictor screening. These methods are well suited for environments where the number of potential regressors is large relative to the sample size, as they reduce model complexity while retaining the most relevant predictive information. By applying these techniques to the component-specific predictor set, the proposed framework allows the forecasting model to select both variables and frequency components that are most informative for predicting inflation at different horizons.

This approach provides a flexible framework for capturing heterogeneous temporal dynamics and for improving forecast performance in data-rich macroeconomic environments. Each of the J detail components and the smooth component is forecast separately using the corresponding component-specific predictor vector, and the aggregate inflation forecast is then obtained by summing all component-level forecasts:

$$\hat{\pi}_{t+h} = \sum_{j=1}^{J} \hat{\pi}_{t+h}^{D_j} + \hat{\pi}_{t+h}^{S_J}, \tag{8}$$

Overall, the proposed methodology integrates multiscale signal decomposition with modern high-dimensional forecasting tools. The wavelet transformation extracts frequency-specific information from macroeconomic variables, while the variable selection and model averaging procedures determine which components are most relevant for prediction. This combination of frequency-specific decomposition and high-dimensional estimation constitutes the core of the proposed approach.

2.2. **Forecast Design.** This section describes the forecasting design adopted in the paper. Following Garcia et al. (2017), Brazilian inflation is forecast under a direct forecasting approach, in which a separate model is estimated for each forecast horizon h using only information available at time t. Let π_{t+h} denote inflation h periods ahead and $\mathbf{x}_t$ denote the vector of predictors observed at time t. The forecasting problem can be written as

$$\pi_{t+h} = \beta_h^{'}\mathbf{x}_t + u_{t+h}, \tag{9}$$

where $\beta \in \mathbb{R}^q$ is a vector of unknown parameters, u_{t+h} is the forecasting error, and $\mathbf{x}_t = (x_{1t}, \ldots, x_{qt})' \in \mathbb{X} \subseteq \mathbb{R}^q$ may include a large number of potential covariates.

To describe how wavelet information enters the forecasting exercise, we partition the predictor vector into two blocks, $\mathbf{x}_t = \left((\mathbf{x}_t^M)', (\mathbf{x}_t^E)'\right)'$, where $\mathbf{x}_t^M$ consists of macroeconomic predictors and $\mathbf{x}_t^E$ contains the expectation variables. This partition is consistent with the high-dimensional dataset of Garcia et al. (2017), which combines 59 macroeconomic variables with 34 variables linked to specialist forecasts.

In the wavelet specification, the multiresolution decomposition is applied only to the macroeconomic variables $\mathbf{x}_t^M$, whereas the expectation variables $\mathbf{x}_t^E$ are kept in levels. This choice is motivated by the role of expert forecasts in the original forecasting environment. Garcia et al. (2017) show that specialist forecasts are important candidate regressors at short horizons especially for $h = 1$ and $h = 2$, while their predictive power fades as the horizon increases.

In the frequency-specific forecasting framework, a separate model is estimated for each component of inflation. For each detail component $\pi_t^{D_j}$ and the smooth component $\pi_t^{S_J}$, the direct forecasting equation is written as

$$\pi_{t+h}^{D_j} = \beta^{D_j}{}'_h \mathbf{z}_t^{D_j} + u_{t+h}^{D_j}, \tag{10}$$

where $\mathbf{z}_t^{D_j}$ is the component-specific predictor vector defined in Section 2.1, $\beta_h^{D_j}{}' \in \mathbb{R}^q$ is a vector of unknown parameters, and $u_{t+h}^{D_j}$ is the forecasting error. An analogous equation is estimated for the smooth component,

$$\pi_{t+h}^{S_J} = \beta^{S_J}{}'_h \mathbf{z}_t^{S_J} + u_{t+h}^{S_J}, \tag{11}$$

where $\mathbf{z}_t^{S_J} = \left((\mathbf{w}_t^{S_J})', (\mathbf{x}_t^E)'\right)'$ collects the smooth components of the macroeconomic variables and the expectation variables in levels.

In addition to the component-specific forecasting framework described above, the empirical implementation adopts a wavelet-based multiscale selection method, denoted by WMS. Following Verona (2026), the final aggregate forecast is not imposed mechanically from a fixed multiscale specification. The motivation is that not all frequency components contribute equally to forecast accuracy at every horizon: high-frequency components tend to carry predictive information primarily at short horizons, whereas low-frequency components tend to become more relevant at longer horizons. Including all components indiscriminately may therefore introduce noise and reduce forecast accuracy. For each forecast horizon, component-specific forecasts are first generated from the candidate high-dimensional models. The WMS then selects, from a predefined set of admissible multiscale aggregations, the specification that minimizes the out-of-sample RMSE. The selected aggregation is then taken as the final aggregate inflation forecast.

2.3. **Model specifications.** All forecasting methods described below are estimated separately for each inflation component using the corresponding component-specific predictor vector defined in Section 2.2. The aggregate inflation forecast is then recovered through the multiscale aggregation rule described above.

2.3.1. *Factor model with targeted predictors.* Factor models offer a parsimonious way of handling a large set of predictors in forecasting problems. Instead of including all candidate variables directly in the forecasting equation, the information is summarized by a small number of latent factors estimated by principal components. Consider the following forecasting model

$$\pi_{t+h} = \sum_{i=1}^{p} \gamma_i' \mathbf{f}_{t-i} + u_{t+h}, \tag{12}$$

where $\mathbf{f}_t$ is a $k \times 1$ vector of common factors extracted from the predictor set, with $k \ll q$.

Following Bai and Ng (2008), the forecasting performance of factor models may improve when the factors are extracted not from the full set of regressors, but from a subset of variables selected according to their predictive content for the corresponding inflation component. This screening step is intended to improve the signal-to-noise ratio in factor estimation and to reduce the influence of noisy variables with weak forecasting ability. In this sense, target factors are standard factor models preceded by a pre-testing step that retains only the most relevant variables for factor extraction.

We follow the implementation in Medeiros and Vasconcelos (2016). In the pre-testing step, we include four lags of the corresponding inflation component as fixed controls and evaluate each candidate predictor individually. For each candidate, we include its current value and three additional lags, and assess its predictive relevance through the associated t-statistics. The variables are then ranked by the absolute value of these statistics, and statistical significance is used as the selection criterion. Principal components are extracted from the selected predictors, and the number of factors included in the final specification is chosen by BIC.

2.3.2. *LASSO and Flex-adaLASSO.* High-dimensional forecasting environments, in which the number of candidate predictors q is large relative to the sample size T, require regularization methods that simultaneously perform variable selection and parameter estimation. The LASSO, introduced by Tibshirani (1996), addresses this issue by adding a L_1 penalty to the least squares objective, which shrinks some coefficients exactly to zero and produces sparse solutions. For each forecast horizon h, the LASSO estimator solves

$$\hat{\beta} = \arg\min_{\beta} \left[\sum_{t=1}^{T} (\pi_{t+h} - \beta' \mathbf{x}_t)^2 + \lambda \sum_{j=1}^{q} |\beta_j| \right] \tag{13}$$

where $\lambda \geq 0$ controls the degree of penalization. We select λ by minimizing the BIC over the regularization path, following Medeiros and Vasconcelos (2016).

The Adaptive LASSO, proposed by Zou (2006), allows the penalty to vary across coefficients. It uses weights constructed from a first-step estimator, usually the LASSO, to reflect the relative importance of the regressors. The adaLASSO is defined as:

$$\hat{\beta} = \arg\min_{\beta} \left[\sum_{t=1}^{T} (\pi_{t+h} - \beta' \mathbf{x}_t)^2 + \lambda \sum_{j=1}^{q} \gamma_j |\beta_j| \right], \tag{14}$$

where $\gamma_j = |\hat{\beta}_j^{\,*}|^{-\tau}$ are adaptive weights constructed from a first-stage LASSO estimator. We implement the adaLASSO in which the exponent τ is treated as a tuning parameter, selected over a grid. The optimal τ is chosen as the value that minimizes the BIC across all models in the grid. This specification, referred to as Flex-adaLASSO, was adopted by Medeiros and Vasconcelos (2016), who note that selecting τ by BIC may reduce the forecasting errors. We also estimate a post-selection ordinary least squares regression (P. OLS), using the predictors selected by the first-stage LASSO estimator in each component-specific equation. After variable selection, re-estimating the model by OLS may reduce shrinkage bias.

2.3.3. *CSR with targeted predictors.* The Complete Subset Regression (CSR), introduced by Elliott et al. (2013), implements an alternative approach to handle model uncertainty in environments with a large number of candidate predictors by estimating all possible combinations of regressors over a subset of fixed size k of K candidate predictors, and combining their forecasts for the corresponding inflation component by taking the average across all regressions in the subset.

We follow the implementation of Garcia et al. (2017), in which the CSR is used together with a targeted pre-testing step. In each estimation window, the corresponding inflation component is regressed on each candidate predictor individually, including its lagged values, and the candidate variables are ranked according to the absolute value of the associated t-statistics. The complete subset regressions are then estimated only on the predictors that are most relevant according to this ranking. This pre-selection is necessary because, even when the subsets are small, the number of possible regressions grows very quickly with the dimension of the candidate set.

We set the targeted subset size to 25 predictors and estimate all combinations with subsets of size 4, so that the final CSR forecast is obtained as the average forecast across all regressions in the restricted model space.

3. Data and Results

In this section, we investigate forecasting methods using a wavelet-based multiscale approach and evaluate their predictive performance for monthly inflation in Brazil, measured by the Brazilian consumer price index (IPCA), the official inflation index in Brazil.

3.1. **Data.** We used the high-dimensional dataset proposed by Garcia et al. (2017). The dataset was obtained from Bloomberg and the Central Bank of Brazil, covering the period from January 2003 to December 2015. It contains 156 monthly observations across 93 Brazilian economic series, and consists of 59 macroeconomic variables and 34 variables linked to specialist forecasts (Garcia et al., 2017). The macroeconomic variables include various inflation and economic activity indices in Brazil, such as measures of labor activity, energy consumption, exchange rates, stock markets, government accounts, and monetary variables. In addition, the dataset includes expert forecasts from the FOCUS survey produced by the Central Bank of Brazil, which aggregates expectation variables, such as the median of the h-period-ahead specialist forecasts, the median of the Top 5 experts (Top5), i.e., the five experts who produced the best forecasts in the previous period, as well as the mean and standard deviation of these Top 5 experts. Forecasts are produced for horizons $h = 1, \ldots, 12$, where $h = 1$ corresponds to a forecast made five days before the IPCA release, $h = 2$ to one month and five days ahead, and $h = 12$ to 11 months and five days ahead. All of the variables included in the models are listed in Table 1. The models are estimated using a nine-year rolling window, with the initial out-of-sample forecast generated for January 2012. Due to the direct forecasting framework, the estimation sample size decreases as the horizon increases, utilizing 108 observations for $h = 1$, 107 for $h = 2$ and decreasing for subsequent horizons. The models are evaluated based on 48 monthly forecast points for each horizon. We include a random walk and an autoregressive model with lag length selection determined by BIC as benchmarks.

No.	Variable	No.	Variable	No.	Variable
1	Brazil CPI IPCA MoM	32	Brazil Monetary Base	63	$t+4$ median
2	FGV Brazil General Prices IGP-	33	Brazil Money Supply M1 Brazil	64	$t+5$ median
3	FGV Brazil General Prices IGP-DI	34	Brazil Money Supply M2 Brazil	65	$t+6$ median
4	FGV Brazil General Prices IGP-M	35	Brazil Money Supply M3 Brazil	66	$t+7$ median
5	Brazil CPI IPCA-15	36	Brazil Money Supply M4 Brazil	67	$t+8$ median
6	Brazil CPI IPCA Median Market Expectations	37	Brazil National Treasury Revenue Total	68	$t+9$ median
7	Brazil Total Electricity Consumption	38	Brazil Social Contribution over Net Profit Tax Income	69	$t+10$ median
8	Brazil Industrial Electricity Consumption	39	Brazil PIS & PASEP Tax Income	70	$t+11$ median
9	BoA Merrill Lynch Economic Conditions Index	40	Brazil Central Government Net Revenue	71	$t+12$ median
10	CNI Brazil Manufacture Industrial Confidence Index	41	Brazil Central Government Revenue from the Central Bank	72	$t+13$ median
11	CNI Brazil Manufacture Industrial Installed Capacity	42	Brazil Central Government Total Expenditures	73	Top5 $t+1$ median
12	Brazil Industrial Production Seasonally Adjusted	43	Brazil National Treasury Gross Revenue	74	Top5 $t+2$ median
13	Brazil Industrial Production Annual	44	Brazil Importing Tax Income	75	Top5 $t+3$ median
14	IBGE Brazil Unemployment Rate	45	BNDES Brazil Income Taxes	76	Top5 $t+4$ median
15	Brazil Unemployment Statistic Male	46	Brazil National Treasury Revenue from Industrialized Products	77	Top5 $t+5$ median
16	Brazil Unemployment Statistic Total	47	Brazil National Treasury Revenues from Other Taxes	78	Top5 $t+6$ median
17	IMF Brazil Unemployment Rate in Percentage	48	Brazil Central Government Revenue from the Social Security	79	Top5 $t+7$ median
18	CNI Brazil Manufacturing Industry Employment	49	Brazil National Treasury Revenue from Import Tax	80	Top5 $t+8$ median
19	Brazil Unemployment Statistic Female	50	Brazil Current Account	81	Top5 $t+9$ median
20	Brazil Industry Working Hours	51	Brazil Trade Balance FOB	82	Top5 $t+10$ median
21	Brazil Average Real Income	52	Brazil Public Net Fiscal Debt as a Percentage of GDP	83	Top5 $t+11$ median
22	Brazil Minimum Wage	53	Brazil Public Net Fiscal Debt	84	Top5 $t+12$ median
23	USD-BRL X-Rate	54	Brazilian Federal Government Domestic Debt	85	Top5 $t+13$ median
24	USD-BRL X-Rate Tourism	55	Brazil Public Net Government & Central Bank Domestic Debt	86	$t+1$ median^2
25	EUR-BRL X-Rate	56	Brazilian States Debt Total Consolidated Net Debt	87	$t+1$ mean
26	Brazil IBOVESPA Index	57	Brazilian States Debt to Foreigners	88	$t+1$ mean^2
27	Brazil Savings Accounts Deposits	58	Brazilian Cities Debt	89	$t+1$ Std
28	Brazil Total Savings Deposits	59	Brazilian Cities Debt to Foreigners	90	$t+12$ median^2
29	Brazil BNDES Long Term Interest Rate	60	$t+1$ median	91	$t+2$ mean
30	Brazil Selic Target Rate	61	$t+2$ median	92	$t+2$ mean^2
31	Brazil Cetip DI Interbank Deposits	62	$t+3$ median	93	$t+2$ Std

TABLE 1. Macroeconomic variables and Focus expectation variables.

3.2. **Empirical Results.** The empirical analysis evaluates the forecasting performance of the proposed wavelet-based multiscale selection (WMS) method relative to a broad set of benchmark models commonly used in the inflation forecasting literature. These include simple univariate specifications, such as the random walk (RW) and autoregressive (AR) models, as well as high-dimensional and shrinkage-based approaches, including factor models, LASSO, adaptive LASSO, post-selection OLS, and complete subset regression (CSR). In addition, we consider survey-based expectations (FOCUS) and a forecast combination based on the median of the top five experts (Top5). Forecasts are generated for horizons ranging from one to twelve months ahead, and performance is assessed using relative root mean squared errors (RMSE), computed as the ratio of the RMSE of the WMS model to that of each competing alternative. Values below one indicate superior predictive accuracy of the proposed method. To evaluate the statistical significance of the differences in predictive performance, we employ the Diebold–Mariano test with the Harvey–Leybourne–Newbold small-sample correction (DM-HLN). The analysis also includes accumulated inflation forecasts over an 11-month horizon, providing an additional assessment of medium-term predictive performance.

The empirical results, reported in Table 2, indicate that the proposed wavelet-based multiscale selection (WMS) method delivers substantial improvements in forecasting accuracy at

Forecast Horizon	RW	AR	FOCUS	Top5	Factors	LASSO	F. aL	P. OLS	CSR
$t+1$	0.20***	0.21***	0.51***	0.50***	0.37***	0.51***	0.49***	0.48***	0.46***
$t+2$	0.53***	0.59***	0.94	1.02	0.95	0.92	1.02	0.96	1.06*
$t+3$	0.57***	0.69**	0.88	0.91	1.07	0.90	1.04	1.03	1.03
$t+4$	0.49***	0.60**	0.81	0.81	0.94	0.80	0.86	0.83	0.91
$t+5$	0.51***	0.69**	0.89	0.85	1.00	0.86	0.92	0.90	0.99
$t+6$	0.49***	0.72**	0.89	0.84*	0.96	0.87*	0.97	0.93	0.99
$t+7$	0.49***	0.78*	0.90	0.83	0.93	0.90	0.98	0.98	1.01
$t+8$	0.54***	0.81**	0.92	0.87	0.91	0.87**	0.97	0.96	1.02
$t+9$	0.62***	0.81***	0.91	0.92	0.89	0.89**	0.94	0.94	1.02
$t+10$	0.69***	0.83**	0.91	0.88	0.97	0.91	0.98	1.00	1.04
$t+11$	0.78**	0.85	0.92	0.93	0.98	0.91	0.96	0.98	1.05
$t+12$	0.89	0.86	0.94	0.96	1.02	0.92	1.00	1.02	1.07
Acc.	0.36***	0.57	0.72	0.73	0.89	0.78	0.95	0.95	1.02

Table 2. Relative forecast accuracy of the wavelet-based multiscale selection method (WMS) against competing models at horizons $h = 1, \ldots, 12$. The entries are relative RMSEs, computed as the ratio of the RMSE of WMS to the RMSE of each competing model. The competing models are: random walk (RW); autoregressive model (AR); factor model (Factors); LASSO (LASSO); Flex-adaLASSO (F. aL); post-selection OLS (P. OLS); complete subset regression (CSR); survey-based expectations (FOCUS); and the median of the top five experts (Top5). Values below one indicate that the WMS outperforms the competing model. The column Acc. reports the relative forecast errors of the 11-month accumulated inflation forecasts. Asterisks denote statistical significance of the DM-HLN test at the 10% (*), 5% (**), and 1% (***) levels.

Forecast Horizon	Components	D1 model	S1 model
$t+1$	$D_1 + S_1$	Factors	Factors
$t+2$	$D_1 + S_1$	Factors	CSR
$t+3$	S_1	—	CSR
$t+4$	$D_1 + S_1$	LASSO	LASSO
$t+5$	$D_1 + S_1$	Factors	CSR
$t+6$	$D_1 + S_1$	Factors	CSR
$t+7$	$D_1 + S_1$	Factors	CSR
$t+8$	$D_1 + S_1$	F. aL	CSR
$t+9$	S_1	—	CSR
$t+10$	$D_1 + S_1$	Factors	CSR
$t+11$	S_1	—	CSR
$t+12$	S_1	—	CSR

Table 3. Component selection of the wavelet-based multiscale selection (WMS) method at each forecast horizon. The column Components indicates which frequency components are included in the aggregate forecast: $D_1 + S_1$ denotes that both the detail and smooth components are retained, while S_1 indicates that only the smooth component is used and the high-frequency detail D_1 is excluded. The columns D1 model and S1 model report the forecasting model assigned to each component by the optimization procedure. A dash (—) indicates that the component was excluded.

short horizons. Against the random walk, the WMS achieves relative RMSEs between 0.20 and 0.89 with statistically significant gains at the 1% level through $h = 10$. Improvements over the autoregressive model are significant through $h = 10$ but diminish in the later horizons. The WMS generally outperforms survey-based expectations. At $h = 1$, the RMSE is reduced by approximately 49% relative to the FOCUS median and 50% relative to the Top5 consensus, both significant at the 1% level. At medium horizons ($h = 3$ through $h = 8$), relative RMSEs against both survey benchmarks range between 0.81 and 0.92. These gains, while economically meaningful, are not individually significant in most cases.

The comparison with high-dimensional models reveals a horizon-dependent pattern. At $h = 1$, the WMS delivers statistically significant gains against all competing specifications, with relative RMSEs below 0.52 for every benchmark. In absolute terms, the WMS achieves an RMSE of 0.48 at $h = 1$, compared with 0.95 for the LASSO, 0.95 for the FOCUS median, and 2.41 for the random walk. The magnitude of the gains at $h = 1$ reflects the informational advantage of separating transitory from persistent dynamics at the shortest horizon. At this horizon, the detail component D_1 isolates high-frequency fluctuations coinciding with monthly price adjustments, while the smooth component S_1 filters the underlying inflation trend. This decomposition enables the model to extract frequency-specific predictive signals, leveraging the predictive power of expectations at the short horizon. At medium horizons, the WMS delivers consistent improvements against the LASSO, and significant reductions in horizons 6, 8 and 9. Against the Flex-adaLASSO and Post-OLS, relative errors remain below unity at most horizons, though without statistical significance. Compared to the CSR model, the WMS dominates at $h = 1$, but the two methods produce comparable forecasts from $h = 5$ onward, with the CSR marginally outperforming at $h = 8$ through $h = 12$. This pattern likely reflects a limitation of the single decomposition level at longer horizons. At longer horizons, the selected specification more often excludes the high-frequency detail component, so that the forecast is increasingly driven by the smooth component, which is always retained by construction.

Table 3 reports the component selection of the WMS at each horizon and provides a frequency-specific interpretation of the gains. The smooth component S_1 is assigned to the CSR in 10 of 12 horizons. The detail component D_1 is retained in 8 of 12 horizons and assigned to the targeted factor model in 6 of those cases, indicating that dimensionality reduction through principal components is better suited for the noisier high-frequency component. The WMS excludes D_1 at $h = 3$, $h = 9$, $h = 11$, and $h = 12$, indicating that the high-frequency component introduces noise rather than predictive information at these horizons. This selective exclusion is consistent with evidence showing that the predictive content of frequency components is horizon-dependent, with higher frequency components being less informative at medium and long horizons (Verona, 2026; Faria and Verona, 2018), and that the relative importance of wavelet details varies over time (Risse, 2019).

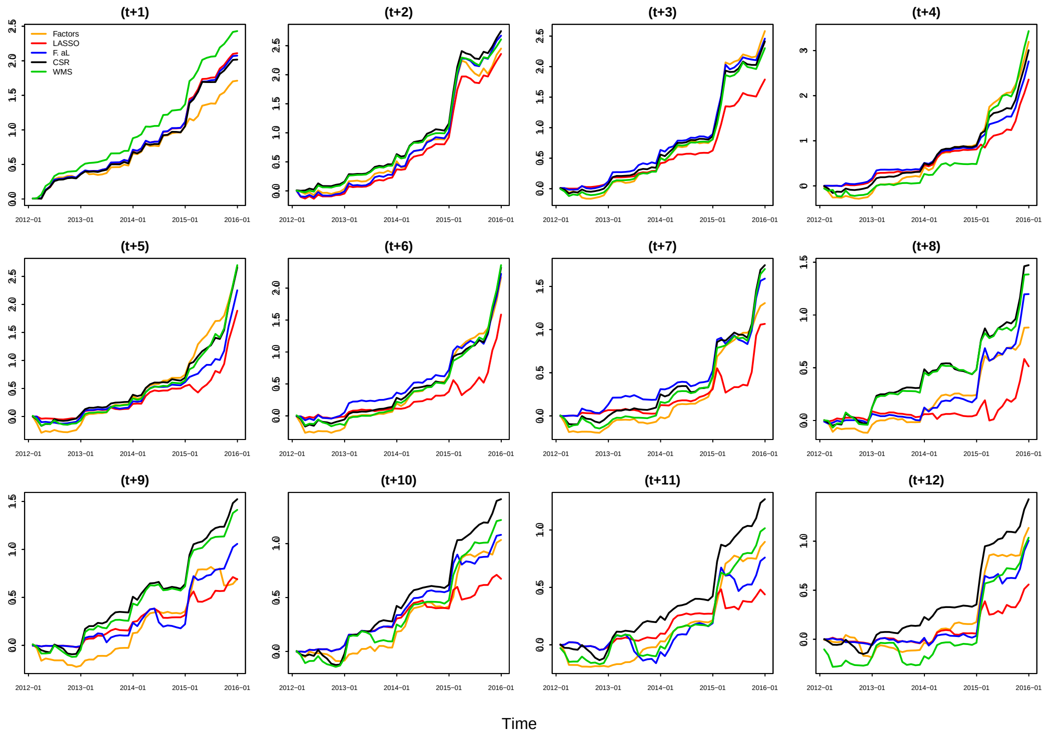

FIGURE 1. Cumulative squared forecast error differences relative to the AR benchmark over the out-of-sample evaluation period, for forecast horizons $h = 1, \ldots, 12$. Each panel corresponds to a specific forecast horizon and reports the time evolution of the cumulative difference in squared forecast errors between the AR model and each competing specification: factor model (Factors); LASSO, Flex-adaLASSO (F. aL), complete subset regression (CSR), and wavelet-based multiscale selection method (WMS). Positive values indicate that the competing model outperforms the AR benchmark. The WMS method exhibits the largest cumulative gains at short horizons, particularly at $h = 1$. At medium and longer horizons, the gains of the WMS converge toward those of the CSR, consistent with the dominance of the smooth component in the selection at these horizons. The figure highlights that the forecasting improvements of the WMS are persistent over time and not driven by isolated episodes.

Figure 1 complements these findings by plotting cumulative squared forecast error differences relative to the AR benchmark. At $h = 1$, the WMS accumulates a persistent advantage over all competing models throughout the evaluation period, confirming that the gains are systematic and not driven by isolated episodes. At medium horizons, the WMS remains among the top-performing models, though its cumulative advantage narrows relative to the CSR and LASSO. At longer horizons ($h = 9$ through $h = 12$), the cumulative error paths of the WMS and the CSR converge, consistent with the relative RMSEs reported in Table 2.

Replacing the Daubechies filter with the Haar filter with $J = 1$ decomposition yields broadly consistent results, with similar patterns at longer horizons. Furthermore, we found that increasing the decomposition level with the Daubechies filter deteriorates accuracy across all horizons. This result may reflect the constraints imposed by the sample size. Higher decomposition levels increased model complexity and reduced the effective signal-to-noise ratio in the component-specific regressions, without delivering offsetting gains in frequency resolution. The $J = 1$ specification therefore provides the most favorable trade-off between frequency resolution and estimation precision given the sample size.

Taken together, the evidence from pointwise accuracy measures, component selection patterns, and cumulative forecast errors provides a coherent picture of the WMS method's contribution. The frequency-specific decomposition of inflation, combined with high-dimensional models matched to each frequency band, delivers its strongest gains at short horizons, where the separation between transitory and persistent dynamics is most informative. The method also identifies which components to retain and which models to assign to each frequency band. At longer horizons, gains diminish as the smooth component dominates, a limitation inherent to the $J = 1$ decomposition.

4. Conclusion

This paper investigates the role of multiscale information in macroeconomic forecasting by combining wavelet-based decompositions with modern high-dimensional estimation techniques. The proposed framework decomposes inflation into frequency components, matches each component with predictors observed at the corresponding frequency, and estimates component-specific forecasting models in a high-dimensional environment. In this way, the approach allows the forecasting model to exploit heterogeneous predictive information across frequencies while preserving a parsimonious structure.

The empirical results indicate that this frequency-specific approach is particularly useful at short horizons. The wavelet-based multiscale selection method (WMS) delivers its largest gains at the first horizon, where it substantially improves upon standard time-series benchmarks, survey-based expectations, and the competing high-dimensional models. More broadly, the WMS consistently outperforms the random walk and autoregressive benchmarks across nearly all horizons, and generally improves upon survey-based forecasts, though these gains lose statistical significance at the longest horizons.

The component selection reveals that this horizon dependence is driven by the selective inclusion of the detail component, which is retained at short and medium horizons but excluded at longer horizons where it may introduce noise rather than predictive content. This indicates that the contribution of multiscale information does not come from mechanically combining all frequency bands, but from allowing the relevance of each component to vary with the forecasting horizon.

The results also highlight the complementarity between different estimation strategies across frequency bands. The targeted factor model is most frequently selected for the high-frequency detail component, while the CSR dominates the smooth component across horizons. This specialization suggests that the gains from the WMS arise not from a single superior model, but from the ability to assign different estimation strategies to different frequency bands according to their predictive content.

The current specification uses a single decomposition level, which provides only two frequency bands. While this is sufficient to generate substantial gains at short horizons, it restricts the scope for selective component exclusion at longer horizons. Increasing the decomposition level did not improve accuracy in the current sample, suggesting that the $J = 1$ specification offers the most favorable trade-off given the available sample size.

From a broader perspective, the paper contributes to the growing literature on data-rich macroeconomic forecasting by showing that improvements do not necessarily require additional variables, but rather more informative representations of existing data. By explicitly accounting for the multiscale nature of economic dynamics, the proposed approach provides a systematic way to extract and exploit information that is otherwise hidden in the time domain.

Several avenues for future research remain. First, the framework could be extended to multivariate settings, including joint modeling of inflation and real activity indicators. Second, integrating wavelet-based representations with nonlinear and machine learning models, such as deep neural networks or transformers, may further enhance predictive performance. Third, exploring real-time implementations and robustness to data revisions would be particularly relevant for policy applications. Finally, another direction concerns the integration of wavelet denoising techniques into multiscale forecasting frameworks. These methods operate at the coefficient level and could provide a finer mechanism for separating signal from noise within each frequency band.

Overall, the findings of this paper suggest that multiscale methods represent a promising and interpretable tool for macroeconomic forecasting, with clear benefits for both empirical research and policy analysis.

References

Alexandridis, A. K., E. Panopoulou, and I. Souropanis (2024). Forecasting exchange rate volatility: An amalgamation approach. *Journal of International Financial Markets, Institutions and Money 97*, 102067.

Bai, J. and S. Ng (2008). Forecasting economic time series using targeted predictors. *Journal of Econometrics 146*(2), 304–317.

Candès, E. J. and M. B. Wakin (2008). An introduction to compressive sampling. *IEEE Signal Processing Magazine 25*(2), 21–30.

Crowley, P. M. (2007). A guide to wavelets for economists. *Journal of Economic Surveys 21*(2), 207–267.

Daubechies, I. (1992). *Ten Lectures on Wavelets*. Society for Industrial and Applied Mathematics.

Donoho, D. L. (1995). De-noising by soft-thresholding. *IEEE Transactions on Information Theory 41*(3), 613–627.

Elliott, G., A. Gargano, and A. Timmermann (2013). Complete subset regressions. *Journal of Econometrics 177*(2), 357–373. Dynamic Econometric Modeling and Forecasting.

Faria, G. and F. Verona (2018). Forecasting stock market returns by summing the frequency-decomposed parts. *Journal of Empirical Finance 45*, 228–242.

Fernandez, V. (2008). Traditional versus novel forecasting techniques: how much do we gain? *Journal of Forecasting 27*(7), 637–648.

Garcia, M. G., M. C. Medeiros, and G. F. Vasconcelos (2017). Real-time inflation forecasting with high-dimensional models: The case of brazil. *International Journal of Forecasting 33*(3), 679–693.

Gençay, R., F. Selçuk, and B. Whitcher (2002). *An Introduction to Wavelets and Other Filtering Methods in Finance and Economics*. Academic Press.

Mallat, S. (1998). *A Wavelet Tour of Signal Processing*. Academic Press.

Medeiros, M. C. and G. F. Vasconcelos (2016). Forecasting macroeconomic variables in data-rich environments. *Economics Letters 138*, 50–52.

Meyer, Y. (1992). *Wavelets and Operators*. Cambridge University Press.

Percival, D. and A. Walden (2000). *Wavelet Methods for Time Series Analysis*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press.

Risse, M. (2019). Combining wavelet decomposition with machine learning to forecast gold returns. *International Journal of Forecasting 35*(2), 601–615.

Rua, A. (2011). A wavelet approach for factor-augmented forecasting. *Journal of Forecasting 30*(7), 666–678.

Rua, A. (2017). A wavelet-based multivariate multiscale approach for forecasting. *International Journal of Forecasting 33*(3), 581–590.

Sengupta, S., T. Chakraborty, and S. K. Singh (2025). Forecasting cpi inflation under economic policy and geopolitical uncertainties. *International Journal of Forecasting 41*(3), 953–981.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological) 58*(1), 267–288.

Verona, F. (2026). Forecasting inflation: The sum of the cycles outperforms the whole. Bank of Finland Research Discussion Papers 1/2026, Bank of Finland, Helsinki.

Zou, H. (2006). The adaptive lasso and its oracle properties. *Journal of the American Statistical Association 101*(476), 1418–1429.