Bournemouth University

BAFES –
Bournemouth Accounting, Finance & Economic Series

NO 07 / 2016

Forecasting UK Income Tax

Ghodsi, Z. & Webster A.

Forecasting UK Income Tax

Zara Ghodsi[a], Allan Webster[a]

[a] *Statistical Research Centre, Bournemouth University,*
89 Holdenhurst Road, Bournemouth BH8 8EB, UK

[1] *zghodsi@bournemouth.ac.uk*

Abstract

The literature on forecasting tax revenues focuses on the need for a body of competing forecasts independent of government, to limit potential political bias. The Office for Budget Responsibility does provide detailed independent forecasts for the UK but there are limited alternatives. The literature on appropriate techniques for forecasting detailed tax revenues is under-developed. In many countries tax revenue forecasts are embedded in a more extensive macro-economic forecasting model. These models lack sufficient precision for revenue forecasting revenues for several specific taxes. Such models are too involved to support a body of competing independent forecasts. In consequence there is an established need for single equation revenue forecasts for specific taxes to complement the macro-economic approach. This study considers the use of a number of (mainly) time series forecasting techniques. We find Recurrent Singular Spectrum Analysis (RSSA) to perform the best of the techniques considered.

Keywords: United Kingdom; Income Tax; Forecasting; Singular Spectrum Analysis; ARIMA; Exponential Smoothing; Neural Networks.

1 Introduction

Two key themes emerge from the existing literature on forecasting tax revenues and the related literature on budget forecasting. Firstly, the literature stresses the need for independent forecasts and, in particular, a body of competing independent forecasts. Secondly, techniques for tax forecasting are acknowledged in the literature as being currently under-developed. Some authors such as Buettner and Kauder argue in favour of tax forecasts embedded within wider macro-economic models [1]. Other authors such as Bretschneider et al have argued in favour of simple econometric models [2].

The debate about appropriate techniques does not just concern issues of accuracy or efficiency of the forecasts but also some pragmatic issues. For example, Leal et al express the view that macro-economic models are often too aggregated to provide detailed revenue projections [3].

In this paper we consider the use of direct estimation using a number of (mainly) time series approaches. We examine the scope for using time series forecasting techniques as a contribution to the body of independent forecasting. Macro-economic models have undeniable advantages but are complex and costly. Only few organisations have the resources to maintain them. The intention of this paper is to assess the performance of a number of different direct (not embedded within a macro model) forecasting techniques. The purpose is to ask the question: Can time series forecasting make a valid contribution to the body of tax revenue forecasts?

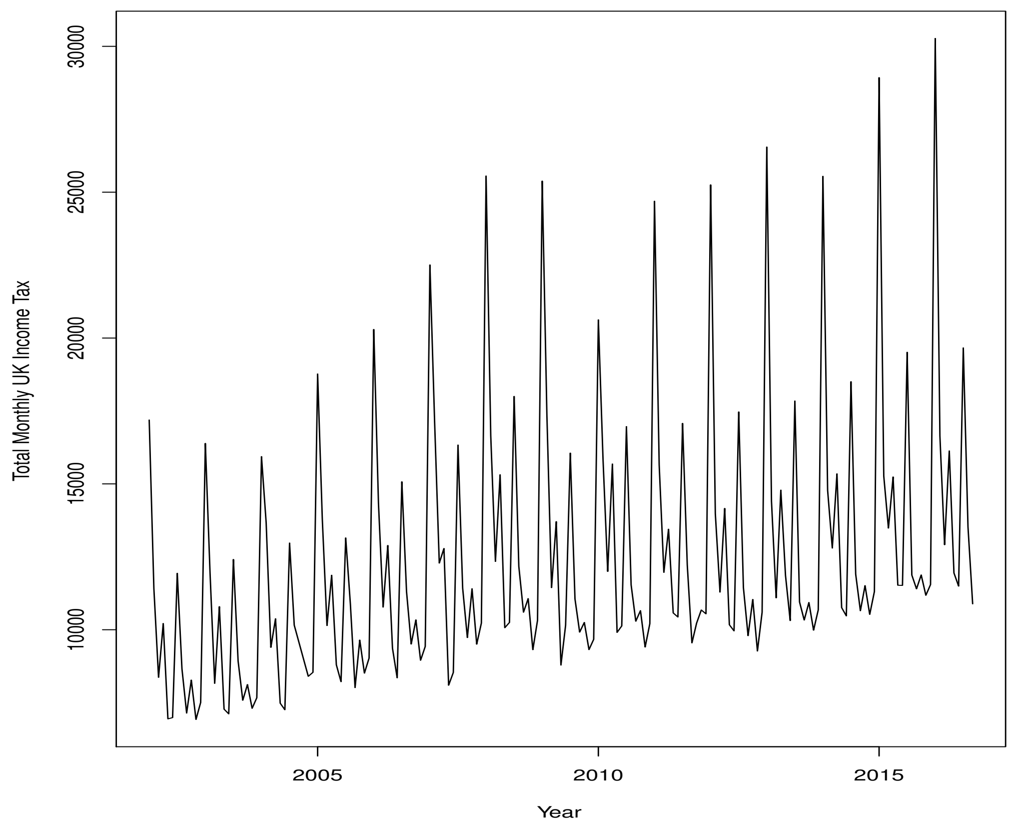

Figure 1: Total monthly UK Income Tax time series (Jan. 2002 - Sep. 2016).

To do this we evaluate the performance of a number of different direct forecasting techniques relative to each other and relative to the forecasts for UK tax revenues produced by the *Office for Budget Responsibility*. The remainder of this paper is organised as follows. Section 2 reviews the relevant literature. Section 3 provides the details of the main forecasting techniques employed. Section 4 describes the data for UK income tax and discusses the measures for evaluating the forecasting performance. Section 5 covers the empirical results and conclusions are provided in section 6.

2 Review of Literature

As Leal et al note the literature on forecasting tax revenues has two main themes [3]. The first of these, as with the wider issue of budget forecasting, is that forecasts are often perceived as biased for politically motivated reasons. For the UK politically motivated bias less of a concern than for some countries since the establishment of the independent Office for Budget Responsibility in 2010. The second main theme that Leal et al identify concerns appropriate forecasting techniques [3]. As they note this literature is under-developed and yet to reach definitive conclusions. The focus of this paper is on this second strand in the literature.

The debate on perceived political bias and appropriate techniques for tax forecasting is not new. Auerbach in a study of government revenue forecasts in the US, found little difference between the performance of government and independent forecasters, suggesting a lack of political bias [4]. The study did, however, find a need for better forecasting performance. In particular the need for an improved dynamic specification was emphasized.

Buettner and Kauder reviewed the practice and performance of revenue forecasts for a number of OECD countries [1]. Their study found two key factors to reduce forecast errors: the use of a macro-economic model and independence from government. They found a number of common strands across countries. In particular, most countries adopted a mixed approach of using a macro-economic model to forecast those taxes most likely to be related to the business cycle and direct (single equation) methods for other taxes.

Chatagny and Soguel provide an interesting bridge between the literature on tax revenue forecasting and the related, wider literature on budget forecasting [5]. In a study of Swiss cantons they found that systematic under forecasting of tax revenues was associated with reduced fiscal deficits.

The literature on budget forecasting necessarily overlaps that of tax revenue forecasting with respect to perceived political bias. There is also significant common ground with respect to the use of macro-economic models for forecasting but there is one important difference. For tax revenues alone direct forecasts not embedded in a macro-economic model are not only feasible but widely used (for at least some taxes). Budget forecasts are, of necessity, more complex. Artis and Marcellino examined the performance of IMF, OECD and the EC, in forecasting budget deficits (in relation to GDP) [6]. They found systematic under forecasting for some countries and systematic over forecasting for others. Jonung and Larch found evidence of systematic political bias in budget forecasts for EU countries [7]. These studies tend to further emphasise the role of independent forecasters.

3 Forecasting Methods

3.1 Singular Spectrum Analysis (SSA)

As proposed in [8] the two complementary stages of the SSA, i.e., the decomposition and reconstruction stages can be carried out in four steps.

Step 1: Embedding
Embedding can be considered as a mapping that transfers a one-dimensional time series $Y_N = (y_1, \ldots, y_N)$ into the multi-dimensional series $X_1, \ldots, X_K$ with vectors $X_i = (y_i, \ldots, y_{i+L-1})^T \in \mathbf{R}^L$, where L, $(2 \leq L \leq N-1)$ is the window length and $K = N - L + 1$. This first step provides the trajectory matrix $\mathbf{X} = [X_1, \ldots, X_K] = (x_{ij})_{i,j=1}^{L,K}$.

Step 2: SVD
In this step we perform the Singular Value Decomposition (SVD) of $\mathbf{X}$ into a sum of rank-one bi-orthogonal elementary matrices, $\mathbf{X} = \mathbf{X}_1 + \ldots + \mathbf{X}_L$, with $\mathbf{X}_i = \sqrt{\lambda_i} U_i V_i^T$ and $i = 1, \ldots, L$. $\lambda_1, \ldots, \lambda_L$ are the eigenvalues of $\mathbf{X}\mathbf{X}^T$ $(\lambda_1 \geq \ldots \geq \lambda_L \geq 0)$, $U_1 \ldots U_L$ the corresponding eigenvectors and $V_1 \ldots V_L$ are the principal components defined as $V_i = X^T U_i / \sqrt{\lambda_i}$.

Step 3: Grouping
The grouping step consists in splitting the elementary matrices in *Step 2* into m disjoint groups and summing the matrices within each group to obtain new matrices, $\mathbf{W}_1, \ldots, \mathbf{W}_m$, so that the trajectory matrix $\mathbf{X}$ can be rewritten as: $\mathbf{X} = \mathbf{W}_1 + \ldots + \mathbf{W}_m$ (see [9] for technical details.).

Step 4: Diagonal averaging
The purpose of diagonal averaging is to transform each new matrix $(\mathbf{W}_{i(i=1,\ldots,m)})$ to the form of a Hankel matrix, which can be subsequently converted to a time series.

The selection of SSA choices of L and r, such that they are optimal, is of critical importance to achieving the most accurate forecasts using the SSA technique. Thus, in this paper, we use the RMSE criterion (see, Section 4.2) to determine the optimal L for decomposing the UK income tax series, and the optimal r for reconstructing the filtered series to be used for forecasting approach. Thus, a combination of L and r which gives the lowest RMSE, provides the optimal choices for the SSA technique.[1].

There are two versions of SSA forecasting approaches called Recurrent (RSSA) and Vector (VSSA). The achieved optimal choices for VSSA and RSSA are presented in Table 1. Below, a concise description of these two versions of forecasting algorithms has been provided.

3.1.1 Recurrent SSA

With the reconstructed series, $\widetilde{y}_i$ $(i = 1, \ldots, N)$, the recurrent SSA forecasts can be obtained as follows:

[1]The optimal SSA code used in this study is available upon request.

Let denote π_i the last component of the eigenvector U_i $(i = 1, \ldots, r)$ and $v^2 = \pi_1^2 + \ldots + \pi_r^2 < 1$. Moreover suppose for any vector $U \in \mathbf{R}^L$ and denote by $U^\triangledown \in \mathbf{R}^{L-1}$ the vector consisting of the first $L-1$ components of the vector U. Then, the h-step ahead forecasts are given by

$$y_i = \begin{cases} \widetilde{y}_i & \text{for } i = 1, \ldots, N \\ \sum_{j=1}^{L-1} \alpha_j y_{i-j} & \text{for } i = N+1, \ldots, N+h \end{cases} \tag{1}$$

where α_j's are wrapped in a vector $A = (\alpha_1, \ldots, \alpha_{L-1})$ that can be computed by $A = (1 - v^2)^{-1} \sum_{i=1}^{r} \pi_i U_i^\triangledown$.

3.1.2 Vector SSA

Consider a matrix, Π, that is given by

$$\Pi = V^\triangledown (V^\triangledown)^T + (1 - v^2) A A^T,$$

where $V^\triangledown = [U_1^\triangledown, ..., U_r^\triangledown]$, v^2 and A are as defined above. Defining a linear operator $\theta^{(v)} : \mathfrak{L}_r \mapsto \mathbf{R}^L$ by the following formula

$$\theta^{(v)} U = \begin{pmatrix} \Pi U^\triangledown \\ A^T U^\triangledown \end{pmatrix}.$$

The h-step ahead forecasts, $y_{N+1}, ..., y_{N+h}$, can be computed in two steps:

In the first step we define a vector Z_i as follows:

$$Z_i = \begin{cases} \widetilde{X}_i & \text{for } i = 1, \ldots, K \\ \theta^{(v)} Z_{i-1} & \text{for } i = K+1, \ldots, K+h+L-1 \end{cases} \tag{2}$$

where, $\widetilde{X}_i$'s are the reconstructed columns of trajectory matrix.

In the second step we construct the matrix Z, as $\mathbf{Z} = [Z_1, ..., Z_{K+h+L-1}]$, and make its diagonal averaging to obtain a series $y_1, ..., y_{N+h+L-1}$ that contain our VSSA forecasts.

3.2 Auto-Regressive Integrated Moving Average (ARIMA)

The Automatic-ARIMA which is an optimised version of Box and Jenkins (1970) ARIMA model referred to as *auto.arima* provided by the forecast package in the R software has been adopted in this study [10].

The three components (p, d, q) are the AR order, the degree of differencing, and the MA order. In Automatic-ARIMA, to determine the number of differences d, KPSS unit root tests, Augmented Dickey-Fuller (ADF) test or the Phillips-Perron (PP) unit root tests are applied which among them the KPSS tests is known to give better forecasts in comparison to the ADF and PP test [11, 12]. To determine the values for the order of autoregressive terms p, and the order of the moving average process q, the algorithm then minimises the Akaike Information

Criterion (AIC). The optimal model is chosen to be the model which represents the smallest AIC. If the time series is nonstationary, then within the ARIMA(p,d,q) process the value of $d \geq 1$ and the Automatic-ARIMA forecasting algorithm accounts for this by taking first differences of the data until the data is stationary. If the data is stationary, then no differencing is required, and so $d = 0$ [11]. Table 1 shows the ARIMA model parameters used for forecasting the total UK income tax.

A non-seasonal ARIMA model may be written as:

$$(1 - \phi_1 B - \ldots \phi_p B^p)(1 - B)^d y_t = c + (1 + \phi_1 B + \ldots + \phi_q B^q)e_t, \tag{3}$$

or

$$(1 - \phi_1 B - \ldots \phi_p B^p)(1 - B)^d (y_t - \mu t^d/d!) = (1 + \phi_1 B + \ldots + \phi_q B^q)e_t, \tag{4}$$

where μ is the mean of $(1 - B)^d(y_t$, $c = \mu(1 - \phi_1 - \ldots - \phi_p)$ and B is the backshift operator. In the R software, the inclusion of a constant in a non-stationary ARIMA model is equivalent to inducing a polynomial trend of order d in the forecast function. It should be noted that when d=0, μ is the mean of y_t. According to hyndman2007automatic, the seasonal ARIMA model can be expressed as:

$$\Phi(B^m)\phi(B)(1 - B^m)^D(1 - B)^d y_t = c + \Theta(B^m)\theta(B)\epsilon_t, \tag{5}$$

where $\Phi(z)$ and $\Theta(z)$ are the polynomials of orders P and Q, and ϵ_t is white noise. Note that here if c $\neq$ 0, there is an implied polynomial of order d+D in the forecast function. As mentioned previously, to determine the values of p and q the AIC of the following form is minimised:

$$AIC = -2log(L) + 2(p + q + P + Q + k), \tag{6}$$

where $k = 1$ if c $\neq$ 0 and 0 otherwise and L represents the maximum likelihood of the fitted model.

3.3 Exponential Smoothing (ETS)

In this study, we rely on the automated Exponential Smoothing (ETS) model provided in the forecast package in R. In this model, the error, trend and seasonal components are considered to determine the best exponential smoothing model from over 30 potential options following optimizing initial values and parameters using the Maximum Liklihood Estimation (MLE) and selecting the best model according to the AIC. Those interested in the several ETS formula's that are evaluated through the forecast package when selecting the best model to fit the residuals are referred to Chapter 7, Table 7.8 in [11]. The ETS model parameters for forecasting UK income tax are reported in Table 1 where α, γ, σ are the ETS smoothing parameters.

3.4 Neural Networks (NN)

Our Neural Network (NN) models are estimated using an automatic forecasting model known as *nnetar* which is provided through the forecast package in R programming code. A detailed description of the model with explanation on the underlying dynamics can be found in [11]. The NN takes the form

$$\hat{y}_t = \hat{\beta}_0 + \sum_{j=1}^{k} \hat{\beta}_j \psi(x_t^{'}.\hat{\gamma}_j), \tag{7}$$

where x_t consist of p lags of y_t and the function ψ has the logistic form

$$\psi(x_t^{'}.\hat{\gamma}_j) = [1 + exp(-\hat{\gamma}_{j0} + \sum_{i=1}^{p} \hat{\gamma}_{ji}.y_{t-1})]^{-1} j = 1, \ldots, k \tag{8}$$

This form is often referred to as a one hidden layer feed forward NN model. As can be seen, the nonlinearity arises through the lagged y_t entering in a flexible way through the logistic functions of (8). The number of logistic functions included, namely k, is known as the number of hidden nodes.

The parameters in the NN model are determined according to a loss function embedded into learning algorithm. This loss function could be for example RMSE as explained below under Section 4.2. The *nnetar* function trains 25 networks by using random starting values. The average of the resulting predictions will be counted as the forecasted points. The NN model parameters are reported in Table 1 where p is the number of lagged inputs, P is the automatically selected value for seasonal time series, and k is the number of nodes in the hidden layer. It is of note that as evident in that table, in all cases the selected NN model has only k=1 hidden node, p=1 lags suggesting that for the UK income tax a simple network model outperforms the complex ones.

Table 1: Forecasting model parameters for total UK income tax.

h	ARIMA(p, d, q)	ETS(α, γ, σ)	NN(p, P, k)	VSSA(L, r)	RSSA(L, r)
1	(2,0,1)(0,1,1)	(0.23,0.30,0.06)	(1,1,1)	(24,13)	(34,14)
3	(2,0,1)(0,1,1)	(0.23,0.30,0.06)	(1,1,1)	(47,21)	(41,12)
6	(2,0,1)(0,1,1)	(0.23,0.30,0.06)	(1,1,1)	(44,13)	(42,12)
9	(2,0,1)(0,1,1)	(0.23,0.30,0.06)	(1,1,1)	(44,20)	(41,12)
12	(2,0,1)(0,1,1)	(0.23,0.30,0.06)	(1,1,1)	(38,16)	(39,12)

4 The Data and Measures for Evaluating Forecast Accuracy

4.1 The Data

This paper employs the monthly UK income tax data from January 2002 to September 2016. Table 2 provides the descriptives statistics of the data. According to that table, the average total monthly UK income tax between January 2002-September 2016 was 12370.62. The maximum value was recorded at 30271.69 in January 2016 and the minimum 6922.83 (in November 2002). The data was tested for normality using the Shapiro-Wilk test and it was found that at a p-value of 0.05 the data is not normally distributed. Also, according to the result of D'Agostino skewness test, at a p-value of 0.05, the data is positively skewed. An analysis of the kurtosis using the Anscombe-Glynn kurtosis test suggests that for this series the kurtosis is not equal to 3 and the series has a Leptokurtic distribution having a high probability for extreme values.

Table 2: Descriptive statistics: Total UK income tax (Jan. 2002 - Sep. 2016).

Mean	Min.	Max.	Std. Dev.	Skew.	Kurtosis
12370.62	6922.83	30271.69	4377.39	1.72	3.27

It should also be noted that the method of linear interpolation has been applied to deal with the two missing values in the time series (i.e. September and October 2004) [13].

The UK income tax series has also been evaluated for certain external shocks creating structural breaks and making the series nonstationary in mean and variance. According to Bai and Perron test for structural breaks, we can see that the time series has been affected by a structural break in December 2006 [14].

4.2 Measures for Evaluating the Forecast Accuracy

Root Mean Squared Error (RMSE)

Mean Squared Error (MSE) is a widely used measure to assess the quality of an estimator. The MSE always yields a non-negative value. The closer measure of MSE to zero defines the better performance of the estimator.

It is more common to use the Root mean square error (RMSE) which simply is the square root of MSE. RMSE is also a popular measure of accuracy to find the average of the squares of the errors or deviations. RMSE has the same unit of measurement as the quantity under study.

In comparing two estimators (e.g. two forecasting methods in this study), it is advised to adopt the Ratio of Root Mean Square Error (RRMSE) [15]. For example:

$$\text{RRMSE} = \frac{RSSA}{ETS} = \frac{\left(\sum_{i=1}^{N}(\widehat{y}_{T+h,i} - y_{T+h,i})^2\right)^{1/2}}{\left(\sum_{i=1}^{N}(\widetilde{y}_{T+h,i} - y_{T+h,i})^2\right)^{1/2}},$$

where, $\widehat{y}_{T+h}$ is the h-step ahead forecast obtained by RSSA, $\widetilde{y}_{T+h}$ is the h-step ahead forecast from the ETS model, and N is the number of the forecasts. If $\frac{RSSA}{ETS}$ is less than 1, then the SSA outperforms ETS by 1-$\frac{RSSA}{ETS}$ percent.

Mean Absolute Percentage Error (MAPE)

The MAPE measure is also quoted in this paper as it is a widely understood criterion for evaluating forecast accuracy. In brief, the lower the MAPE result, the better the forecast.

$$\text{MAPE} = \frac{1}{N}\sum_{t=1}^{N}|100 \times \frac{y_{T+h} - \widehat{y}_{T+h,i}}{y_{T+h}}|,$$

where y_{T+h} represents the actual data corresponding to the h step ahead forecast, and $\widehat{y}_{T+h,i}$ is the h step ahead forecasts obtained from a particular forecasting model.

5 Empirical Results

In this paper, the $\frac{2}{3}^{rd}$ of the data has been considered as in-sample for model training and the $\frac{1}{3}^{rd}$ of the data has been set aside as out-of-sample for evaluating the forecasting accuracy. To effectively evaluate the performance of the forecasting models both in the short and long run, the data was forecasted at horizons of $h = 1$, 3, 6, 9 and 12 steps ahead which correspond to 1, 3, 6, 9 and 12 months ahead. Table 3 reports the RMSE results for the out-of-sample forecasts of total UK income tax using VSSA, RSSA, ARIMA, ETS and NN.

Table 3: Out-of-sample RMSE results for total UK income tax.

h	VSSA	RSSA	ARIMA	ETS	NN
1	795.63	800.15	858.96	1168.78	6276.50
3	751.90	698.86	820.99	1207.35	6007.29
6	731.60	667.90	762.11	983.85	7020.22
9	693.58	676.59	794.98	1253.00	6774.38
12	711.44	706.91	837.98	1167.54	7200.37
Average	736.83	710.08	815.01	1156.10	6655.75

based on the RMSE criterion, RSSA outperforms VSSA, ARIMA, ETS and NN by recording the lowest forecasting error at all horizons except for $h = 1$ which VSSA reports a better result. The nonparametric ETS model is the second worst performer and outperforms the NN model by 82%. Also, the less variation seen in the results reported by SSA confirms that SSA is the most stable model in this case. The higher performance of the SSA technique could be explained by its capability in reducing the noise level of a time series. SSA is a specialised filtering method with the ability to decompose the UK income tax series and analyse the eigenvalues to precisely identify and divide the noise from the signal. The appropriateness of the separation between signal and noise obtained using SSA was confirmed by the very small values of w-correlation, confirming that the signal and its corresponding noise are almost w-orthogonal.

According to the MAPE criterion reported in Table 4, it is evident that the NN model is the worst performer at all horizons with an overall average MAPE of 45.36%. The MAPE value of 5.5% reported for RSSA highlights the most accurate forecasts obtained by this model followed by VSSA as the second best model for forecasting UK income tax with an average MAPE of 5.73%. Accordingly, the superior performance of SSA technique at all forecasting horizons portrays SSA's capabilities of providing comparatively more accurate forecasts in both short and long run.

Table 4: Out-of-sample MAPE results for total UK income tax.

h	VSSA	RSSA	ARIMA	ETS	NN
1	6.11%	6.13%	6.13%	8.02%	42.58%
3	5.78%	5.34%	5.84%	8.56%	39.81%
6	5.29%	5.16%	6.03%	6.79%	47.11%
9	5.70%	5.29%	5.77%	9.16%	48.13%
12	5.77%	5.57%	6.17%	8.66%	50.54%
Average	5.73%	5.50%	5.99%	8.24%	45.63%

Having reported the RSSA as the best forecasting model for the UK income tax series, the RRMSE values are obtained as $\frac{RSSA}{Alternative method}$. To ensure that the achieved results are not chance occurrences, the RRMSE results are further tested for statistical significance using the modified Diebold-Mariano test [16]. Accordingly, we found that all the RRMSE results are statistically significant at all horizons at a p-value of 0.05 and thus provides solid evidence for the inferences we have made. However, the RRMSE results attributed to the $\frac{RSSA}{VSSA}$ are only significant at a p-value of 0.1 highlighting the close performance of these two versions of the SSA technique.

Table 5: The ratio of Out-of-sample RMSE results for total UK income tax.

h	$\frac{RSSA}{VSSA}$	$\frac{RSSA}{ARIMA}$	$\frac{RSSA}{ETS}$	$\frac{RSSA}{NN}$
1	1.01**	0.93*	0.68*	0.13*
3	0.93**	0.85*	0.58*	0.12*
6	0.91**	0.88*	0.68*	0.10*
9	0.98**	0.85*	0.54*	0.10*
12	0.99**	0.84*	0.61*	0.10*
Average	0.96	0.87	0.62	0.11

Note:* indicates results are statistically significant based on Diebold-Mariano at $p = 0.05$.
** indicates results are statistically significant based on Diebold-Mariano at $p = 0.1$.

At the next step, the weighted correlation (w-correlation) statistic is used to show the appropriateness of the various decompositions achieved by RSSA According to [17], the w-correlation statistic shows the dependence between two series and can be calculated as:

$$\rho_{12}^{(w)} = \frac{\left(Y_N^{(1)}, Y_N^{(2)}\right)_w}{\parallel Y_N^{(1)} \parallel_w \parallel Y_N^{(2)} \parallel_w},$$

where $Y_N^{(1)}$ and $Y_N^{(2)}$ are two time series, $\parallel Y_N^{(i)} \parallel_w = \sqrt{\left(Y_N^{(i)}, Y_N^{(i)}\right)_w}$, $\left(Y_N^{(i)}, Y_N^{(j)}\right)_w = \sum_{k=1}^{N} w_k y_k^{(i)} y_k^{(j)} \quad (i, j = 1, 2)$, w_k=min$\{k, L, N-k\}$ (here, assume $L \leq N/2$).

Therefore, if the obtained w-correlation is close to 0, this confirms that the corresponding series are w-orthogonal and that the two components are adequately separable [17]. However, if the w-correlation between two reconstructed series are large, this confirms that the series are far from being w-orthogonal, and are therefore not very well separable and the components should be considered as one group.

Table 6: w-correlations between signal and residuals at different forecasting horizons.

Series	1	3	6	9	12
Total UK income tax	0.020	0.016	0.014	0.015	0.022

The w-correlation values reported in Table 6, further explains the outstanding performance of the RSSA model. According to the small values of that table, the RSSA forecasting algorithm is highly successful in separating the signal from the noise found in the total UK income tax series.

6 Conclusion

The literature on tax revenue forecasting and the related literature on budget forecasting focus on perceived political bias in forecasts. This results in a body of scholarly opinion in favour of forecasts that are produced independently of government and, in particular, a body of independent forecasts. Macro-economic models are widely used for forecasting overall tax revenues and, often, for revenues from those individual taxes that are most closely linked to overall economic activity. However, they lack sufficient precision to be used to forecast revenues from many individual taxes. Another limitation is the cost and complexity of building and maintaining a macro-economic model. For these reasons a number of authors have emphasized the need for effective direct (not embedded in a wider model) techniques for forecasting revenues from specific taxes. In this study we considered a number of different time series (mainly) direct forecasting techniques. Using UK monthly data for income tax receipts from January 2002 to September 2016, we considered the performance of five different forecasting techniques with respect to income tax receipts. Of these we found Recurrent Singular Spectrum Analysis (RSSA) to perform the best. RSSA outperforms the other forecasting techniques with respect to both Root Mean Squared Error (RMSE) and Mean Absolute Percentage Error (MAPE).

References

[1] Buettner, T., & Kauder, B. (2010). Revenue forecasting practices: differences across countries and consequences for forecasting performance. Fiscal Studies, 313-340.

[2] Bretschneider, S. I., Gorr, W. L., Grizzle, G., & Klay, E. (1989). Political and organizational influences on the accuracy of forecasting state government revenues. International Journal of Forecasting, 5(3), 307-319.

[3] Leal, T., Prez, J. J., Tujula, M., & Vidal, J. P. (2008). Fiscal forecasting: lessons from the literature and challenges. Fiscal Studies, 29(3), 347-386.

[4] Auerbach, A. J. (1999). On the performance and use of government revenue forecasts. National Tax Journal, 767-782.

[5] Chatagny, F., & Soguel, N. C. (2012). The effect of tax revenue budgeting errors on fiscal balance: evidence from the Swiss cantons. International Tax and Public Finance, 19(3), 319-337.

[6] Artis, M., & Marcellino, M. (2001). Fiscal forecasting: The track record of the IMF, OECD and EC. The Econometrics Journal, 4(1), 20-36.

[7] Jonung, L., & Larch, M. (2006). Improving fiscal policy in the EU: the case for independent forecasts. Economic Policy, 21(47), 492-534.

[8] Golyandina, N., Nekrutkin, V., & Zhigljavsky, A. (2001). Analysis of time series structure: SSA and related techniques. Chapman & Hall/CRC.

[9] Hassani, H. (2007). Singular Spectrum Analysis (SSA): Methodology and comparison. *Journal of Data Science, 5(2)*, 239–257.

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

[11] Hyndman, R. J. & Athanasopoulos G. (2013). Forecasting: principles and practice, OTexts: Australia. Available via: `www.OTexts.com/fpp`. [Accessed: 16 June 2014]

[12] Hyndman, R. J. (2014). Unit root tests and ARIMA models. Available via: `http://robjhyndman.com/hyndsight/unit-root-tests/`. [Accessed: 16 June 2014]

[13] Meijering, E. (2002). A chronology of interpolation: from ancient astronomy to modern signal and image processing. Proceedings of the IEEE, 90(3), 319-342

[14] Bai J., & Perron P. (2003). Computation and analysis of multiple structural change models. *Journal of Applied Econometrics, 18(1)*, 1–22.

[15] Hassani, H., Heravi, H., & Zhigljavsky, A. (2009). Forecasting European industrial production with Singular Spectrum Analysis. *International Journal of Forecasting, 25(1)*, 103–118.

[16] Harvey, D. I., Leybourne, S. J., & Newbold, P. (1997). Testing the equality of prediction mean squared errors. *International Journal of Forecasting, 13(2)*, 281-291.

[17] Hassani, H., Heravi, S., & Zhigljavsky, A. (2009). Forecasting European industrial production with singular spectrum analysis. International journal of forecasting, 25(1), 103-118.