
# Forecasting UK Consumer Price Inflation using Inflation Forecasts

Hossein Hassani, Emmanuel Sirimal Silva

PII: S1090-9443(18)30182-0
DOI: https://doi.org/10.1016/j.rie.2018.07.001
Reference: YREEC 765

To appear in: *Research in Economics*

Received date: 4 June 2018
Accepted date: 2 July 2018

# Forecasting UK Consumer Price Inflation using Inflation Forecasts

Hossein Hassani[1,*] and Emmanuel Sirimal Silva[2]

[1] *Research Institute of Energy Management and Planning, University of Tehran, Iran*

[2] *Fashion Business School, London College of Fashion, University of the Arts London, UK*

[*] Corresponding Author: *hassani.stat@gmail.com*

August 1, 2018

**Abstract**

The inflation rate is a key economic indicator for which forecasters are constantly seeking to improve the accuracy of predictions, so as to enable better macroeconomic decision making. Presented in this paper is a novel approach which seeks to exploit auxiliary information contained within inflation forecasts for developing a new and improved forecast for inflation by modelling with Multivariate Singular Spectrum Analysis (MSSA). Unlike other forecast combination techniques, the key feature of the proposed approach is its use of forecasts, i.e. data into the future, within the modelling process and extracting auxiliary information for generating a new and improved forecast. We consider real data on consumer price inflation in UK, obtained via the Office for National Statistics. A variety of parametric and nonparametric models are then used to generate univariate forecasts of inflation. Thereafter, the best univariate forecast is considered as auxiliary information within the MSSA model alongside historical data for UK consumer price inflation, and a new multivariate forecast is generated. We find compelling evidence which shows the benefits of the proposed approach at generating more accurate medium to long term inflation forecasts for UK in relation to the competing models. Finally, through the discussion, we also consider Google Trends forecasts for inflation within the proposed framework.

**Keywords:** Consumer price inflation; Auxiliary information; Multivariate Singular Spectrum Analysis; Parametric; Nonparametric; Forecast.



## 1 Introduction

Inflation forecasting has a wide and long ranging history. Especially in countries like UK, where inflation targeting has prevailed since 1992, causing policy makers to anticipate inflation [1]. Regardless of its importance over a long time period, inflation forecasting remains a key challenge for Central Banks [2]. As such, there remains an interest in developing comparatively more effective approaches for improving the accuracy of inflation forecasts.

To this end, there has been considerable interest in forecast combination approaches as viable tools for inflation forecasting. Whilst it is beyond the mandate of this paper to review all such efforts, we find it pertinent to refer the reader to few recent and successful attempts at exploiting forecast combination techniques within the field of inflation forecasting. Uncertainty is rife in the current economic climate, and forecast combination methods can be used to hedge against bad forecast performance of single models during such times of crisis [3]. A popular example of forecast combination in the literature is through Dynamic Model Average (DMA). The DMA technique was exploited in [4] for forecasting US inflation. Here, the authors found the DMA technique to be significantly better than benchmark regression models and selected time varying coefficient models. In [2], both the DMA technique and Dynamic Model Selection (DMS) were used to improve inflation forecasts for USA and the Euro area. In contrast, the Bayesian Model Averaging (BMA) technique is yet another popular form of forecast combination and was utilized in [5] for forecasting US inflation forecasting. The BMA model was seen outperforming simple equalweighted averaging. Those interested in more varied applications of forecast combination techniques are referred to [6–8] and the references therein.

Nevertheless, a review of the relevant literature points out towards the existence of a strong heterogeneity in the predictive performances of the models used for inflation forecasting [2]. Thus, our interest delves into the search for a new and improved approach for improving the accuracy of inflation forecasts. At the outset, it is noteworthy that forecast combination techniques extract multiple forecasts from various models and then combine same using some weights. For example, the DMA technique allows several models to hold at given points in time and then averages their respective forecasts [2]. In contrast, the proposed method takes an entirely different approach and presents an alternative form of forecast combination for time series analysis and forecasting. We hope the difference will be clearer to the reader in the explanations which follow and in Section 2.

In brief, our interest lies in determining the possibility and effectiveness of exploiting auxiliary information contained within inflation forecasts. We seek to combine historical data for UK consumer price inflation with a forecast for consumer price inflation using a multivariate tool which can extract the signals contained with the given forecast and use this new information to generate a more improve forecast for UK consumer price inflation. The development of multivariate models for inflation forecasting is warranted as studies have evidenced the difficulties in improving inflation forecasts with simple univariate models alone [2, 9].

Accordingly, we begin by forecasting the UK consumer price inflation series at different horizons using popular and powerful univariate time series analysis and forecasting models such as ARIMA, Exponential Smoothing (ETS), Neural Networks (NN), and Trigonometric Box-Cox ARMA Trend Seasonal (TBATS). The choice of univariate models is motivated by its previous applications in most cases. For example, it is well known that ARIMA models have been a popular contender for inflation forecasting both historically and in the recent past [10–13]. The ETS technique too has been utilised effectively in inflation forecasting studies [13–15] whilst applications of NN based approaches are documented in [16–19]. However, to the best of our knowledge, the TBATS model [20] has not been evaluated in the context of inflation forecasting, and thus, it will be interesting to see how this performs. Once the univariate forecasts are computed, we propose using the best performing univariate forecast as auxiliary information, coupled with historical UK consumer price inflation data within a multivariate framework for producing a new and improved CPI forecast. It is noteworthy that the best performing univariate forecast represents data into the future, and therefore we require a multivariate tool that can model time series with different series lengths to further the objective of this study.

We opt for the nonparametric Multivariate Singular Spectrum Analysis (MSSA) [21] technique as a tool which can suitably fit the requirements of the proposed methodology. This is because: first, the Vertical MSSA (VMSSA) approach can model time series with different series lengths [22]; secondly, MSSA is a filtering and signal extraction technique which has the capability of extracting any auxiliary information [21] contained within a given forecast. MSSA is now experiencing a surge in applications in a variety of fields, and those interested are referred to the work in [23–26] and references therein.

The remainder of this paper is organised as follows. Section 2 presents the various forecasting methods and the MSSA-based methodology for forecasting UK consumer price inflation. Section 3 introduces the data, the forecasting exercise and the metrics used for evaluating forecast accuracy. The results are presented in Section 4 and the paper concludes in Section 5.

# 2 Methodology

## 2.1 Autoregressive Integrated Moving Average (ARIMA)

The ARIMA model used here for generating univariate forecasts for UK consumer price inflation is commonly referred to as 'auto.arima' and can be accessed via the forecast package in $R$. Those interested in a detailed description of this optimized algorithm are referred to [27]. In brief, the modeling process begins by repeating KPSS [29] tests to determine the number of differences $d$. Thereafter, the Akaike Information Criterion (AIC) is minimized following differencing $d$ times to determine the values of $p$ (number of autoregressive terms) and $q$ (number of lagged forecast errors in the forecasting equation). Instead of considering every possible combination of $p$ and $q$, the algorithm opts to traverse the model space via a stepwise search. It is noteworthy that the algorithm relies on a corrected version of the AIC (AIC$_c$) as indicated below:

$$AIC = -2log(L) + 2(p + q + P + Q + k), \tag{1}$$

$$AIC_c = AIC + \frac{2(p + q + k + 1)(p + q + k + 2)}{T - p - q - k - 2}. \tag{2}$$

where $k = 1$ if c $\neq 0$ and 0 otherwise, $L$ is the maximum likelihood of the data and the last term in parentheses is the number of parameters in the model (this includes $\sigma^2$ which is the variance of the residuals).

Thereafter, the algorithm searches the following four ARIMA models, to achieve what is called the 'current model': ARIMA*(2,d,2)*, ARIMA*(0,d,0)*, ARIMA*(1,d,0)* and ARIMA*(0,d,1)* for the model which minimises the AIC$_c$. If $d = 0$ then the constant $c$ is included; if $d \geq 1$, then the constant $c$ is set to zero. In addition, the model also evaluates variations on the current model by varying $p$ and $q$ by +/-1 and including/excluding $c$. The steps following on from the minimisation of the AIC are repeated until no lower AIC$_c$ can be found.

As seen in the next section, we consider monthly UK consumer price inflation data, and therefore we provide a brief expansion of the seasonal ARIMA model alone. In doing so we mainly follow [27]. The seasonal ARIMA model can be expressed as:

$$\Phi(B^m)\phi(B)(1-B^m)^D(1-B)^d y_t = c + \Theta(B^m)\theta(B)\epsilon_t, \tag{3}$$

where $\Phi(z)$ and $\Theta(z)$ are the polynomials of orders $P$ and $Q$, and $\epsilon_t$ is white noise. If, $c \neq 0$, there is an implied polynomial of order $d + D$ in the forecast function.

As explained in [28] point forecasts can then be obtained as follows. Begin by expanding the seasonal ARIMA equation so that $y_t$ is on the left hand side with all other terms on the right. Then, rewrite the ARIMA equation and replace $t$ with $T + h$ and finally, on the right hand side of this equation replace future observations by their forecasts, future errors by zero, and past errors by the corresponding residuals. Eventually, use the forecasting horizon $h = 1$ month ahead for example to calculate all forecasts for that horizon.

## 2.2 Exponential Smoothing (ETS)

The ETS model from the forecast package in *R* is automated to consider the error, trend and seasonal components in choosing the best exponential smoothing model from 30 possible options by optimizing initial values and parameters using the Maximum Likelihood Estimator and selecting the best model based on the AIC. This ETS algorithm overcomes limitations from the previous models of exponential smoothing which failed to provide a method for easily calculating prediction intervals [31]. The algorithms that generate point forecasts for UK consumer price inflation[1] and state space equations for each of the models in the ETS framework[2] can be found in [28] where a more detailed description of ETS is available.

## 2.3 Neural Networks (NN)

An automatic NN forecasting model known as nnetar which is provided through the forecast package in *R* is used for UK consumer price inflation forecasting

[1] https://www.otexts.org/sites/default/files/fpp/images/Table7-8.png
[2] https://www.otexts.org/sites/default/files/fpp/images/Table7-10.png

in this paper. Those interested in a detailed explanation on how the nnetar model is operated are referred to [28]. In brief, the package considers fitting feed-forward networks with one hidden layer. The parameters in the neural network model are selected based on a loss function embedded into the learning algorithm. The *nnetar* algorithm trains 25 networks by using random starting values and then obtains the average of the resulting predictions to compute the forecast.

A seasonal *nnetar* model is represented by the notation NNAR$(p, P, k)_m$, a model that has inputs $(y_{t-1}, y_{t-2}, \ldots, y_{t-p}, y_{t-m}, y_{t-2m}, y_{t-Pm})$, and $k$ neurons in the hidden layer. If the values of $p$ and $P$ are not specified (which is the case in this paper), they are automatically selected. For seasonal time series, the default values are $P = 1$ and $p$ is chosen from the optimal linear model fitted to the seasonally adjusted data. If $k$ is not specified, it is set to $k = (p+P+1)/2$ (rounded to the nearest integer).

## 2.4 Trigonometric Box-Cox ARMA Trend Seasonal Model (TBATS)

The TBATS model is an automated exponential smoothing state space model with Box-Cox transformation, ARMA error correction, Trend and Seasonal components. The result is a technique which is aimed at providing accurate forecasts for time series with complex seasonality. However, as reported in the next section, when the automated model was fitted on the data, the algorithm automatically opted for a BATS model (Exponential Smoothing State Space Model With Box-Cox Transformation, ARMA Errors, Trend And Seasonal Components) over TBATS. As noted in [20] BATS model is the most obvious generalization of the traditional seasonal innovations models to allow for multiple seasonal periods. A detailed description of the BATS model can be found in [20].

After generating univariate forecasts from the above models, we determine the best performing univariate forecast for UK consumer price inflation based on a loss function. Then, this best performing forecast is selected as auxiliary information with the MSSA model introduced next.

## 2.5 Multivariate Singular Spectrum Analysis with Auxiliary Information (MSSA(AI))

Figure 1 summarised the inputs in the multivariate system. The observations represented via $(cpi_1^{(1)}, ..., cpi_N^{(1)}) \equiv (cpi_1^{(2)}, ..., cpi_N^{(2)}) \equiv (cpi_1^{(3)}, ..., cpi_N^{(3)})$ as these represent the historical data for UK consumer price inflation. Then, the observations within $(cpi_{N+1}^{(3)}, ..., cpi_{N+h}^{(3)})$ represents the best univariate out-of-sample forecast for UK consumer price inflation and thus data into the future. In fact, the information contained in $(cpi_{N+1}^{(3)}, ..., cpi_{N+h}^{(3)})$ could represent any auxiliary information and the multivariate system can incorporate additional variables in the modelling process.

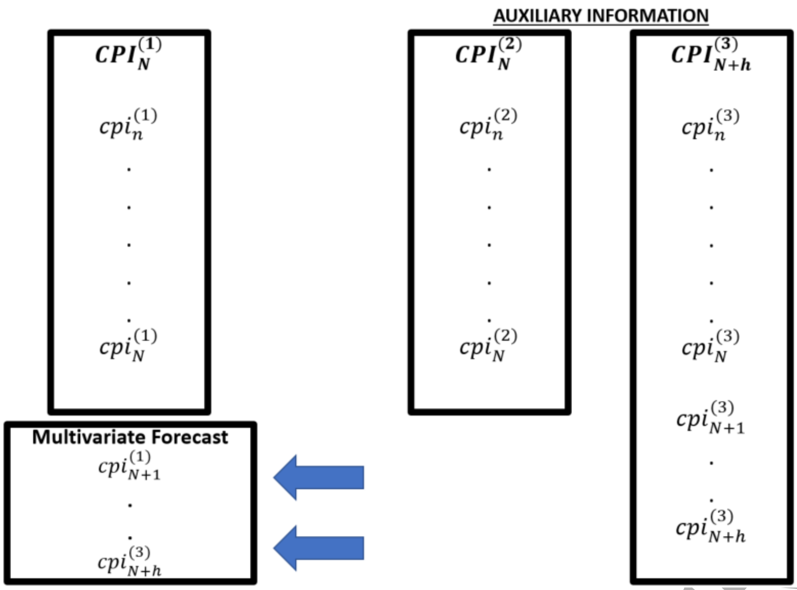


Figure 1: A graphical illustration of the MSSA process and objective.

(1) Consider two time series $CPI_N^{(2)}$ and $CPI_{N+h}^{(3)}$ with an identical frequency. Here, $(cpi_1^{(2)}, ..., cpi_N^{(2)})$ and $(cpi_1^{(3)}, ..., cpi_N^{(3)})$ represents historical data for UK consumer price inflation. Note: $cpi_{N+h}^{(3)}$ represents $cpi_N^{(1)}$ plus the best $h$-step ahead univariate forecast for that same variable. This $h$-step ahead univariate forecast represents data into the future.
(2) Call upon Vertical MSSA which can consider data with different series lengths and a time lag into the future, for developing an improved multivariate forecast for consumer price inflation. Our aim is to obtain a multivariate $h$-step ahead forecast for $CPI_N^{(1)}$ as represented by $\hat{cpi}_{N+1}^{(1)}, \ldots, \hat{cpi}_{N+h}^{(1)}$.
(3) Exploit MSSA's filtering and signal extraction capabilities for modelling and extracting information in $CPI_{N+h}^{(3)}$ (which represents data into the future) and $CPI_N^{(2)}$, for generating a new and improved forecast for the variable in $CPI_N^{(1)}$.

The MSSA technique begins with the decomposition stage which has two steps known as embedding and Singular Value Decomposition (SVD). Initially, through embedding we create the trajectory matrices $\mathbf{X}^{(i)}$ $(i = 1, 2, 3)$ of the one-dimensional time series $CPI_N^{(2)}$ and $CPI_{N+h}^{(3)}$ respectively. As such, we will have 2 different $L_i \times K_i$ trajectory matrices $\mathbf{X}^{(i)}$ $(i = 1, 2, 3)$, where $\mathbf{X}^{(1)}$ will take the form:

$$\mathbf{X}^{(1)} = (x_{ij})_{i,j=1}^{L,K} = \begin{pmatrix} y_1 & y_2 & \cdots & y_K \\ y_2 & y_3 & \cdots & y_{K+1} \\ \vdots & \vdots & \ddots & \vdots \\ y_L & y_{L+1} & \cdots & y_N \end{pmatrix}. \quad (4)$$

A similar trajectory matrix as in Equation (1) can be constructed for $\mathbf{X}^{(2)}$ to represent the data in $Y_N^{(2)}$. Finally, the trajectory matrix $\mathbf{X}^{(3)}$ which incorporates the official forecast can be constructed as:

$$\mathbf{X}^{(3)} = (x_{ij})_{i,j=1}^{L,K+h} = \begin{pmatrix} y_1 & y_2 & \cdots & y_K & y_{K+1} & \cdots & \omega_{K+h} \\ y_2 & y_3 & \cdots & y_{K+1} & y_{K+2} & \cdots & \omega_{K+h+1} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ y_L & y_{L+1} & \cdots & y_N & \omega_{N+1} & \cdots & \omega_{N+h} \end{pmatrix}, \quad (5)$$

where $\omega_1, \ldots, \omega_h$ represents the official forecast.

Thereafter, a new block Hankel trajectory matrix, $\mathbf{X}_H$ is constructed. Assume $L_1 = L_2 = \ldots = L_M = L$ where $M$ is the number of time series. Therefore, we have different values of $K_i$ ($K_i = N_i - L_i + 1$) and series length $N_i$, but similar $L_i$. The result of this step is:

$$\mathbf{X}_H = \left[\ \mathbf{X}^{(1)} : \ \ \mathbf{X}^{(2)} : \ \ \cdots \ \ : \mathbf{X}^{(M)}\ \right].$$

Hence, the structure of the matrix $\mathbf{X}_H\mathbf{X}_H^T$ is as follows:

$$\mathbf{X}_H\mathbf{X}_H^T = \mathbf{X}^{(1)}\mathbf{X}^{(1)^T} + \cdots + \mathbf{X}^{(M)}\mathbf{X}^{(M)^T}. \quad (6)$$

As it appears from the structure of the matrix $\mathbf{X}_H\mathbf{X}_H^T$ in MSSA, we do not have any cross-product between Hankel matrices $\mathbf{X}^{(i)}$ and $\mathbf{X}^{(j)}$. Moreover, in this format, the sum of $\mathbf{X}^{(i)}\mathbf{X}^{(i)^T}$ provides the new block Hankel matrix. Note also that performing the SVD of $\mathbf{X}_H$ in MSSA yields $L$ eigenvalues as in univariate SSA.

Then, in the reconstruction stage we are faced with two steps known as grouping and diagonal averaging. Initially, we group the eigenvalues from the SVD process as either signal or noise (there are several approaches for grouping, see [22,32]) and then perform diagonal averaging on the signal components to reconstruct a new, less noisy time series which can be used for forecasting. A more detailed description of the theory underlying decomposition and reconstruction with MSSA and the forecasting process can be found in [22] and is therefore not reproduced here.

Finally, we present the VMSSA forecasting algorithms used in this paper, and in doing so we mainly follow [22].

## VMSSA Recurrent (VMSSA-R) Forecasts

Let us have two series with different length $Y_{N_i}^{(i)} = (y_1^{(i)}, \ldots, y_{N_i}^{(i)})$ and corresponding window length $L_i$, $1 < L_i < N_i, i = 1, 2$. The VMSSA-R forecasting algorithm for the $h$-step ahead forecast is as follows.

1. For a fixed value of $K$, construct the trajectory matrix $\mathbf{X}^{(i)} = [X_1^{(i)}, \ldots, X_K^{(i)}] = (x_{mn})_{m,n=1}^{L_i,K}$ for each single series $Y_{N_1}^{(1)}$, and $Y_{N_2}^{(2)}$ separately.
2. Construct the block trajectory matrix $\mathbf{X}_V$ as follows:

$$\mathbf{X}_V = \begin{bmatrix} \mathbf{X}^{(1)} \\ \mathbf{X}^{(2)} \end{bmatrix}. \quad (7)$$

3. Denote $\lambda_{V_1} \geq \ldots \geq \lambda_{V_{L_{sum}}} \geq 0$ are the eigenvalues of the $\mathbf{X}_V \mathbf{X}_V^T$, where $L_{sum} = L_1 + L_2$.

4. Let $\mathbf{U}_{V_j} = (U_j^{(1)}, U_j^{(2)})^T$ be the $j^{th}$ eigenvector of the $\mathbf{X}_V \mathbf{X}_V^T$, where $U_j^{(i)}$ with length $L_i$ corresponds to the series $Y_{N_i}^{(i)}$ $(i = 1, 2)$.

5. Consider $\widehat{\mathbf{X}}_V = [\widehat{X}_1 : \ldots : \widehat{X}_K] = \sum_{i=1}^{r} U_{V_i} U_{V_i}^T \mathbf{X}_V$ as the reconstructed matrix achieved from $r$ eigentriples:

$$\widehat{\mathbf{X}}_V = \left[ \begin{array}{c} \widehat{\mathbf{X}}^{(1)} \\ \widehat{\mathbf{X}}^{(2)} \end{array} \right]. \tag{8}$$

6. Consider matrix $\widetilde{\mathbf{X}}^{(i)} = \mathcal{H}\widehat{\mathbf{X}}^{(i)}$ $(i = 1, 2)$ as the result of the Hankelization procedure of the matrix $\widehat{\mathbf{X}}^{(i)}$ obtained from the previous step, where $\mathcal{H}$ is a Hankel operator.

7. Assume $U_j^{(i)\triangledown}$ denotes the vector of the first $L_i - 1$ components of the vector $U_j^{(i)}$ and $\pi_j^{(i)}$ is the last component of the vector $U_j^{(i)}$ $(i = 1, 2)$.

8. Select the number of $r$ eigentriples for the reconstruction stage that can also be used for forecasting purpose.

9. Define matrix $\mathbf{U}^{\triangledown(1,2)} = \left( U_1^{\triangledown(1,2)}, \ldots, U_r^{\triangledown(1,2)} \right)$, where $U_j^{\triangledown(1,2)}$ is as follows:

$$U_j^{\triangledown(1,2)} = \left[ \begin{array}{c} U_j^{(1)\triangledown} \\ U_j^{(2)\triangledown} \end{array} \right]. \tag{9}$$

10. Define matrix $\mathbf{W}$ as follows:

$$\mathbf{W} = \left[ \begin{array}{cccc} \pi_1^{(1)} & \pi_2^{(1)} & \cdots & \pi_r^{(1)} \\ \pi_1^{(2)} & \pi_2^{(2)} & \cdots & \pi_r^{(2)} \end{array} \right]. \tag{10}$$

11. If the matrix $\left( \mathbf{I}_{2\times 2} - \mathbf{W}\mathbf{W}^T \right)^{-1}$ exists and $r \leq L_{sum} - 2$, then the $h$-step ahead VMSSA forecasts exist and is achieved by the following formula:

$$\left[ \hat{y}_{j_1}^{(1)}, \hat{y}_{j_2}^{(2)} \right]^T = \begin{cases} \left[ \tilde{y}_{j_1}^{(1)}, \tilde{y}_{j_2}^{(2)} \right], & j_i = 1, \ldots, N_i \\ \\ \left( \mathbf{I}_{2\times 2} - \mathbf{W}\mathbf{W}^T \right)^{-1} \mathbf{W} \mathbf{U}^{\triangledown 2^T} \mathbf{Z}_h, & j_i = N_i + 1, \ldots, N_i + h, \end{cases} \tag{11}$$

where, $\mathbf{Z}_h = \left[ Z_h^{(1)}, Z_h^{(2)} \right]^T$ and $Z_h^{(i)} = \left[ \hat{y}_{N_i - L_i + h + 1}^{(i)}, \ldots, \hat{y}_{N_i + h - 1}^{(i)} \right]$ $(i = 1, 2)$. It should be noted that equation (11) indicates that the $h$-step ahead forecasts of the refined series $\hat{Y}_{N_i}^{(i)}$ are obtained by a multi dimensional linear recurrent formula (LRF). For the univariate case, there is only a one dimensional LRF.

### VMSSA Vector (VMSSA-V) Forecasts

Let us have items (1)-(10) of VMSSA-R.

1. Define vectors $Z_i$ as follows:

$$Z_i = \begin{cases} \widetilde{X}_i & \text{for } i = 1, \ldots, k \\ \mathcal{P}^{(\nu)} Z_{i-1} & \text{for } i = k+1, \ldots, k+h+L_{\max}-1, \end{cases} \tag{12}$$

   where, $L_{\max} = \max\{L_1, \ldots, L_M\}$.

2. Constructing the matrix $\mathbf{Z} = [Z_1 : ... : Z_{K+h+L_{\max}-1}]$ and making its hankelization. Using this calculation we obtain $\hat{y}_1^{(i)}, \ldots, \hat{y}_{N+h+L_{\max}}^{(i)}$ $(i = 1, \ldots, M)$.

3. The numbers $\hat{y}_{N_i+1}^{(i)}, \ldots, \hat{y}_{N_i+h}^{(i)}$ $(i = 1, \ldots, M)$ form the $h$ step ahead VMSSA-V forecasts.

# 3 Data

The data used in this study was obtained via the Office for National Statistics in UK and relates to annual rates of Consumer Prices Index including owner occupiers housing costs (CPIH) measured monthly from January 2006 to May 2018 (Figure 2)[3]. It is noteworthy that the CPIH is the lead inflation index in UK and is recognised as the most comprehensive measure of inflation as it includes owner occupiers housing costs and Council Tax, which are excluded from the Consumer Price Index (CPI) [34].

Table 1 reports some key descriptive statistics for this data. This shows that the data is positively skewed and not normally distributed. Accordingly, it is evident that during the period in question, UK inflation as per the CPIH has averaged at 2.25% with a standard deviation of +/-0.99%. The median shows that inflation as per the CPIH was at or below 2.40% during half of the months over this period with an inter quartile range of +/- 1%. The data was also tested for normality using the Shapiro-Wilk test and as the test statistic was highly statistically significant it lead to a rejection of the null hypothesis, thereby indicating that the CPIH series is not normally distributed. As such, the median rate of inflation for UK at 2.40% is a more accurate measure of central tendency during the period in question.

Table 1: Descriptive statistics for CPIH data (Jan 2006 - May 2018).

| **Mean** | **Median** | **SD** | **IQR** | **SW** |
|---|---|---|---|---|
| 2.25% | 2.40% | 0.99% | 1% | <0.01 |

*Note*: SD is the standard deviation. IQR is the interquartile range. SW is the $p$-value for the Shapiro-Wilk test for normality.

[3] `https://www.ons.gov.uk/economy/inflationandpriceindices/timeseries/l55o/mm23`


Figure 2: UK monthly consumer price inflation time series.

In Table 2 we provide a summary of the forecasting exercise carried out in this study. We consider different out-of-sample horizons with different sampling periods because research has shown that the forecast performance of models are episodic and sensitive to these factors [2]. Therefore, if the forecasting results under such conditions can indicate a single model as being superior across all horizons, then we can have more confidence in the robustness of our results.

Table 2: Summary of forecasting exercise.

| $h$ | **In-Sample Period** | **Out-of-Sample Period** | $\hat{N}$ |
|---|---|---|---|
| $h = 3$ | Jan 2006 - May 2015 | Jun 2015 - Aug 2015 | 3 |
| $h = 6$ | Jan 2006 - May 2015 | Jun 2015 - Nov 2015 | 6 |
| $h = 12$ | Jan 2006 - May 2015 | Jun 2015 - May 2016 | 12 |
| $h = 24$ | Jan 2006 - May 2015 | Jun 2015 - May 2017 | 24 |
| $h = 36$ | Jan 2006 - May 2015 | Jun 2015 - May 2018 | 36 |

*Note*: $\hat{N}$ is the number of out-of-sample forecasts.

In line with good practice, and to enable replication of the forecasting results presented in the following Section, in Table 3 we present readers with the models used for obtaining the out-of-sample forecasts for UK consumer price inflation in this paper.

Table 3: Summary of fitted models for out-of-sample forecast generation.

| $h$ | ARIMA | ETS | NN | TBATS | MSSA(AI) |
|---|---|---|---|---|---|
| $h = 3$ | ARIMA(0,1,2)(0,0,2) | ETS(A,Ad,N) | NNAR(3,1,2) | BATS(0.557, 0,0) | MSSA-V(48,12) |
| $h = 6$ | ARIMA(0,1,2)(0,0,2) | ETS(A,Ad,N) | NNAR(3,1,2) | BATS(0.557, 0,0) | MSSA-V(36,8) |
| $h = 12$ | ARIMA(0,1,2)(0,0,2) | ETS(A,Ad,N) | NNAR(3,1,2) | BATS(0.557, 0,0) | MSSA-V(48,5) |
| $h = 24$ | ARIMA(0,1,2)(0,0,2) | ETS(A,Ad,N) | NNAR(3,1,2) | BATS(0.557, 0,0) | MSSA-R(48,9) |
| $h = 36$ | ARIMA(0,1,2)(0,0,2) | ETS(A,Ad,N) | NNAR(3,1,2) | BATS(0.557, 0,0) | MSSA-V(48,8) |

*Note*: $ARIMA(p,d,q)$ is where $p$ is the autoregressive part, $d$ is the degree of first differencing and $q$ is the order of the moving average. $ETS(A, Ad, N)$ refers to an additive damped trend method. $NNAR(p, P, k)$ is where $p$ indicates lagged inputs, $P$ takes a default value of 1 for seasonal time series, and $k$ shows the neurons in the hidden layer. $MSSA(L, r)$ refers to the MSSA(AI) choices.

## 3.1 Metrics

The accuracy of forecasts are distinguished based on the Root Mean Squared Error (RMSE) and the Ratio of the RMSE criteria. Both these criteria are frequently cited loss functions, see for example [23–26] and references therein.

$$RMSE = \left(\frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2\right)^{\frac{1}{2}}, \tag{13}$$

where, $Y_i$ is the actual value, $\hat{Y}_i$ refers to a forecast from a given model, and $n$ is the number of the forecasts. Likewise, the ratio of the RMSE can be easily calculated as:

$$RRMSE = \frac{RMSE_{MSSA(AI)}}{RMSE_{Benchmark}}, \tag{14}$$

where $RMSE_{MSSA(AI)}$ refers to the RMSE from the MSSA(AI) model and $RMSE_{Benchmark}$ refers to the RMSE for the competing forecast. Then, if the RRMSE is less than 1, the MSSA (AI) approach outperforms the benchmark model by 1-RRMSE%.

# 4 Results

The out-of-sample forecasting RMSE results are reported in Table 4. As noted above, we generate out-of-sample forecasts for UK consumer price inflation for the short, medium and long run and thereby present a clear overview of the model performance for policy makers and forecasters interested in this output.

The first observation is that a feed-forward NN model with one hidden layer fails to provide the best forecast for UK consumer price inflation across any of the horizons. Secondly, if one was interested in obtaining univariate forecasts for UK consumer price inflation, then we are able to recommend ETS as the most appropriate model at $h = 3$ and $h = 6$ steps-ahead whilst BATS produces the best univariate forecast at $h = 12$ steps-ahead. In the very long run, at $h = 24$ and $h = 36$ steps-ahead, we find ARIMA to be the best univariate model. In other words, for short term UK inflation forecasting we can recommend ETS as the most appropriate model, for medium term forecasting, ETS and BATS (depending on the horizon of interest), and for long term forecasting, ARIMA.

Table 4: Out-of-sample forecasting RMSE for UK consumer price inflation.

| $h$ | $\hat{N}$ | **ARIMA** | **ETS** | **NN** | **BATS** | **MSSA(AI)** |
|---|---|---|---|---|---|---|
| $h = 3$ | 3 | 0.087 | **0.082** | 0.177 | 0.085 | 0.082 |
| $h = 6$ | 6 | 0.295 | **0.129** | 0.312 | 0.143 | 0.116 |
| $h = 12$ | 12 | 0.606 | 0.210 | 0.299 | **0.201** | 0.122 |
| $h = 24$ | 24 | **0.629** | 0.922 | 0.865 | 0.906 | 0.355 |
| $h = 36$ | 36 | **0.560** | 1.521 | 1.441 | 1.503 | 0.337 |

*Note*: $\hat{N}$ is the number of out-of-sample forecasts. Shown in bold font are the best univariate forecasts at each horizon. The best univariate forecast is then used as additional information within the MSSA(AI) model.

The key highlight in Table 4 is the performance of the MSSA(AI) models. These MSSA models have used the best performing univariate forecasts for each horizon as auxiliary information along side the historical CPIH data for generating the new and improved multivariate forecasts for UK consumer price inflation. Interestingly, we notice that there is no real gain made with the MSSA(AI) approach when seeking to forecast UK inflation in the very short run. This result is understandable as the whole premise underlying the proposed MSSA(AI) approach is that there should be auxiliary information or signals which can be extracted in a given forecast. Thus, at $h = 3$ steps-ahead, it is not surprising that the model does not seem to be able to extract sufficient information from the best performing ETS forecast for improving the MSSA(AI) predictions for UK inflation. However, when considering the RMSE results for the medium to long term forecasting, we can see that the MSSA(AI) forecasts clearly outperform the competing univariate model forecasts based on this loss function.

In line with good practice, all out-of-sample forecasts are evaluated for statistically significant differences using the Hassani-Silva (HS) test for comparing predictive accuracy in [33]. These results and the RRMSE output for UK consumer price inflation forecasting are reported in Table 5. Based on these results, we can make the following conclusions.

In the very short run, i.e., at $h = 3$ steps-ahead, the results indicate that the MSSA(AI) forecasts are 6%, 54%, and 3% better than the forecasts from ARIMA, NN, and BATS models. However, the ETS forecast is seen outperforming the MSSA(AI) forecast by 1%. Even though none of the outcomes are statistically significant, there is no justification for opting for a comparatively complex multivariate modelling approach when seeking to forecast UK consumer price inflation in the very short run as the MSSA(AI) forecast cannot outperform the best univariate forecast in this instance.

In contrast, at $h = 6$ steps-ahead, the MSSA(AI) forecasts are 61%, 10%, 63%, and 19% better than the ARIMA, ETS, NN, and BATS forecasts, respectively. As with the previous case, we find no evidence of statistically significant differences between the MSSA(AI) forecasts and those obtained via the univariate models. The failure to pick up statistically significant outcomes at $h = 3$

and $h = 6$ steps-ahead is not entirely surprising given the very small sample sizes. Yet, when considering the gains attainable via the MSSA(AI) approach in relation to the competing models at this particular horizon, it is suitable to consider modelling with MSSA to improve the accuracy of forecasts here.

When forecasting at $h = 12$ steps-ahead, we find that the MSSA(AI) forecasts are statistically significantly better than forecasts from ARIMA, ETS, NN and BATS with respective accuracy gains of 80%, 42%, 59%, and 39%, respectively. Here, it is interesting that even with a sample size of 12 out-of-sample observations, the HS test is able to pick up significant differences between the MSSA(AI) and competing forecasts. This further demonstrates the true power of the approach underlying the MSSA(AI) model at providing significantly better inflation forecasts for UK at $h = 6$ steps-ahead.

In terms of long term forecasts for UK consumer price inflation, at $h = 24$ steps-ahead, we find evidence indicating the MSSA(AI) forecasts are statistically significantly better than those from ARIMA, ETS, NN and BATS models with respective accuracy gains of 44%, 62%, 59%, and 39%. In the very long run, at $h = 36$ steps-ahead, we find the MSSA(AI) forecasts are once again statistically significantly better than the competing forecasts with accuracy gains of 40%, 78%, 77%, and 78% in relation to forecasts from ARIMA, ETS, NN, and BATS respectively.

Overall, we find conclusive evidence which indicates that the proposed MSSA(AI) approach for forecasting UK inflation with inflation forecasts is viable and worthy of careful consideration. In particular, our findings show that the MSSA(AI) approach can produce statistically significant accuracy gains when used to forecast at $h = 12$, $h = 24$ and $h = 36$ steps-ahead, whilst considerable accuracy gains are attainable at $h = 6$ steps-ahead too (with no statistically significant differences). At the same time, we do not find evidence for its use for forecasting UK consumer price inflation in the very short run.

Table 5: Out-of-sample forecasting RRMSE for UK consumer price inflation.

| $h$ | $N$ | $\frac{MSSA(AI)}{ARIMA}$ | $\frac{MSSA(AI)}{ETS}$ | $\frac{MSSA(AI)}{NN}$ | $\frac{MSSA(AI)}{BATS}$ |
|---|---|---|---|---|---|
| $h = 3$ | 3 | 0.94 | 1.01 | 0.46 | 0.97 |
| $h = 6$ | 6 | 0.39 | 0.90 | 0.37 | 0.81 |
| $h = 12$ | 12 | 0.20*** | 0.58* | 0.41* | 0.61* |
| $h = 24$ | 24 | 0.56** | 0.38** | 0.41* | 0.39* |
| $h = 36$ | 36 | 0.60*** | 0.22*** | 0.23*** | 0.22*** |

*Note*: $N$ is the number of observations being forecasted. All outcomes are tested for statistically significant differences between the distributions of MSSA(AI) forecasts and a competing forecast based on the Hassani-Silva (HS) test [33]. *** indicates the results are statistically significant at $p = 0.01$, ** at $p = 0.05$, and * at $p = 0.10$.

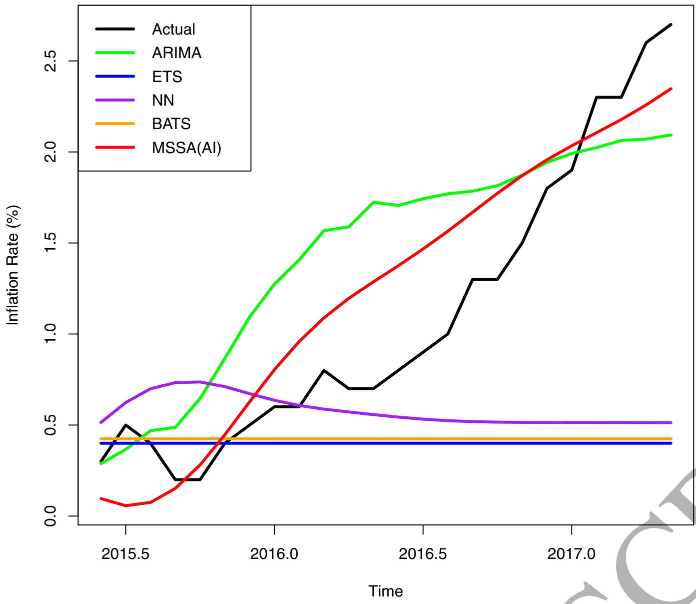


Figure 3: Out-of-sample forecasts for UK consumer price inflation at $h = 24$ steps-ahead.

Finally, in Figures 3 and 4 we provide a graphical representations of the out-of-sample forecasts for UK consumer price inflation in the long run from all models considered in this study. This clearly demonstrates the inferiority of the NN, BATS and ETS models at generating forecasts for UK consumer price inflation in the long run whilst also demonstrating the superiority of forecasts from the MSSA(AI) model in relation to those from the best performing univariate model, ARIMA. In relation to the views expressed in [2, 9] with regard to univariate model's capabilities at inflation forecasting, our findings are in line with theirs when the forecasting horizon is beyond $h = 3$ steps-ahead, as the univariate models are clearly experiencing difficulties modeling and forecasting inflation at these horizons. However, at $h = 3$ steps-ahead, we find the univariate approach can be more reliable than the multivariate approach proposed in this paper.

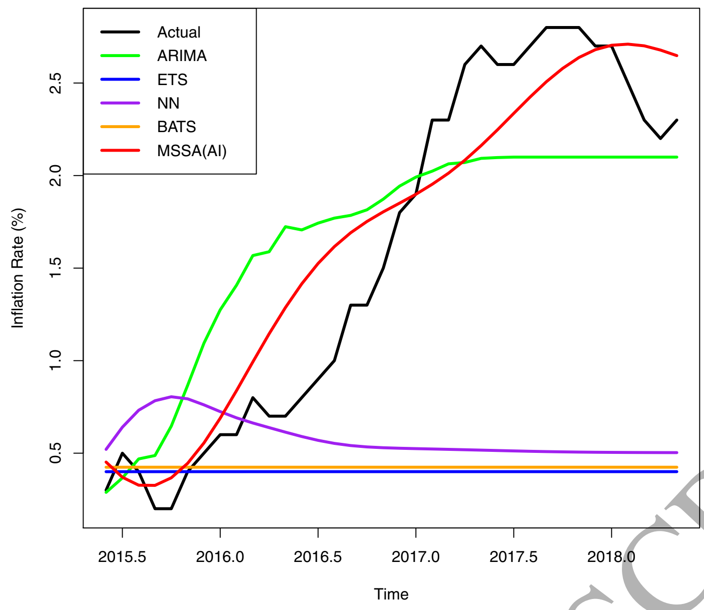


Figure 4: Out-of-sample forecasts for UK consumer price inflation at $h = 36$ steps-ahead.

# 5 Discussion

The MSSA(AI) approach for forecasting UK consumer price inflation has several modeling capabilities which are worthy of discussion. The importance of its varied modelling capabilities are further enhanced by the prevalence of Big Data (see for example, [35–37] and references therein) which generates more rapid information that could be useful for improving forecast accuracy.

## Option 1: Use of Purely Model Based Forecasts for Inflation

This option is evaluated in detail in this paper. Here, we propose generating forecasts for UK consumer price inflation using univariate models and then select the best performing forecast as auxiliary information within the MSSA(AI) framework. It is noteworthy that forecasts from other multivariate models can also be considered as auxiliary information.

## Option 2: Use of Official Forecasts for Inflation

Central Banks and National Statistical Institutes can sometimes be trapped in traditions [38] which limits their willingness to do away with entirely new approaches to modeling and forecasting. The MSSA(AI) approach can be useful under such circumstances as it does not require Central Bankers to completely do away with their existing forecasting approaches which are used to publish official forecasts for inflation. Instead, they can consider modelling their official forecast as auxiliary information within the MSSA(AI) model to generate a new and improved forecast.

## Option 3: Use of Judgemental Forecasts for Inflation

The MSSA(AI) model is for exploiting auxiliary information in forecasts. Thus, it is able to also consider judgemental or even professional forecasts as auxiliary information and seek to extract signals contained within such forecasts for generating a new and improved multivariate forecast.

## Option 4: Use of Additional Variables within MSSA(AI) Framework

Given that MSSA(AI) is a multivariate model capable of handling multiple time series, there is room to further develop the accuracy of forecasts by combining leading indicators for a variable of interest alongside historical data and a forecast. Noteworthy is the fact that multiple forecasts can also be modeled.

Here, we consider Google Trends for inflation in UK (Figure 5) as an additional variable in the MSSA(AI) model. In this case, we forecast the last 36 observations for Google Trends using the same univariate models and then select the best univariate forecast for Google Trends, the previously best MSSA(AI) forecast for UK consumer price inflation and historical data as inputs. The RRMSE's from this forecasting exercise are reported via Table 6. The new forecast (MSSA(AI):GT) is able to outperform the initial MSSA(AI) forecast by 4% in this example. Whilst this gain is not statistically significant, there is evidence that more research into modeling multiple forecasts within the MSSA(AI) framework could lead to the generation of more accurate forecasts.

Figure 6 shows the out-of-sample forecasts from MSSA(AI) and MSSA(AI):GT. This simple example illustrates that as the forecasting horizon expands beyond 2017, the MSSA(AI):GT model can generate forecasts which are comparatively better aligned with the actual series.

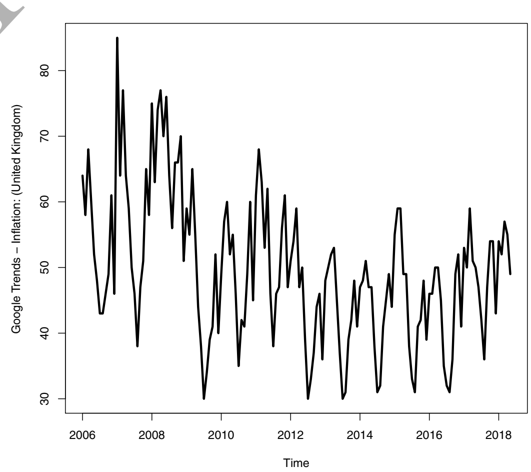


Figure 5: Google Trends for inflation in UK.

Table 6: Out-of-sample forecasting RRMSE for UK consumer price inflation with Google Trends.

| $h$ | $N$ | $\frac{MSSA(AI):GT}{ARIMA}$ | $\frac{MSSA(AI):GT}{ETS}$ | $\frac{MSSA(AI):GT}{NN}$ | $\frac{MSSA(AI):GT}{BATS}$ | $\frac{MSSA(AI):GT}{MSSA(AI)}$ |
|---|---|---|---|---|---|---|
| $h = 36$ | 36 | $0.58^{***}$ | $0.21^{***}$ | $0.22^{***}$ | $0.22^{***}$ | 0.96 |

*Note*: $N$ is the number of observations being forecasted. All outcomes are tested for statistically significant differences between the distributions of MSSA(AI) forecasts and a competing forecast based on the Hassani-Silva (HS) test [33]. *** indicates the results are statistically significant at $p = 0.01$

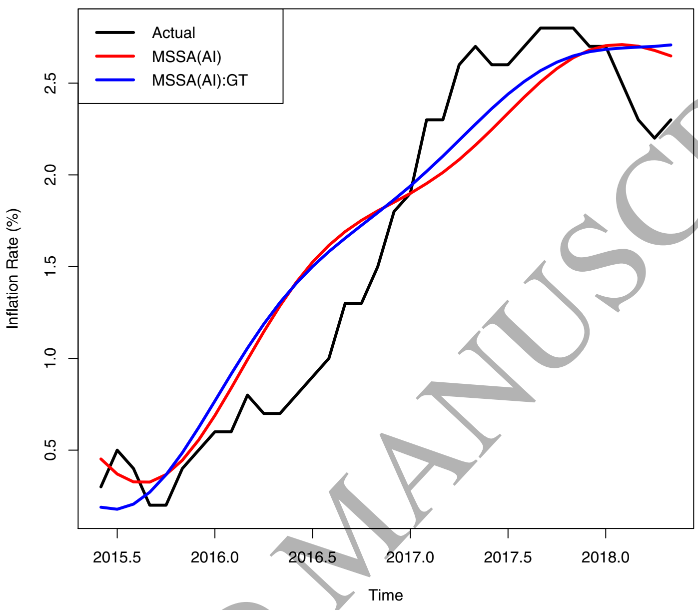


Figure 6: Out-of-sample forecasts for UK consumer price inflation at $h = 36$ steps-ahead.

# 6 Conclusion

This paper begins with the aim of exploiting the auxiliary information contained within forecasts of UK consumer price inflation for generating a new forecast which is more accurate than the initial forecast. In other words, the new multivariate approach considers modeling and extracting the signals from a given forecast for UK inflation and combines this information with historical data for UK inflation to produce a new and comparatively more accurate forecast. We exploit the nonparametric MSSA technique as the tool for achieving this goal.

The forecasting exercise begins by modeling and forecasting UK consumer price inflation using a variety of parametric and nonparametric, optimized univariate forecasting models. The univariate forecasting exercise shows that, for short term ($h = 3$ steps-ahead) UK inflation forecasting, ETS is the most appropriate model, whilst for medium term forecasting, ETS (at $h = 6$ steps-ahead) and BATS (at $h = 12$ steps-ahead) are best, with ARIMA reporting the best univariate forecast in the long run ($h = 24$ and $h = 36$ steps-ahead).

Thereafter, we consider the best univariate forecast for UK consumer price inflation as auxiliary information within the MSSA modeling approach and combine this univariate forecast with historical inflation data at each horizon to generate a new and improved inflation forecast. Overall, we find conclusive evidence which supports the proposed MSSA(AI) approach for forecasting UK inflation. The findings of this study clearly demonstrates that the underlying methodology is worthy of careful consideration by Central Bankers and forecasters alike.

In particular, our findings show that the MSSA(AI) approach can produce statistically significant accuracy gains when used to forecast at $h = 12$, $h = 24$ and $h = 36$ steps-ahead, whilst considerable accuracy gains are attainable at $h = 6$ steps-ahead (but with no statistically significant differences). However, it is noteworthy that we do not find evidence for its use for forecasting UK consumer price inflation in the very short run.

In conclusion, the introduction of this novel approach opens several research avenues. For example, there is a need for comparing the results with more varied univariate and multivariate forecasts for inflation in order to show the consistency of the MSSA(AI) approach as a reliable inflation forecasting tool. Moreover, automation of the MSSA(AI) algorithm can be extremely useful given the importance of technology and algorithms in driving forecast generation in the modern age. The automation process should carefully consider the selection of MSSA(AI) choices of Window Length $L$ and the number of eigenvalues $r$ which can be time consuming otherwise. Finally, extensive simulation studies are needed with consideration given to forecasts with different distributions and accuracy levels to determine the sensitivity and generalisability of the MSSA(AI) approach.

## References

[1] Cogley, T., Morozov, S., and Sargent, T. J. (2005). Bayesian fan charts for U.K. inflation: Forecasting and sources of uncertainty in an evolving monetary system. *Journal of Economic Dynamics and Control*, **29**(11), 1893–1925.

[2] Di Filippo, G. (2015). Dynamic Model Averaging and CPI Inflation Forecasts: A Comparison between the Euro Area and the United States. *Journal of Forecasting*, **34**, 619–648.

[3] Hubrich, K., and Skudelny, F. (2017). Forecast Combination for Euro Area Inflation: A Cure in Times of Crisis? *Journal of Forecasting*, **36**(5), 515–540.

[4] Koop, G., and Korobilis, D. (2012). Forecasting inflation using dynamic model averaging. *International Economic Review*, **53**(3), 867–886.

[5] Wright, J. H. (2009). Forecasting US inflation by Bayesian model averaging. *Journal of Forecasting*, **28**(2), 131–144.

[6] Bjrnland, H. C., Gerdrup, K., Jore, A. S., Smith, C., and Thorsrud, L. A. (2012). Does Forecast Combination Improve Norges Bank Inflation Forecasts? *Oxford Bulletin of Economics and Statistics*, **74**(2), 163–179.

[7] Melo, L. F., Loaiza, R. A., and Villamizar-Villegas, M. (2016). Bayesian combination for inflation forecasts: The effects of a prior based on central banks estimates. *Economic Systems*, **40**(3), 387–397.

[8] Simionescu, M., Gavurova, B., and Smrcka, L. (2017). Combined forecasts to improve Survey of Profession Forecasters predictions for quarterly inflation in the U.S.A. *Economic Research-Ekonomska Istraivanja*, **30**(1), 789–805.

[9] Orphanides, A., and van Norden S. (2005). The reliability of inflation forecast based on output gap estimates in real time. *Journal of Money, Credit, and Banking*, **37**(3), 583–600.

[10] Iqbal, M., and Naveed, A. (2016). Forecasting Inflation: Autoregressive Integrated Moving Average Model. *European Scientific Journal*, **12**(1), 83–92.

[11] Pincheira, P. M., and Medel, C. A. (2015). Forecasting Inflation with a Simple and Accurate Benchmark: The Case of the US and a Set of Inflation Targeting Countries. *Czech Journal of Economics and Finance*, **65**(1), 2–29.

[12] Carlo, T. C., and Marcal, E. F. (2016). Forecasting Brazilian inflation by its aggregate and disaggregated data: a test of predictive power by forecast horizon. *Applied Economics*, **48**(50), 4846–4860.

[13] Silva, E. S., Hassani, H., Otero, J., and Beneki, C. (2018). Forecasting Inflation Under Varying Frequencies. *Electronic Journal of Applied Statistical Analysis*, **11**(1), 307–339.

[14] Jere, S., and Siyanga, M. (2016). Forecasting Inflation Rate of Zambia Using Holts Exponential Smoothing. *Open Journal of Statistics*, **6**, 363–372.

[15] Rich, R., and Steindel, C. (2005). A Review of Core Measure and an Evaluation of Its Measures. *Federal Reserve Bank of New York Staff Report*, **236**, (December).

[16] Gupta, S., and Kashyap, S. (2015). Forecasting inflation in G-7 countries: an application of artificial neural network. *Foresight*, **17**(1), 63–73.

[17] McAdam, P., and McNelis, P. (2015). Forecasting inflation with thick models and neural networks. *Economic Modelling*, **22**(5), 848–867.

[18] Thakur, G. S. M., Bhattacharyya, R., and Mondal, S. S. (2016). Artificial Neural Network Based Model for Forecasting of Inflation in India. *Fuzzy Information and Engineering*, **8**(1), 87–100.

[19] Moshiri, S., and Cameron, N. (2000). Neural network versus econometric models in forecasting inflation. *Journal of Forecasting*, **19**(3), 201–217.

[20] De Livera, A. M., Hyndman, R. J., and Snyder, R. D. (2011). Forecasting time series with complex seasonal patterns using exponential smoothing. *Journal of the American Statistical Association*, **106**(496), 1513–1527.

[21] Sanei, S., and Hassani, H. (2015). *Singular spectrum analysis of biomedical signals.* United States: CRC Press.

[22] Hassani, H., and Mahmoudvand, R. (2013). Multivariate Singular Spectrum Analysis: A general view and new Vector forecasting approach. *International Journal of Energy and Statistics*, **1**(1), 55–83.

[23] Silva, E. S., Hassani, H., and Heravi, S. (2018). Modeling European industrial production with multivariate singular spectrum analysis: A cross-industry analysis. *Journal of Forecasting*, **37**(3), 371–384.

[24] Silva, E. S., Ghodsi, Z., Ghodsi, M., Heravi, S., and Hassani, H. (2017). Cross country relations in European tourist arrivals. *Annals of Tourism Research*, **63**, 151–168.

[25] Hassani, H., Soofi, A. S., and Zhigljavsky, A. (2013). Predicting inflation dynamics with singular spectrum analysis. *Journal of the Royal Statistical Society A*, **76**(3), 743–760.

[26] Hassani, H., Silva, E. S., Gupta, R., and Das, S. (2018). Predicting global temperature anomaly: A definitive investigation using an ensemble of twelve competing forecasting models. *Physica A: Statistical Mechanics and its Applications*, **509**, 121–139.

[27] Hydman, R. J., and Khandakar, Y. (2008). Automatic time series forecasting: The forecast package for R. *Journal of Statistical Software*, **27**(3), 1–22.

[28] Hyndman, R. J., and Athanasopoulos, G. (2018). *Forecasting: Principles and practice.* 2nd ed., OTexts. Available via: `https://otexts.org/fpp2/`. [Accessed: 23.06.2018].

[29] Kwiatkowski, D., Phillips, P. C. B., Schmidt, P., and Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root. *Journal of Econometrics*, **54**(13), 159–178.

[30] Cleveland, R. B., Cleveland, W. S., McRae, J. E., and Terpenning, I. J. (1990). STL: A Seasonal-Trend Decomposition Procedure Based on Loess. *Journal of Official Statistics*, **6**(1), 3–73.

[31] Makridakis, S. G., Wheelwright, S. C. and Hyndman, R. J. (1998). *Forecasting: Methods and applications.* Wiley.

[32] Hassani, H., Ghodsi, Z., Silva, E. S., and Heravi, S. (2016). From nature to maths: Improving forecasting performance in subspace-based methods using genetics Colonial Theory. *Digital Signal Processing*, **51**, 101–109.

[33] Hassani, H., and Silva, E. S. (2015). A Kolmogorov-Smirnov Based Test for Comparing the Predictive Accuracy of Two Sets of Forecasts. *Econometrics*, **3**(3), 590–609.

[34] Office for National Statistics (2018). Methodology: Consumer Price Inflation (includes all 3 indices CPIH, CPI and RPI) QMI Available via: `https://www.ons.gov.uk/economy/inflationandpriceindices/methodologies/consumerpriceinflationincludesall3indicescpihcpiandrpiqmi`. Accessed: [04.07.2018].

[35] Hassani, H., and Silva, E. S. (2018). Big Data: a big opportunity for the petroleum and petrochemical industry. *OPEC Energy Review*, **42**(1), 74–89.

[36] Hassani, H., Huang, X., and Silva, E. S. (2018). Digitalisation and Big Data Mining in Banking. *Big Data and Cognitive Computing*, In Press.

[37] Hassani, H., Huang, X., Silva, E. S., and Ghodsi, M. (2016). A review of data mining applications in crime. *Statistical Analysis and Data Mining: The ASA Data Science Journal*, **9**(3), 139–154.

[38] Hassani, H., Saporta, G., and Silva, E. S. (2014). Data Mining and Official Statistics: The Past, the Present, and the Future. *Big Data*, **2**(1), 34–43.