A state-dependent linear recurrent formula with application to time series with structural breaks

Donya Rahmani1 | Damien Fay2

1School of Mathematics, University of Southampton, Highfield, Southampton, UK
2Retail Science Lead, Infor Retail, Atlanta, Georgia, USA

Correspondence Donya Rahmani, School of Mathematics, University of Southampton, Highfield, Southampton, UK. Email: d.rahmani@soton.ac.uk

Abstract
An underlying assumption in Multivariate Singular Spectrum Analysis (MSSA) is that the time series are governed by a linear recurrent continuation. However, in the presence of a structural break, multiple series can be transferred from one homogeneous state to another over a comparatively short time breaking this assumption. As a consequence, forecasting performance can degrade significantly. In this paper, we propose a state-dependent model to incorporate the movement of states in the linear recurrent formula called a State-Dependent Multivariate SSA (SD-MSSA) model. The proposed model is examined for its reliability in the presence of a structural break by conducting an empirical analysis covering both synthetic and real data. Comparison with standard MSSA, BVAR, VAR and VECM models shows the proposed model outperforms all three models significantly.

KEYWORDS
forecasting structural breaks, linear recurrent formula, multivariate singular spectrum analysis, multivariate time series, state-dependent models, structural break, time series forecasting

1 | INTRODUCTION

This paper examines an adaptation of the Singular Spectral Analysis (SSA) model in the presence of a structural break. Time series are well known to exhibit abnormal behaviour, such as clustering of outliers, non-linearities, and the particular focus here, structural breaks. Structural breaks may be conceptualised in two ways, the time series fluctuates between states of equilibrium to a state of recovering from a shock, or as a result of natural limit cycles within the economic system as proposed in Beaudry et al. (2015). Real world time series are also known to exhibit non-stationary and heteroscedastic behaviour which can render standard linear time series analysis methods (e.g., ARIMA) inappropriate (Kantz & Schreiber, 2004). This has led to the development of several non-linear and embedding techniques and motivated the development of SSA (Broomhead & King, 1986). SSA is a non-parametric modelling technique which makes no underlying assumptions about the distribution or stationarity of the time series and so can handle non-linearity by removing noise from the time series using eigen decomposition methods (Golyandina et al., 2001).

SSA assumes that a time series evolves according to an underlying state space system, with the movement in that state space described by a linear recurrent formula (LRF) (Golyandina et al., 2001). In addition, an underlying assumption in SSA is that an observed time series is one realisation from a single LRF. However, this usually does not account for the presence of a structural break.
Either the structural break dynamics are not apparent due to an insufficient history or the structural break is part of another process that also evolves the state space, albeit at a rapid pace. To account for the structural break while continuing to use SSA one must therefore assume an LRF exists that determines the evolution before the break, the break itself in which evolution occurs due to a different process/shock, and the evolution after the break is described either by a new LRF or the old LRF (which is starting from a new point in the state space). Ignoring the structural break and estimating the LRF from the entire data incurs forecast performance degradation for reasons described in Theorem 1. The solution proposed here is to change the LRF from a fixed recurrent equation to an adaptive recurrent equation. Such a model is known as a state-dependent model (SDM). The SDM was introduced by Priestley (1980) to target shifts in the dynamics of a time series from one state to another in the presence of a shock and, to the best of our knowledge, its use in SSA is novel.

The aim of this paper is to incorporate the SDM into the SSA framework to prevent shifts in a time series state from introducing a persistent bias in model forecasts (as will be shown later). In addition, we present results in the multivariate setting in which several inter-related economic time series are to be forecast together. The resulting model is shown to greatly improve Multivariate SSA (MSSA) (Golyandina et al., 2001) in the presence of a structural break, while still maintaining MSSA’s ability to model a time series in the embedded time domain with noise rejection.

The proposed model is tested on synthetic data and the monthly Industrial Production Indicators (IPI) of France, Germany and the UK and monthly Unemployment Rate (UR) series for the EU and USA. To provide a better understanding of its performance with respect to both standard MSSA and classical autoregressive models, the results are benchmarked against the standard VAR, Bayesian VAR (BVAR), VECM and MSSA models. The remainder of the paper is organised as follows: Section 2 presents related work, Section 3 extends MSSA and proposes the SD-MSSA (State-Dependent Multivariate Singular Spectrum Analysis) model. Section 4 presents the empirical results and discusses the findings. Finally, Section 5 draws conclusions and suggests avenues for future work.

2 | RELATED WORK

In this section we give a brief overview of SSA and focus on common models and features derived from time series in the presence of a structural break. In the last two paragraphs, we discuss research related to the core addition; state dependent model.

SSA was first introduced in the 80’s by Broomhead and King (1986) and Fraedrich (1986) as a time series approach to model dynamic behaviour through embedding delays as inspired by Taken’s Theorem. Since then, the theoretical foundations of SSA for time series analysis and also forecasting have been established in a series of papers (Allan & Smith, 1996; Danilov & Zhigljavsky, 1997; Golyandina et al., 2001; Groth & Ghil, 2011; Golyandina et al., 2001; Kapl & Mueller, 2010; Patterson et al., 2011; Oropeza & Sachchi, 2011; Vautard et al., 1992; You et al., 2000; Zhang & Hui, 2012; Zhao et al., 2011), and the references therein.

The embedded representation of a time series used in SSA decomposes it into a set of data-adaptive orthonormal components. These components can be projected into a lower dimension and then reconstructed to form a smoother time series, which can then be used for explaining structure and forecasting (Viljoen & Nel, 2010). Of particular interest, this paper focuses on an extension of SSA called Multivariate SSA. MSSA seeks commonalities between the embedding delay covariance matrices of several time series which are called common modes. Essentially, the applicability of MSSA centres on the existence (or not) of these common modes. It is assumed that common mode exists which explain simultaneous variations in a system of multiple time series (Viljoen & Nel, 2010). For example, when the modes in the time series are well-matched to each other then the MSSA performance can improve significantly; especially when a single time series alone is not sufficient to obtain good (univariate) estimates of the modes. As noted in Viljoen and Nel (2010) two or more time series may share common modes but may also exhibit specific modes (specific to just one time series).

Given the common modes characteristic, MSSA has been found to perform well in some domains such as noise reduction in EEG signals (Groth & Ghil, 2011), anomaly detection and traffic patterns in computer networks (Babaie et al., 2014), spatial statistics (Awichi & Müller, 2013), reconstruction of shared oscillations and analysis of oscillatory modes of several time series (Groth & Ghil, 2011), phase synchronisation analysis (Groth et al., 2015), and many others.

Although in all the examples mentioned above MSSA is used as a filtering and reconstruction technique, it can also be used as a powerful forecasting technique. For example, Patterson et al. (2011) show that MSSA, as a non-parametric tool, can be useful for forecasting real-time data subject to a revision process, when compared to those time series models which involve restrictive
assumptions such as linearity, normality and stationarity. Moreover, applying MSSA on the UK’s IPI time series was found to improve forecasting accuracy over standard parametric models like VAR and ARMA. More comparisons between MSSA (and SSA) forecasting and other classical models like autoregressive models (ARIMA and GARCH) and random walk models, through a wide range of financial and economic time series, can be found in previous works (Hassani et al., 2013; Hassani et al., 2012; Heravi et al., 2004; Kapl & Mueller, 2010; Osborn et al., 1999; Patterson et al., 2011; Skare & Buterin, 2015; Zhang & Hui, 2012).

MSSA has been also combined with other methods to target specific features of particular datasets. For example, Menezes et al. (2014) combine MSSA with Periodic AutoRegressive models (PAR), to forecast the monthly average wind speed of two regions of Northeast Brazil (Hydrological time series). Their outcome shows PAR(P)-MSSA is far superior than the two others (PAR(P) and PAR(P)-SSA), since MSSA takes into account the spatial dependence between the two stations. Multi-scale SSA (MS-MSSA) is yet another improvement on MSSA and proposed by You et al. (2000). They define a moving window proportional to series length and use a wavelet transform to detect regular part of the lag-covariance matrix. Despite the fact that empirical results show a similar behaviour of MS-MSSA and wavelet analysis, MS-MSSA shows faster performance on tracking sharper regime transitions in the evolution of a non-linear system than either MSSA or wavelet analysis (You et al., 2000).

MSSA forecasting is conducted via the recurrent coefficients of an LRF (Formula 2 in Section 3). There exist several variations of the LRF, such as the nearest subspace LRF approach when there is an approximate separability in which the selected linear space cannot be the trajectory space of the time series anymore (Golyandina et al., 2001) and the minimal recurrent formula, which removes the extraneous roots of the characteristic polynomial of the LRF (reduction method).

Previous works (Gelman et al., 2013; Hendry & Mizon, 2005; Kim et al., 2004; Pesaran et al., 2006) discuss different approaches to dealing with structural breaks in several econometric models. Hendry and Mizon (2005) point out that when a structural break occurs forecasts invariably become quite bad as usually the underlying assumptions about the data generating process (DGP) become invalid. It is then suggested that a reset type approach works best in which one re-estimates the model states given only the new data (e.g., an intercept can be reset as it can be estimated from the data directly). However, for SSA one has a problem as there is no clear relationship between the LRF coefficients and the data (the SVD clouds the relationship, Section 3.2) and in addition we would like to retain information pre-break especially as it relates to the seasonal pattern of the time series. For the LRF (Golyandina et al., 2001) find that in the presence of abrupt changes, the number of the perturbed extraneous roots increases significantly which results in a less precise forecast. Furthermore, while the time series before and after the break may locally follow a LRF (or two separate LRF’s) the combined time series may not. To detect such changes, Golyandina et al. (2001) defined a heterogeneity function and an eigentriple rearrangement by measuring the discrepancy between two (or more) different states of a time series. They have also examined the root function of the characteristic polynomial of the LRF. This gives a dynamic description of the sequence of the homogeneous part of the series which is affected by the eigentriple rearrangement and structural changes. Although the system post-break evolves according to the same dynamics, MSSA is unable to estimate these parameters correctly (as we shall prove in Theorem 4.1). Note that it is not that the data pre- and post-structural break are necessarily samples from a different process but rather estimating the LRF does not take the break into account and this leads to a biased estimate of the model. Therefore, we require not a fixed state space system as in SSA but one which adapts to the changes in the system as in State Dependent Models (SDM).

The SDM (Priestley, 1982) uses a sequential structure to distinguish between states transition and can be used in forecasting and detecting non-linearity in time series. According to Priestley (1980) using state-dependent models has advantages in two aspects. First, they can be used directly with prediction and forecasting, second, since they can be fitted to data without specific prior assumptions about the form of the non-linearity, they can provide an overview of the inherent non-linear characteristic in a given series (Priestley, 1982). An SDM operates by splitting the state spaces into a large number of small segments, and models the process as locally linear inside each segment.

In this paper, the movement of LRF parameters are tracked by using an SDM in such a way that the parameters can change not only over time but also over the states of the system. In other words, the LRF’s behaviour in the presence of a structural break is explained by a non-linear function connecting the current states to its previous states (and the previous states of other series) and random shocks. The whole process is called SD-MSSA since the parameters are no longer fixed for all time series and will be updated recursively via the Extended Kalman Filter (EKF). In the next section, we briefly describe the MSSA model and then present theoretical foundations of SD-MSSA.
3 \ | \ METHODOLOGY

In this section, we first present MSSA and the LRF. Then, after having shown that SSA is biased in the presence of a structural break, we propose the state dependent LRF to alleviate the bias.\(^1\)

3.1 | MSSA model

Let \( Y = \{ Y_i^{(m)} : m = 1, \ldots, M \} \) be a multivariate time series with \( M \) series of length \( N \). The first step in MSSA involves embedding the time series into what is called the trajectory matrix, \( X_M \), with columns \( X_i^{(m)} = [y_i^{(m)} \ldots y_i^{(m)}] \), \( i = 1, \ldots, N - L + 1 \). Here, \( L \) is called embedding dimension or window length and is the maximum delay applied to the series to form the embeddings. These embeddings\(^2\) are then stacked to form the block trajectory matrix as

\[
X_M = [X^{(1)} \ldots X^{(M)}] = \\
\begin{pmatrix}
Y_1^{(1)} & \cdots & Y_{N-L+1}^{(1)} \\
Y_1^{(2)} & \cdots & Y_{N-L+1}^{(2)} \\
\vdots & \ddots & \vdots \\
Y_1^{(M)} & \cdots & Y_{N-L+1}^{(M)}
\end{pmatrix}
\]

The next step involves estimating a lower dimensional reconstruction of \( X_M \) via the singular value decomposition as \( \hat{X}_M = [\hat{X}^{(1)} \ldots \hat{X}^{(M)}] = \sum_{j=1}^r U_{Mj}^T U_{Mj} X_M \), where \( U_{Mj} = (u_{1j} \ldots u_{Lj})^T \) is the left singular vectors of \( X_M X_M^T \) with \( \lambda_j \) as the corresponding eigenvalues and \( r \) is the number of components retained. To project \( X_M \) back into the time domain, a Hankelization process (averaging across the anti-diagonals of the matrix) can be applied on each of the blocks as follows, \( \hat{X}_N^{(m)} = \mathcal{H}(\hat{X}^{(m)}) ; m = 1, \ldots, M \), which transforms \( Y_N^{(m)} \) into a new time series of the same length. The final step is making a forecast from the smoothed time series. The underlying assumption here is that \( \hat{y}_N^{(m)} \) satisfies a LRF. Assume the data point \( \hat{y}_N^{(m)} \) becomes available for the \( m \)th time series this would result in adding a new column to the trajectory matrix of \( X^{(m)} \) with \( L - 1 \) known elements, \( \{ y_{N-L+2}^{(m)} \ldots y_{N+1}^{(m)} \}^T \). Those first \( L - 1 \) rows of \( X^{(m)} \) can form a linear projection onto the \( L \)th row as

\[
y_{N+1}^{(m)} = \phi_1 y_{N}^{(m)} + \cdots + \phi_{L-1} y_{N-L+2}^{(m)},
\]

which is also defined as an \( L \)-continuation of the series where the coefficients vector \( (\phi_1, \ldots, \phi_{L-1}) \) is applied to the linear space \( \mathcal{L}_r = \text{span}\{ U_{Mj} \} \). Note that the choice of \( L \) depends on the structure of the data and no single rule for \( L \) can cover all applications. However, in general \( L \) should be proportional to the periodicity of the data, large enough to obtain sufficiently separated components but not greater than \( N/2 \). Full discussion of the choice of this parameter is given in Golyandina et al. (2001).

To obtain the coefficients vector, \( (\phi_1, \ldots, \phi_{L-1}) \), let \( U_{Mj}^\tau \) denotes the vector of the first \( L - 1 \) rows of the eigenvectors \( U_{Mj} \), and \( \pi_{Mj} \) denotes the last row of the eigenvectors \( U_{Mj} \), then if \( e_L \notin \mathcal{L}_r \) it can be simply proved that \( \pi_{Mj}^2 + \pi_{Mj}^3 + \cdots + \pi_{Mj}^L < 1 \), see (Golyandina et al., 2001). It is also shown that there exists a unique vector, \( (\hat{y}_N^{(m)}, \ldots, \hat{y}_{N-L+2}^{(m)})^T \in \mathcal{L}_r \), which corresponds to any vector \( (Y_N^{(m)}, \ldots, Y_{N-L+2}^{(m)})^T \) with coefficients, \( h_j \), such that

\[
(Y_N^{(m)}, \ldots, Y_{N-L+2}^{(m)})^T = (\hat{y}_N^{(m)}, \ldots, \hat{y}_{N-L+2}^{(m)})^T + e_L^T \sum_{j=1}^r h_j U_{Mj}.
\]

Multiplying both sides by \( e_L \) and \( U_{Mj} \) respectively, results in

\[
\sum_{j=1}^r h_j = \sum_{j=1}^r (\hat{y}_N^{(m)}, \ldots, \hat{y}_{N-L+2}^{(m)})^T U_{Mj}^\tau + \hat{y}_N^{(m)} \sum_{j=1}^r \pi_{Mj},
\]

and

\[
\hat{y}_N^{(m)} = \sum_{j=1}^r h_j \pi_{Mj} = \sum_{j=1}^r (\hat{y}_N^{(m)}, \ldots, \hat{y}_{N-L+2}^{(m)})^T U_{Mj}^\tau \pi_{Mj} + \hat{y}_N^{(m)} \sum_{j=1}^r \pi_{Mj}^2
\]

\[
= \left[1 - \sum_{j=1}^r \pi_{Mj}^2 \right]^{-1} \left[ \sum_{j=1}^r (\hat{y}_N^{(m)}, \ldots, \hat{y}_{N-L+2}^{(m)})^T U_{Mj}^\tau \pi_{Mj} \right],
\]

which means the LRF coefficients, \( \Phi = (\phi_{L-1}, \ldots, \phi_1)^T \), can be expressed as

---

\(^1\)We caution the reader that although SSA and AR/VAR models share similarities, they are fundamentally quite different and in our experience drawing comparisons can hinder understanding.

\(^2\)In principle, because of the embedding theorem (Takens, 1981), the time delayed version of a time series is generically sufficient to reconstruct the dynamics of the underlying systems if enough delayed coordinates are used (Cao et al., 1998).
Equation (2) is the core of this paper and has many interesting properties (see Golyandina & Zhigljavsky, 2013). First, it can be used with the chain rule of forecasting time series to give multi-step ahead forecasts. Second, it is not a typical AR($L$ - 1) model as the residual is not iid. More importantly, $\Phi$ captures the dynamics of the time series (specifically, it defines the linear recurrence Golyandina and Zhigljavsky (2013)) and gives an optimal multi-step ahead forecast while an AR model typically gives the optimal 1-step ahead forecast (this is the major advantage of SSA as a forecasting tool). An example, to be used later, is shown in Figure 1a which shows three time series with very different characteristics, but yet all have been generated from the exact same LRF.

Before turning to the theoretical effect of a structural break on the estimation of an LRF, we give a brief intuitive illustration of an LRF’s behaviour in the presence of a structural break. Figure 1b shows a cloud map of 100 bootstrapped LRF estimates of $\phi_1(t)$, estimated at time $t$, for the French Vehicle time series (used later in Section 4.2.2). It is evident that, before the structural break (at time 226), the LRF’s distribution has a lower variance and is more concentrated around the mean (white area). One would expect that after a break that the variance will temporarily increase; however, the increased variance appears to persist and this is worrying. This would seem to imply that once a structural break has occurred we should discard pre-break data which is obviously an unsatisfactory outcome.

$\Phi = \frac{1}{\sqrt{C_0}} \sum_{j=1}^{r} \sum_{m} U_{Mj}^{(j)}.$ (2)

3 A trivial example is when $r = 1$, that is, the residual has seasonal components which are auto-correlated, that is, not iid.

3.2 Linear recurrent formula in the presence of a structural break

As shown above, the LRF is the main vehicle through which future forecasts are obtained. In the following theorem we assume a single LRF dictates the evolution of a time series. We then examine the effect of a structural break which affects only the mean of the time series. It is convenient mathematically to assume the break affects the series before the break instead of after which is justified as both segments follow the same LRF.

**Theorem 1** (LRF Theorem). Let $\bar{Y} \in \mathcal{R}^{N \times M}$ be a multivariate time series which follows a LRF without any structural breaks. Assume that the observed time series is a back-shifted version of $\bar{Y} \in \mathcal{R}^{N \times M}$ such that $Y_m = \bar{Y}_m - \delta_m \ \forall t < \tau_m$, where each time series experiences a structural shift in the mean, $\delta_m$, at time $\tau_m$. Let $C_X$ and $\bar{C}_X$ denote the covariance matrices of the trajectories of $Y$ and $\bar{Y}$ respectively. Then the deviation between the first eigenvalue of $C_X$ and $\bar{C}_X$ is $\lambda_1 - \bar{\lambda}_1 \approx TL(Q^2 - 2Qy)$. Furthermore, this deviation decreases as $O(k/N)$.

**Lemma 2.** A time series, $y^{(m)}_1$, with an intercept, $c^{(m)}$ implies $\bar{C}_X$ has an eigenvector $u_1 \approx [1,1,...,1]^T/\sqrt{L}$ with corresponding eigenvalue $\lambda_1 \approx L(N - L + 1)\sum mc^{(m)}$. (2)

**Proof.** First note that for $t > \max(\tau_m)$ all structural shifts have occurred and the observed time series (by assumption) is a linear continuation of $\bar{Y}$, not $Y$. Thus, $\bar{C}_X$ contains the correct covariance structure for the time series going forward. The covariance matrix of the trajectory matrix can be
explicitly written in terms of the original time series without any structural shifts, \( y_k^{(m)} \), as

\[
\bar{c}_{ij} = \begin{cases} 
\sum_{m=1}^{M} \sum_{k=1}^{N-L+1} (\bar{y}_k^{(m)} - \delta_m 1_{A_m})^2, & i = j, \\
\sum_{m=1}^{M} \sum_{k=1}^{N-L+1} \bar{y}_k^{(m)} \bar{y}_{k+i-j}^{(m)}, & i \neq j.
\end{cases}
\]

(3)

Note that \( C_X \) is different from \( \tilde{C}_X \), in just a shift in the mean for the first \( \tau_m \) elements and so can be expressed as

\[
c_{ij} = \begin{cases} 
\sum_{m=1}^{M} \sum_{k=1}^{N-L+1} (\tilde{y}_k^{(m)} - \delta_m 1_{A_m})^2, & i = j, \\
\sum_{m=1}^{M} \sum_{k=1}^{N-L+1} \tilde{y}_k^{(m)} (\tilde{y}_{k+i-j}^{(m)} - \delta_m 1_{A_m}), & i \neq j.
\end{cases}
\]

(4)

where \( l = \min(i, j) \) and \( 1_{A_m} \) is a step function with \( 1_{A_m} = 1 \) for \( A_m = \{ t < \tau_m \} \) and zero otherwise. Note that structural break by definition takes the system from one mean to another such that \( \delta_m > \delta_k \). Expanding these terms results in

\[
E(C_X) \approx E(\tilde{C}_X) + (T Q^2 - 2T \tilde{y}) J_{L \times L} \approx E(\tilde{C}_X) + TL(Q^2 - \tilde{y}) \frac{1}{\sqrt{L}} 1_L 1_L^T \frac{1}{\sqrt{L}},
\]

(7)

where \( T \) is a row vector with the numbers of samples prior to the structural breaks, \( Q \) is a column vector of amplitudes of the structural breaks, \( 1_L \) is vector of ones with \( L \) elements, and \( J_{L \times L} \) is a square matrix of ones. Thus, the net effect of a structural break is that the first eigenvalue of the covariance matrix has a bias of approximately \( TL(Q^2 - Q \tilde{y}) \). From Lemma 2, below, this bias can be seen to tend to zero with \( N \) as

\[
\frac{\lambda_1 - \bar{\lambda}_1}{\bar{\lambda}_1} \approx \frac{TL(Q^2 - 2Q \tilde{y})}{(N - L + 1)\sum_m c^{(m)}} = O(k/N),
\]

(8)

where \( k \) is some constant.

Theorem 1 states that the first eigenvalue of \( C_X \) will contain a bias which slowly decays as new samples arrive (see Figure 2b for an empirical example of how this affects the LRF coefficients). For this reason a dynamic LRF is proposed based on a state dependent model.

### 3.3 State-dependent format of linear recurrent formula parameters

In this section, the LRF parameters are viewed as state parameters which are allowed to recursively evolve based on the observations. Let \( Y_N^{(m)} = \{ y_N^{(m)}, \ldots, y_N^{(m)} \} \) be the state vector of the \( m \)th time series. Following from (Priestley, 1982) the propagation coefficients are then defined as a function of \( Y_N^{(m)} \), so equation (1) takes the following form

\[
y_N^{(m)} = \phi_1(Y_N^{(m)}) y_N^{(m)} + \phi_2(Y_N^{(m)}) y_{N-1}^{(m)} + \ldots + \phi_{L-1}(Y_N^{(m)}) y_{N-L+2}^{(m)}.
\]

(9)

Note that in contrast to the LRF in MSSA, which assumes the same LRF coefficients for all series, SD-LRF allocates
different LRF coefficients to each series. In addition, \( \{ \phi_u(Y^{(m)}_N) \} = \{ \phi^{(m)}_u \} ; u = 1, \ldots, L - 1 \), are assumed to be analytic functions of \( Y^{(m)}_{N+1} \) which change smoothly over time and so can be expressed via a Taylor’s series expansion as

\[
\phi^{(m)}_{u,N+1} = \phi^{(m)}_{u,N} + \Delta \phi^{(m)}_{u,N+1} - \phi^{(m)}_{u,N+1-u}, \quad u = 1, \ldots, L - 1,
\]

where \( \Delta \phi^{(m)}_{u,N+1} = y^{(m)}_{N+1-u} - y^{(m)}_{N+1-(u-d)} \), \( y_{N+1,u} \) is the gradient, and \( d \) is the seasonal length of the time series. The gradient parameters \( \phi^{(m)}_{u,N+1,1}, \ldots, \phi^{(m)}_{u,N+1,L-1} \) are unknown hyper-parameters which are assumed to change slowly following a random walk as

\[
\Gamma^{(m)}_{N+1} = \Gamma^{(m)}_N + V^{(m)}_{N+1}, \quad (11)
\]

while \( \Gamma^{(m)}_{N+1} = (\gamma^{(m)}_{N+1,1}, \ldots, \gamma^{(m)}_{N+1,L-1}) \) and \( V^{(m)}_{N+1} \) is a sequence of independent matrix-valued random variables such that \( V^{(m)}_{N+1} \sim N(0, \Sigma_{V^{(m)}}) \). With some modifications, equation (1) can be rewritten in state space form in which the state-vector is no longer \( Y^{(m)}_N \), but is replaced by the state-dependent coefficients augmented with the unknown gradients

\[
\theta^{(m)}_{N+1} = [\phi^{(m)}_{N+1,1}, \ldots, \phi^{(m)}_{N+1,L-1}, \gamma^{(m)}_{N+1,1}, \ldots, \gamma^{(m)}_{N+1,L-1}]^T,
\]

where \( \theta^{(m)}_{N+1} \) is the vector of all unknown parameters of the model (and of course a function of \( Y^{(m)}_N \)). Finally, replacing \( \phi_u \) with (12) in (9) results in the following state space model

\[
y_{N+1} = H^*_{N+1} \theta^{(m)}_{N+1} + \epsilon_{N+1}, \quad \theta^{(m)}_{N+1} = F^*_N \theta^{(m)}_N + W_{N+1},
\]

where \( H^*_N = (y_N, \ldots, y_{N-L+2}, 0, \ldots, 0) \) is an inherent characteristic of the system capturing its motion. \( F^*_N = F^*_N(Y^{(m)}_{N+1}) \) is a state transition matrix characterised by a corresponding block diagonal form and open to modification over time. \( W_{N+1} \) is the measurement noise and \( \epsilon_i \) is the process noise. The state vector is then estimated using the Extended Kalman filter (see Appendix A1 for the full algorithm).
4 | EMPIRICAL RESULTS

In this section, we perform experiments on (i) synthetic and (ii) real time series. In the first set of experiments, we begin with a bivariate time series model generated from a VAR(2) examining the forecasting performance of the LRF in the presence of structural breaks and results in detail. The analysis is then extended to a more complicated Data Generating Process (DGP).

The second set of experiments is conducted on three sets of real time series, IPI: from I.N.S.E.E (Institute National de la Statistique et des Etudes Economiques) for France, from Statistisches Bundesamt Wiesbaden for Germany, from the Office for National Statistics (ONS) for the UK and UR: from Eurostat for 30 countries in Europe and from Bureau of Labour statistics for the 48 states of the USA.

4.1 | Synthetic data experiments

In the first set of experiments, we simulate a multivariate time series from the VAR model, which has been found in several studies to efficiently model the dynamical behaviour of multivariate series (Lütkepohl, 2005).

Let \( Y_t = (y_{1t}, ..., y_{Mt}) \)' denote an \((1 \times M)\) vector of time series, thus the basic \(p\)-lag, VAR(p) model can be expressed as

\[
Y_t = \delta + \Pi_1 Y_{t-1} + \Pi_2 Y_{t-2} + ... + \Pi_p Y_{t-p} + \epsilon_k, \quad (14)
\]

where \( \{\Pi_i \in \mathbb{R}^{M \times M}\}_{i=1}^p \) are coefficient matrices, \( \epsilon_k \in \mathbb{R}^{M \times 1} \) is white noise zero mean vector process with covariance matrix \( \Sigma \), and \( \delta \) is an intercept. Next, we insert a structural break as a shift in the mean level of the multivariate time series simulated by the VAR model mentioned above. The length of each time series is 250 with the shift occurring at time 150. Our DGP consists of three models: (1) two time series have been simulated and shifted to the same level (Figure 2a), (2) two time series have been simulated but only one has been shifted (Figure 2d), (3) six time series have been simulated and each pair have been shifted to different levels, Figure 3a. For simplicity we used \( p = 2 \), which means an embedding dimension of two for both MSSA and SD-MSSA. The goal of this experiment is to show the exponential decay in the empirical LRF coefficients and show how SD-LRF can recursively track the changes and alleviate the exponential decay.

4.1 | DGP 1

When a structural break hits both time series at the same time, as shown in Figure 2a, the LRF coefficient suddenly jumps from 1.005 to 1.03 and then starts decaying slowly back to its pre-structural break values, Figure 2b. On the other side, SD-LRF’s coefficient grows comparatively faster, from 1.00 to 1.20 but decays exponentially as time moves on as can be seen from Figure 2c. Moreover, the response time for SD-LRF (10% to 90% of its final value) is 20 steps while for LRF this is far greater and is 70 steps.

4.1 | DGP 2

In the case that only one of the two time series has been shifted (Figure 2d) the LRF coefficient decays faster than DGP 1 and therefore, the response time drops from 70 steps to 60 steps (Figure 2e). We emphasise once again that MSSA finds one set of coefficients for both time series which is a trade off between computational cost and forecast accuracy. In contrast, SD-LRF finds two sets of coefficients and only the coefficient corresponding to the shifted time series decays while the other one remains stable (Figure 2e).

4.1 | DGP 3

The third simulation consists of 6 time series with different shift levels: zero-inflated, moderately inflated and highly-inflated shift, Figure 3a. Once again LRF coefficients display a slow decay over time (Figure 3b) whereas SD-LRF assigns coefficients specifically to the time series related to the changes which in almost all cases decay significantly faster than LRF. As shown in Theorem 1, the presence of a structural break in the mean, \( \delta \), is related to the eigenvector of \([1,1,...,1]^T / \sqrt{L} \) in \( U_l \). Consequently, the empirical LRF coefficients show an exponential decay and die out very slowly as we move away from the shift and it can be proportional to their linear continuation of the time series pre-break.

4.2 | Real time series experiments

In this section, we examine the performance of SD-LRF to forecast real time series with structural breaks.

4.2.1 | Data

The first datasets are the IPI time series for France, Germany and the UK. Apart from structural changes, IPI...
series share dynamical characteristics which make them quite suitable for multivariate modelling. The data represent eight major components of real industrial production in France, Germany and the UK. These are seasonally adjusted series which measure real output for all facilities located in the manufacturing of Electric and Gas (E&G) utilities, Chemicals, Fabricated Metals (F.M.), Vehicle, Food Products (F.P.), Basic Metals (B.M.), Electrical Machinery (E.M.) and Machinery. These indicators show movements in production output and highlight structural developments as well as short-term changes in the economy. Although we consider only those eight of the two digit industries in this study, these eight industries account for more than 50% of the total industrial production in each country. The same eight industries have been considered by Heravi et al. (2004) and Osborn et al. (1999). Table 1 shows a descriptive analysis of IPI for France, Germany and the UK.

In all cases, our sample period ends in February 2014. However, the data for France start from January 1990, for Germany from January 1991 and for the UK start from January 1998. Most series (France and Germany) present a long period of growth in the 1990s and up to the last recession of 2008–2009. For the UK, however, most series show a period of stagnation in the early 2000s and recession in 2008. For almost all the series, the steepest drop can be seen around 2008–2009, which is attributed to the banking crisis and subsequent recession.6

To test collinearity and cointegration between the series the Variance Inflation Factor (VIF)7 has been measured. Based on Table 1, almost half of the series are shown to be strongly correlated and a quarter of them

---

6We used the Bai and Perron test (Bai & Perron, 2003) for structural breaks. It tests the null hypothesis that there is no structural break in the data against the alternative that there is a known structural break at time $t$. We have reported the last break point in Table 1. As can be seen in all cases the estimated break date occurs during the 2008 economic downturn.

7VIF regresses an explanatory variable towards the rest of variables and subsequently measures the square of the multiple correlation coefficients by $VIF = \frac{1}{1-R^2}$. A rule of thumb is that if VIF goes above 10, then multi-collinearity is high while for $5<VIF<10$ multi-collinearity is moderate.
are moderately correlated. In particular France and German's economies show higher correlation possibly due to a larger industrial component than the UK's.

The second datasets are the EU Unemployment Rate (UR) which is a collection of quarterly series based on the quarterly results of the EU Labour Force Survey and taken from Eurostat. Unemployment is among a number of critical variables whose evolution has been continuously subject to close analysis by economic authorities and academics alike. Indeed, there is a large body of literature that deals with the estimation of (econometrics and time series) models aimed at understanding the determinants of this variable, and evaluating their ability to produce accurate forecasts; a few recent examples include, inter alia, Milas and Rothman (2008), Lahiani and Scaillet (2009), Fendel et al. (2011), Trendle (2002) and Ball et al. (2015).

Table 2 shows a descriptive analysis of the UR over the time period 1998 to 2017 (79 quarters for 30 EU Member States). Note that the UR is expressed in percentage. The first half of the data are used for training and the remaining are used as a hold-out set for forecast performance evaluation. As can be seen from Table 2 there are some missing periods for most of the countries. In this study, we used a cubic interpolation function in MATLAB to impute the missing values. Note that these quarterly time series are fairly short which make them suitable to evaluate SD-LRF performance on forecasting short time series. Most of the unemployment series presents a short period of fall in the 1990s and up to 2008–2009. Depending on the countries the UR's was seen to grow considerably during the 2008 recession. Table 2

---

Table 1: Descriptive analysis of industrial production indicators for France, Germany and UK

|       | Mean in-sample | Mean out-of-sample | Mean overall | SD  | Skew. | Last break point (BP) | SW (p value) | VIF |
|-------|----------------|--------------------|--------------|-----|-------|-----------------------|--------------|-----|
| France E&G | 290 84.74 | 96.03 | 87.47 | 9.74 | -0.49 | 155 | 0.00** | 6.71* |
| Chemical | 290 96.38 | 102.58 | 97.87 | 8.51 | -0.66 | 247 | 0.00** | 5.21* |
| F.M. | 290 116.88 | 102.55 | 113.42 | 11.93 | 0.03 | 226 | 0.00** | 71.94** |
| Vehicle | 290 112.75 | 95.86 | 108.68 | 25.68 | 0.24 | 225 | 0.00** | 10.28** |
| F.P. | 290 97.53 | 100.01 | 98.13 | 2.81 | -0.49 | 246 | 0.00** | 2.77 |
| B.M. | 290 119.39 | 101.50 | 115.07 | 12.08 | -0.24 | 226 | 0.00** | 80.17** |
| E.M. | 290 109.76 | 103.31 | 108.20 | 10.94 | 0.27 | 228 | 0.00** | 27.28** |
| Machinery | 290 114.26 | 106.32 | 112.35 | 14.40 | 0.45 | 228 | 0.00** | 13.45** |
| Germany E&G | 278 98.38 | 97.18 | 98.13 | 4.80 | 0.22 | 221 | 0.006 | 2.13 |
| Chemical | 278 82.72 | 99.53 | 86.23 | 11.67 | -0.03 | 233 | 0.00** | 40.57** |
| F.M. | 278 86.87 | 105.55 | 90.77 | 13.77 | 0.29 | 215 | 0.00** | 53.25** |
| Vehicle | 278 78.47 | 106.96 | 84.41 | 19.87 | 0.04 | 208 | 0.00** | 18.41** |
| F.P. | 278 83.77 | 99.89 | 87.13 | 10.51 | -0.15 | 205 | 0.00** | 28.59** |
| B.M. | 278 97.22 | 99.33 | 97.66 | 9.11 | 0.11 | 215 | 0.00** | 5.81* |
| E.M. | 278 71.69 | 103.17 | 78.26 | 20.12 | 0.38 | 215 | 0.00** | 47.64** |
| Machinery | 278 89.78 | 107.19 | 93.42 | 14.57 | 0.51 | 215 | 0.00** | 22.06** |
| UK E&G | 194 96.27 | 95.82 | 96.11 | 5.22 | -0.45 | 130 | 0.01** | 2.10 |
| Chemical | 194 118.12 | 102.98 | 112.74 | 8.68 | -0.57 | 130 | 0.00** | 6.04* |
| F.M. | 194 117.54 | 101.65 | 111.89 | 8.89 | -0.63 | 130 | 0.00** | 17.44** |
| Vehicle | 194 121.78 | 109.04 | 117.25 | 12.95 | -1.48 | 128 | 0.00** | 2.15 |
| F.P. | 194 100.57 | 100.90 | 100.69 | 2.49 | 0.52 | 121 | 0.00** | 1.69 |
| B.M. | 194 124.18 | 99.83 | 115.52 | 15.26 | 0.00 | 130 | 0.03 | 9.13* |
| E.M. | 194 122.36 | 100.62 | 114.63 | 13.58 | -0.07 | 132 | 0.01** | 4.78 |
| Machinery | 194 98.50 | 100.29 | 99.14 | 6.89 | -0.26 | 124 | 0.01** | 1.80 |

Note: SD (Standard deviation); Skew. (skewness); BP (Bai–Perron) test for structural break; SW (Shapiro–Wilk) test for normality; VIF (variance inflation rate).

**Results are statistically significant at \( \alpha = 0.01 \). High level of multi-collinearity.

*Results are statistically significant at \( \alpha = 0.05 \). Moderate level of multi-collinearity.

---

8https://appsso.eurostat.ec.europa.eu/nui/submitViewTableAction.do
shows that the UR in countries like Cyprus, Greece, Ireland, Portugal and Spain grows significantly after the 2008 recession while countries like Belgium, Latvia, Lithuania and Norway seem to have a stable UR. On the other hand, in all countries, the null hypothesis of a unit root at the 1% level confirms that these series are also non-stationary. Looking at the colinearity results, 28 out of 30 series are shown to be strongly correlated and the two remaining series are moderately correlated. This can be mainly related to the EU’s economy which has a larger impact as a single market.

The third datasets are the mainland USA monthly UR’s from January 1976 to December 2013 (456 months/48 states). The first 304 points (1976–2000) are used for training the model and the remaining points are used as a test set for forecast performance evaluation. Almost all states show a significant growth in unemployment during the 2008 recession, especially West Virginia, Louisiana

### Table 2: Descriptive analysis of unemployment rate for Europe

| Country          | N  | Mean in-sample | Mean out-of-sample | Mean overall | SD  | Skew. | Last break point (BP) | SW (p value) | VIF  |
|------------------|----|----------------|--------------------|--------------|-----|-------|----------------------|--------------|------|
| Austria          | 79 | 3              | 4.61               | 5.34         | 4.88| 0.84  | −0.58                | 59           | 0.00** 10.73** |
| Belgium          | 79 | 3              | 7.91               | 7.96         | 7.93| 0.77  | −0.31                | 57           | 0.07 8.15* |
| Bulgaria         | 79 | 8              | 12.15              | 10.31        | 11.47| 3.99  | 0.54                 | 42           | 0.02* 41.98** |
| Croatia          | 79 | 16             | 12.18              | 14.91        | 13.18| 2.56  | −0.04                | 42           | 0.36 27.78** |
| Cyprus           | 79 | 21             | 4.88               | 12.57        | 7.70 | 4.38  | 0.91                 | 56           | 0.00** 87.44** |
| Czech Republic   | 79 | 0              | 7.28               | 5.74         | 6.72 | 1.57  | −0.63                | 68           | 0.00** 56.43** |
| Denmark          | 79 | 3              | 4.89               | 6.78         | 5.58 | 1.29  | 0.08                 | 65           | 0.17 18.59** |
| Estonia          | 79 | 6              | 10.30              | 8.68         | 9.71 | 3.51  | 0.43                 | 66           | 0.06 41.17** |
| France           | 79 | 15             | 8.94               | 9.70         | 9.22 | 1.27  | 0.53                 | 59           | 0.02* 27.15** |
| Germany          | 79 | 21             | 9.03               | 5.02         | 7.55 | 2.25  | −0.06                | 65           | 0.01** 109.25** |
| Greece           | 79 | 0              | 10.18              | 23.16        | 14.94| 6.85  | 0.77                 | 68           | 0.00** 291.21** |
| Hungary          | 79 | 3              | 7.30               | 8.34         | 7.68 | 2.05  | 0.49                 | 64           | 0.00** 63.28** |
| Iceland          | 79 | 15             | 3.23               | 4.94         | 3.86 | 1.94  | 1.04                 | 66           | 0.00** 26.23** |
| Ireland          | 79 | 4              | 5.90               | 11.49        | 7.95 | 3.91  | 0.63                 | 67           | 0.00** 241.97** |
| Italy            | 79 | 0              | 8.69               | 11.05        | 9.55 | 2.11  | 0.00                 | 56           | 0.01** 72.71** |
| Latvia           | 79 | 8              | 12.09              | 12.31        | 12.17| 3.57  | 0.27                 | 58           | 0.29 74.04** |
| Lithuania        | 79 | 8              | 11.66              | 11.38        | 11.55| 4.03  | −0.19                | 66           | 0.02* 68.13** |
| Luxembourg       | 79 | 15             | 3.73               | 5.66         | 4.44 | 1.44  | −0.17                | 61           | 0.02* 12.47** |
| Malta            | 79 | 14             | 6.50               | 5.73         | 6.22 | 1.00  | −0.73                | 42           | 0.00** 14.47** |
| Netherlands      | 79 | 6              | 3.55               | 6.11         | 4.49 | 1.54  | 0.55                 | 56           | 0.00** 70.23** |
| Norway           | 79 | 6              | 3.48               | 3.73         | 3.57 | 0.62  | 0.04                 | 68           | 0.26 8.97* |
| Poland           | 79 | 6              | 14.35              | 8.44         | 12.18| 4.78  | 0.40                 | 68           | 0.00** 117.50** |
| Portugal         | 79 | 0              | 6.58               | 13.19        | 9.00 | 3.84  | 0.50                 | 68           | 0.00** 200.11** |
| Romania          | 79 | 3              | 7.03               | 6.58         | 6.86 | 1.09  | −0.73                | 42           | 0.00** 4.58* |
| Slovakia         | 79 | 0              | 15.30              | 12.32        | 14.20| 3.26  | 0.02                 | 68           | 0.02* 85.88** |
| Slovenia         | 79 | 3              | 6.27               | 8.68         | 7.15 | 1.56  | 0.33                 | 68           | 0.26 42.55** |
| Spain            | 79 | 0              | 12.61              | 22.28        | 16.16| 5.79  | 0.20                 | 68           | 0.00** 192.12** |
| Sweden           | 79 | 9              | 6.70               | 7.63         | 7.04 | 1.27  | −0.05                | 66           | 0.07 18.97** |
| Switzerland      | 79 | 36             | 3.66               | 4.74         | 4.06 | 0.81  | −0.35                | 46           | 0.01** 22.91** |
| UK               | 79 | 4              | 5.62               | 6.49         | 5.93 | 1.19  | 0.96                 | 65           | 0.00** 69.59** |

*Note: SD (standard deviation); Skew. (skewness); BP (Bai–Perron) test for structural break; SW (Shapiro–Wilk) test for normality; VIF (variance inflation rate).

**Results are statistically significant at $\alpha = 0.01$. Moderate level of multi-collinearity.

*Results are statistically significant at $\alpha = 0.05$. High level of multi-collinearity.
### TABLE 3 Descriptive analysis of unemployment rate for USA

| State         | N  | Mean in-sample | Mean out-of-sample | Mean overall | SD   | Skew. | Last break point (BP) | SW (p value) | VIF |
|---------------|----|----------------|--------------------|--------------|------|-------|-----------------------|--------------|-----|
| Alabama       | 456 | 7.08           | 5.97               | 6.70         | 2.42 | 0.89  | 309                   | 0.00**       | 169.47** |
| Arkansas      | 456 | 6.79           | 6.19               | 6.59         | 1.46 | 0.27  | 309                   | 0.00**       | 115.36** |
| Arizona       | 456 | 6.27           | 6.60               | 6.38         | 1.92 | 0.86  | 254                   | 0.00**       | 37.66**  |
| California    | 456 | 7.28           | 7.92               | 7.50         | 2.02 | 0.73  | 229                   | 0.00**       | 216.75** |
| Colorado      | 456 | 5.37           | 6.13               | 5.63         | 1.59 | 0.27  | 254                   | 0.00**       | 52.44**  |
| Connecticut   | 456 | 5.12           | 6.16               | 5.48         | 1.83 | 0.39  | 310                   | 0.00**       | 87.74**  |
| Delaware      | 456 | 5.39           | 5.28               | 5.35         | 1.85 | 0.36  | 245                   | 0.00**       | 63.33**  |
| Florida       | 456 | 6.39           | 6.57               | 6.45         | 1.96 | 0.64  | 221                   | 0.00**       | 164.78** |
| Georgia       | 456 | 5.67           | 6.64               | 6.00         | 1.74 | 0.95  | 306                   | 0.00**       | 73.62**  |
| Iowa          | 456 | 4.89           | 4.65               | 4.80         | 1.49 | 0.86  | 307                   | 0.00**       | 87.21**  |
| Idaho         | 456 | 6.26           | 5.57               | 6.02         | 1.54 | 0.24  | 320                   | 0.00**       | 170.02** |
| Illinois      | 456 | 6.97           | 7.31               | 7.08         | 1.99 | 0.74  | 308                   | 0.00**       | 78.74**  |
| Indiana       | 456 | 6.05           | 6.66               | 6.26         | 2.38 | 0.75  | 308                   | 0.00**       | 150.20** |
| Kansas        | 456 | 4.48           | 5.43               | 4.80         | 0.99 | 0.70  | 308                   | 0.00**       | 49.78**  |
| Kentucky      | 456 | 6.89           | 7.18               | 6.99         | 1.93 | 0.67  | 308                   | 0.00**       | 143.44** |
| Louisiana     | 456 | 7.87           | 5.84               | 7.17         | 2.25 | 0.88  | 309                   | 0.00**       | 21.38**  |
| Massachusetts | 456 | 5.65           | 5.91               | 5.74         | 1.83 | 0.39  | 311                   | 0.00**       | 175.80** |
| Maryland      | 456 | 5.38           | 5.30               | 5.36         | 1.37 | 0.42  | 254                   | 0.00**       | 102.84** |
| Maine         | 456 | 6.00           | 5.82               | 5.94         | 1.57 | 0.14  | 307                   | 0.00**       | 97.79**  |
| Michigan      | 456 | 8.27           | 8.39               | 8.31         | 2.95 | 0.67  | 307                   | 0.00**       | 164.82** |
| Minnesota     | 456 | 4.89           | 5.30               | 5.03         | 1.35 | 0.67  | 309                   | 0.00**       | 105.79** |
| Missouri      | 456 | 5.82           | 6.38               | 6.01         | 1.64 | 0.69  | 309                   | 0.00**       | 72.89**  |
| Mississippi   | 456 | 8.06           | 7.79               | 7.97         | 2.02 | 0.66  | 225                   | 0.00**       | 49.23**  |
| Montana       | 456 | 6.27           | 4.83               | 5.78         | 1.26 | 0.07  | 320                   | 0.00**       | 313.48** |
| North Carolina| 456 | 5.28           | 7.29               | 5.97         | 2.07 | 0.96  | 310                   | 0.00**       | 138.96** |
| North Dakota  | 456 | 4.36           | 3.37               | 4.02         | 0.94 | 0.66  | 218                   | 0.00**       | 44.92**  |
| Nebraska      | 456 | 3.44           | 3.81               | 3.57         | 0.93 | 0.78  | 309                   | 0.00**       | 59.24**  |
| New Hampshire | 456 | 4.46           | 4.59               | 4.50         | 1.46 | 0.43  | 314                   | 0.00**       | 92.66**  |
| New Jersey    | 456 | 6.39           | 6.59               | 6.46         | 1.94 | 0.37  | 249                   | 0.00**       | 197.35** |
| New Mexico    | 456 | 7.38           | 5.79               | 6.84         | 1.42 | −0.23 | 308                   | 0.00**       | 138.83** |
| Nevada        | 456 | 6.38           | 7.75               | 6.85         | 2.61 | 1.19  | 251                   | 0.00**       | 102.02** |
| New York      | 456 | 6.86           | 6.49               | 6.73         | 1.55 | 0.17  | 314                   | 0.00**       | 109.55** |
| Ohio          | 456 | 6.91           | 6.86               | 6.89         | 2.14 | 1.07  | 224                   | 0.00**       | 150.15** |
| Oklahoma      | 456 | 5.48           | 5.04               | 5.33         | 1.45 | 0.47  | 309                   | 0.00**       | 42.91**  |
| Oregon        | 456 | 7.12           | 7.75               | 7.33         | 1.93 | 0.67  | 249                   | 0.00**       | 92.42**  |
| Pennsylvania  | 456 | 6.84           | 6.20               | 6.62         | 1.83 | 0.92  | 248                   | 0.00**       | 155.22** |
| Rhode Island  | 456 | 6.14           | 7.44               | 6.58         | 2.27 | 0.61  | 312                   | 0.00**       | 228.58** |
| South Carolina| 456 | 6.01           | 7.72               | 6.60         | 2.05 | 0.91  | 305                   | 0.00**       | 77.56**  |
| South Dakota  | 456 | 3.86           | 3.83               | 3.85         | 0.79 | 0.61  | 309                   | 0.00**       | 61.53**  |
| Tennessee     | 456 | 6.64           | 6.85               | 6.71         | 2.01 | 1.00  | 307                   | 0.00**       | 157.62** |
| Texas         | 456 | 6.26           | 6.19               | 6.24         | 1.25 | 0.35  | 288                   | 0.00**       | 85.34**  |
| Utah          | 456 | 5.09           | 5.16               | 5.11         | 1.60 | 0.69  | 226                   | 0.00**       | 68.09**  |
| TABLE 3 (Continued) |
|----------------------|
| **N** | **Mean in-sample** | **Mean out-of-sample** | **Mean overall** | **SD** | **Skew.** | **Last break point (BP)** | **SW (p value)** | **VIF** |
| Virginia | 456 | 4.80 | 4.67 | 4.76 | 1.29 | 0.13 | 224 | 0.00** | 127.45** |
| Vermont | 456 | 5.02 | 4.57 | 4.87 | 1.41 | 0.62 | 308 | 0.00** | 75.83** |
| Washington | 456 | 7.18 | 7.02 | 7.13 | 1.82 | 0.66 | 218 | 0.00** | 125.66** |
| Wisconsin | 456 | 5.41 | 5.99 | 5.61 | 1.85 | 0.92 | 308 | 0.00** | 88.94** |
| West Virginia | 456 | 9.52 | 5.98 | 8.31 | 3.11 | 1.04 | 311 | 0.00** | 124.84** |
| Wyoming | 456 | 5.34 | 4.51 | 5.06 | 1.54 | 0.62 | 214 | 0.00** | 74.44** |

Note: SD (standard deviation); Skew. (skewness); BP (Bai–Perron) test for structural break; SW (Shapiro–Wilk) test for normality; VIF (variance inflation rate).

**Results are statistically significant at \( \alpha = 0.01 \). High level of multi-collinearity.

*Results are statistically significant at \( \alpha = 0.05 \). Moderate level of multi-collinearity.

| TABLE 4 | RMSE for forecasting industrial production indicators for France, Germany and UK |
|----------|--------------------------------------------------------------------------------|
| **h** | **UK** | **France** | **Germany** |
| | 1 | 3 | 6 | 12 | 1 | 3 | 6 | 12 | 1 | 3 | 6 | 12 |
| E&G | | | | |
| SD-MSSA | 2.94* | 3.02* | 3.32 | 3.43 | 3.80* | 3.88* | 4.02 | 4.22* | 3.58 | 3.76* | 3.79* | 3.40* |
| MSSA | 3.14 | 3.48 | 3.32 | 3.53 | 3.99 | 4.05 | 4.15 | 4.44 | 3.73 | 4.69 | 5.63 | 4.82 |
| VAR | 3.43 | 4.69 | 5.17 | 5.37 | 4.10 | 4.57 | 4.98 | 7.61 | 3.88 | 4.76 | 5.59 | 5.74 |
| VECM | 3.31 | 4.42 | 4.72 | 5.95 | 4.41 | 5.21 | 5.71 | 6.70 | 3.51 | 5.05 | 5.67 | 5.13 |
| BVARSV | 3.38 | 6.13 | 7.05 | 8.65 | 4.94 | 5.90 | 7.03 | 8.16 | 3.72 | 4.64 | 5.05 | 5.88 |
| Chemical | | | | |
| SD-MSSA | 2.69 | 4.15* | 5.50* | 5.25* | 3.41 | 4.66 | 4.80* | 6.09* | 2.00* | 3.83* | 3.56* | 4.38* |
| MSSA | 2.71 | 4.78 | 7.17 | 8.33 | 3.24* | 5.88 | 7.88 | 7.48 | 2.46 | 4.84 | 6.69 | 5.57 |
| VAR | 2.65 | 4.41 | 6.69 | 8.34 | 3.06 | 5.68 | 8.26 | 8.88 | 2.16 | 4.01 | 5.91 | 6.12 |
| VECM | 2.61* | 4.45 | 6.75 | 8.35 | 2.90 | 5.22 | 7.68 | 7.94 | 2.12 | 4.12 | 6.31 | 6.84 |
| BVARSV | 2.91 | 6.65 | 7.01 | 7.68 | 3.27 | 4.35* | 5.61 | 6.55 | 2.36 | 2.88 | 4.14 | 5.43 |
| F.M. | | | | |
| SD-MSSA | 3.07 | 3.73* | 5.47* | 6.68* | 4.04 | 4.57 | 4.44* | 8.39* | 2.67 | 4.38* | 3.85* | 9.14* |
| MSSA | 2.38* | 4.55 | 7.57 | 10.03 | 3.19 | 6.90 | 11.38 | 13.89 | 2.08* | 5.26 | 10.12 | 9.94 |
| VAR | 2.87 | 5.52 | 8.61 | 10.56 | 3.20 | 6.68 | 11.55 | 15.00 | 2.24 | 5.14 | 8.56 | 9.51 |
| VECM | 2.39 | 4.39 | 7.40 | 9.83 | 3.04 | 6.09 | 10.32 | 13.02 | 2.29 | 5.30 | 8.99 | 11.31 |
| BVARSV | 2.77 | 5.54 | 7.33 | 9.75 | 3.34 | 4.55* | 6.35 | 9.22 | 2.43 | 4.76 | 4.65 | 9.46 |
| Vehicle | | | | |
| SD-MSSA | 7.10 | 8.19* | 10.58* | 16.51* | 7.27 | 7.78* | 9.73* | 15.35* | 6.03 | 7.34* | 9.69* | 14.70* |
| MSSA | 5.83* | 11.46 | 17.61 | 17.82 | 6.76 | 12.26 | 18.25 | 17.81 | 6.59 | 9.96 | 13.33 | 13.32 |
| VAR | 7.58 | 14.20 | 18.98 | 18.47 | 6.59 | 12.71 | 21.49 | 25.46 | 5.97 | 9.82 | 14.21 | 12.64 |
| VECM | 5.92 | 10.61 | 16.62 | 17.41 | 6.07* | 10.63 | 17.00 | 17.48 | 6.09 | 9.18 | 13.18 | 13.93 |
| BVARSV | 7.24 | 10.51 | 13.77 | 18.05 | 7.31 | 10.18 | 14.19 | 16.18 | 6.12 | 7.40 | 10.31 | 15.32 |
| F.P. | | | | |
| SD-MSSA | 1.81 | 2.20 | 2.43* | 2.65* | 1.28 | 1.41 | 1.36* | 1.21* | 1.51 | 1.53* | 1.56* | 1.65* |
| MSSA | 1.83 | 2.20 | 2.96 | 4.57 | 1.29 | 1.40 | 1.80 | 2.06 | 1.49* | 1.55 | 1.79 | 2.03 |
| VAR | 1.82 | 2.76 | 3.81 | 5.10 | 1.64 | 2.15 | 2.64 | 3.20 | 1.99 | 2.28 | 2.89 | 4.33 |
| VECM | 1.74 | 2.57 | 3.30 | 5.19 | 1.59 | 1.74 | 1.97 | 2.58 | 1.94 | 1.91 | 2.02 | 2.32 |
| BVARSV | 1.91 | 3.15 | 3.30 | 3.63 | 1.77 | 1.90 | 1.91 | 2.37 | 1.84 | 3.02 | 4.12 | 5.04 |

(Continues)
Table 3 shows that the null hypothesis of a unit root at the 1% level cannot be rejected for any of the states. Also, the normality test (SW) confirms that these time series are not normally distributed. Last but not least, the collinearity test confirms that the USA unemployment series are strongly correlated. This can be mainly related to the spatial dependency and geographical homogeneity which determines an association between the commuting flow variables and spatial interdependencies between states.

4.2.2 | Forecasting results

In this section, we evaluate the forecasting performance of SD-MSSA relative to standard MSSA and different time series models, such as Vector AutoRegressive (VAR<sup>9</sup>), Bayesian VAR (BVAR<sup>10</sup>) and Vector Error Correction Model (VECM<sup>11</sup>) for IPI and UR time series in Tables 4–6.

Our algorithm settings are as follows for all experiments. The window length is set to values between 2 and 50 and in the process, we evaluate all possible values of r for r < L. For the monthly series in almost all case we found L's between 12 to 36 result in better forecasting accuracy. As mentioned before, we used multivariate bootstrap SSA (Golyandina et al., 2001) to generate initial values for SD-LRF. The maximum number of epochs is set to 1000. The selected range for the smoothing factor in this study is between 10<sup>-3</sup> to 10<sup>-5</sup>.

Forecast accuracy is measured using the prediction root-mean-square error (RMSE) and mean absolute error (MAE). However, since these measures give quantitatively similar results and to conserve space, we only report the RMSE, as this is the most frequently quoted measure in forecasting. For assessing the statistical significance of the forecasting methods we use the Diebold-Mariano test (DM test) (Diebold & Mariano, 2002).

Table 4 (Continued)

| h | UK | | France | | | | Germany |
|---|---|---|---|---|---|---|---|
| B.M. | | | | | | | |
| SD-MSSA | 6.51 | 8.19 | 12.25 | 12.52 | 4.22 | 4.86<sup>*</sup> | 4.34 | 10.36 | 5.02 | 6.06<sup>*</sup> | 6.64<sup>*</sup> | 11.93<sup>*</sup> |
| MSSA | 5.48 | 9.27 | 13.46 | 13.84 | 3.05 | 7.19 | 12.15 | 14.65 | 3.53<sup>*</sup> | 8.39 | 14.07 | 14.17 |
| VAR | 5.23 | 8.56 | 13.82 | 15.81 | 2.99 | 6.85 | 12.06 | 15.52 | 3.88 | 8.38 | 13.35 | 12.43 |
| VECM | 5.28 | 8.13 | 12.86 | 14.88 | 2.92 | 6.43 | 10.94 | 13.62 | 3.63 | 7.67 | 12.89 | 13.67 |
| BVAR<sub>SV</sub> | 5.10<sup>*</sup> | 8.96 | 9.87<sup>*</sup> | 10.27<sup>*</sup> | 3.47 | 4.98 | 6.36* | 9.62* | 4.10 | 6.08 | 7.96 | 12.38 |
| E.M. | | | | | | | |
| SD-MSSA | 4.71 | 6.25* | 7.88* | 9.50* | 3.09 | 4.10* | 3.86* | 9.79* | 2.63 | 4.35* | 3.97* | 5.03* |
| MSSA | 4.41<sup>*</sup> | 6.71 | 10.55 | 12.92 | 2.97 | 5.01 | 8.82 | 14.56 | 2.15<sup>*</sup> | 5.03 | 7.55 | 6.45 |
| VAR | 4.68 | 7.87 | 11.65 | 14.13 | 2.83<sup>*</sup> | 4.98 | 9.27 | 15.01 | 2.31 | 4.98 | 9.07 | 12.35 |
| VECM | 4.42 | 7.18 | 10.95 | 14.77 | 3.21 | 5.47 | 9.32 | 13.81 | 2.39 | 5.10 | 9.00 | 12.09 |
| BVAR<sub>SV</sub> | 5.18 | 6.85 | 7.93 | 10.54 | 3.07 | 4.38 | 4.69 | 10.67 | 2.48 | 4.64 | 5.08 | 7.23 |
| Machinery | | | | | | | |
| SD-MSSA | 4.45 | 5.29 | 7.64 | 9.97 | 4.46 | 5.08* | 5.93* | 13.41* | 3.95 | 5.45* | 5.67* | 7.77* |
| MSSA | 3.64<sup>*</sup> | 5.33 | 8.14 | 12.01 | 4.13 | 6.86 | 11.95 | 20.71 | 3.73 | 6.23 | 11.19 | 12.64 |
| VAR | 3.98 | 6.43 | 10.10 | 15.42 | 4.04 | 7.63 | 13.46 | 21.67 | 3.67 | 5.29 | 9.01 | 12.95 |
| VECM | 3.97 | 6.08 | 9.87 | 14.81 | 4.15 | 7.64 | 12.97 | 19.80 | 4.26 | 6.14 | 10.40 | 15.76 |
| BVAR<sub>SV</sub> | 4.22 | 5.96 | 11.63* | 8.99<sup>*</sup> | 3.94 | 5.58 | 6.65 | 14.47 | 3.93 | 5.11 | 6.53 | 8.84 |

Note: The table reports RMSE for five forecasting methods, SD-MSSA, MSSA, VAR, VECM and BVAR<sub>SV</sub>.

*RMSE results are statistically significant at α = 0.05 based on the DM test.

and New Mexico. However, states like North and South Carolina and Nevada display 10% reduction in their unemployment rate.

Descriptive analysis in Table 3 shows that the null hypothesis of a unit root at the 1% level cannot be rejected for any of the states. Also, the normality test (SW) confirms that these time series are not normally distributed. Last but not least, the collinearity test confirms that the USA unemployment series are strongly correlated. This can be mainly related to the spatial dependency and geographical homogeneity which determines an association between the commuting flow variables and spatial interdependencies between states.

4.2.2 | Forecasting results

In this section, we evaluate the forecasting performance of SD-MSSA relative to standard MSSA and different time series models, such as Vector AutoRegressive (VAR<sup>9</sup>), Bayesian VAR (BVAR<sup>10</sup>) and Vector Error Correction Model (VECM<sup>11</sup>) for IPI and UR time series in Tables 4–6.

Our algorithm settings are as follows for all experiments. The window length is set to values between 2 and 50 and in the process, we evaluate all possible values of r for r < L. For the monthly series in almost all case we found L's between 12 to 36 result in better forecasting accuracy. As mentioned before, we used multivariate bootstrap SSA (Golyandina et al., 2001) to generate initial values for SD-LRF. The maximum number of epochs is set to 1000. The selected range for the smoothing factor in this study is between 10<sup>-3</sup> to 10<sup>-5</sup>.

Forecast accuracy is measured using the prediction root-mean-square error (RMSE) and mean absolute error (MAE). However, since these measures give quantitatively similar results and to conserve space, we only report the RMSE, as this is the most frequently quoted measure in forecasting. For assessing the statistical significance of the forecasting methods we use the Diebold-Mariano test (DM test) (Diebold & Mariano, 2002). The
| h     | 1      | 3      | 6      | 12     | 1      | 3      | 6      | 12     | 1      | 3      | 6      | 12     |
|-------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| **Austria** | **0.35** | 0.45  | 0.45  | 0.39* | **SD-MSSA 0.37** | 1.38* | 2.56* | 5.09* | **SD-MSSA 0.25** | 0.40  | 0.50  | 0.44* |
| **MSSA** | 0.36  | 0.41  | 0.47  | 0.51  | **MSSA 0.62** | 1.58  | 3.97  | 7.12  | **MSSA 0.25** | 0.42  | 0.58  | 0.72  |
| **Belgium** | **0.46** | 0.58  | 0.63  | 0.53  | **SD-MSSA 0.56** | 0.82  | 1.09* | 0.45* | **SD-MSSA 0.47** | 0.82  | 1.21  | 1.00* |
| **MSSA** | 0.46  | 0.59  | 0.67  | 0.56  | **MSSA 0.52** | 0.84  | 1.45  | 2.22  | **MSSA 0.41** | 0.65* | 1.22  | 2.00  |
| **Bulgaria** | **0.63** | 0.74  | 1.10  | 1.22* | **SD-MSSA 0.55** | 0.61  | 0.58* | 0.88* | **SD-MSSA 0.63** | 1.51  | 2.35* | 2.46* |
| **MSSA** | 0.60  | 0.69  | 1.34  | 2.14  | **MSSA 0.67** | 0.63  | 1.41  | 1.30  | **MSSA 0.50** | 0.82  | 0.65  | 0.94  |
| **Croatia** | **1.21** | 1.84  | 1.50* | 1.94* | **SD-MSSA 0.54** | 0.74  | 1.52  | **0.84** | **SD-MSSA 0.36** | 0.46  | 0.54* | 0.33* |
| **MSSA** | 1.07* | 1.14* | 1.99  | 3.13  | **MSSA 0.44** | 0.63* | 1.30  | 2.52  | **MSSA 0.40** | 0.46  | 0.65  | 0.94  |
| **Czech** | **0.80** | 2.37  | 2.79* | 4.02* | **SD-MSSA 0.50** | 0.97  | 1.10* | 1.45* | **SD-MSSA 0.47** | 0.87  | 1.09* | 0.97* |
| **MSSA** | 0.30  | 0.52  | 1.04  | 1.59  | **MSSA 0.90** | 1.62  | 2.45  | 3.41  | **MSSA 0.54** | 0.82  | 1.22  | 1.86  |
| **Denmark** | **0.24** | 0.50  | 0.82* | 0.43* | **SD-MSSA 0.80** | 0.85* | 0.71* | 1.02* | **SD-MSSA 0.67** | 0.50* | 1.12* | 1.05* |
| **MSSA** | 0.30  | 0.52  | 1.04  | 1.59  | **MSSA 0.90** | 1.62  | 2.45  | 3.41  | **MSSA 0.54** | 0.82  | 1.22  | 1.86  |
| **Estonia** | **0.84** | 0.97* | 0.81* | 1.40* | **SD-MSSA 0.73** | 0.76  | 0.76* | 0.82* | **SD-MSSA 0.25** | 0.46  | 0.66* | 0.63* |
| **MSSA** | 1.38  | 1.95  | 2.82  | 3.35  | **MSSA 0.75** | 0.77  | 0.87  | 0.91  | **MSSA 0.23** | 0.44  | 0.82  | 0.84  |
| **France** | **0.26** | 0.55  | 0.69  | 0.49* | **SD-MSSA 0.26** | 0.37  | 0.39* | 0.20* | **SD-MSSA 0.32** | 0.37  | 0.36  | 0.38  |
| **MSSA** | 0.28  | 0.39* | 0.56  | 0.57  | **MSSA 0.26** | 0.34  | 0.51  | 0.93  | **MSSA 0.35** | 0.35  | 0.32  | 0.34* |
| **Germany** | **0.22** | 0.30  | 0.30  | 0.20  | **SD-MSSA 0.30** | 0.74  | 0.77* | 0.84* | **SD-MSSA 0.24** | 0.56  | 0.75  | 0.31* |
| **MSSA** | 0.23  | 0.25  | 0.30  | 0.31  | **MSSA 0.22** | 0.49* | 1.09  | 1.95  | **MSSA 0.24** | 0.56  | 0.84  | 1.21  |

**Note:** The table reports RMSE for two forecasting methods, SD-MSSA and MSSA. *RMSE results are statistically significant at $\alpha=0.05$ based on DM test.
| TABLE 6  | RMSE for forecasting unemployment rate for the USA |
|----------|--------------------------------------------------|
|          | h | RMSE for SD-MSSA | RMSE for MSSA | RMSE for SD-MSSA | RMSE for MSSA | RMSE for SD-MSSA | RMSE for MSSA | RMSE for SD-MSSA | RMSE for MSSA | RMSE for SD-MSSA | RMSE for MSSA |
|          | 1  | 3  | 6  | 12 | 1  | 3  | 6  | 12 | 1  | 3  | 6  | 12 |
| Alabama  | SD-MSSA | 0.11* | 0.34 | 0.71 | 0.99* | 0.08 | 0.23* | 0.36* | 0.72* | 0.09 | 0.28* | 0.58 | 0.66* |
|          | MSSA   | 0.23 | 0.40 | 0.74 | 1.41 | 0.10 | 0.43 | 0.49 | 0.91 | 0.15 | 0.50 | 0.58 | 1.12 |
| Arkansas | SD-MSSA | 0.06 | 0.16 | 0.22* | 0.63* | 0.09 | 0.26* | 0.42* | 0.70* | 0.11 | 0.36 | 0.69 | 0.91 |
|          | MSSA   | 0.08 | 0.23 | 0.38 | 0.70 | 0.11 | 0.41 | 0.79 | 0.89 | 0.10 | 0.45 | 0.56 | 0.98 |
| Arizona  | SD-MSSA | 0.09 | 0.26* | 0.56* | 0.94* | 0.22 | 0.35 | 0.40* | 0.52* | 0.15 | 0.34 | 0.50* | 1.04* |
|          | MSSA   | 0.19 | 0.55 | 0.77 | 1.24 | 0.09 | 0.39 | 0.68 | 0.86 | 0.16 | 0.38 | 0.74 | 1.41 |
| California | SD-MSSA | 0.19 | 0.29* | 0.53 | 0.68* | 0.11 | 0.34* | 0.65* | 0.88* | 0.18 | 0.25 | 0.26* | 0.53* |
|          | MSSA   | 0.08 | 0.54 | 0.56 | 1.18 | 0.28 | 0.68 | 0.81 | 1.63 | 0.11 | 0.41 | 0.79 | 0.89 |
| Colorado | SD-MSSA | 0.18 | 0.26 | 0.42* | 0.78* | 0.10 | 0.27* | 0.35* | 0.60* | 0.21 | 0.26* | 0.38* | 0.55* |
|          | MSSA   | 0.13 | 0.29 | 0.67 | 1.08 | 0.12 | 0.42 | 0.55 | 0.90 | 0.06* | 0.36 | 0.42 | 0.82 |
| Connecticut | SD-MSSA | 0.17 | 0.19* | 0.42 | 0.52* | 0.15 | 0.42 | 0.60 | 0.71* | 0.10 | 0.34* | 0.48* | 1.12* |
|          | MSSA   | 0.13 | 0.40 | 0.52 | 0.84 | 0.18 | 0.52 | 0.70 | 1.13 | 0.19 | 0.64 | 0.71 | 1.42 |
| Delaware | SD-MSSA | 0.10 | 0.28* | 0.27* | 0.60* | 0.39 | 0.73 | 0.81* | 0.69* | 0.12 | 0.28 | 0.43 | 0.55* |
|          | MSSA   | 0.14 | 0.41 | 0.51 | 0.95 | 0.38 | 0.70 | 0.99 | 1.38 | 0.12 | 0.33 | 0.48 | 0.72 |
| Florida  | SD-MSSA | 0.15 | 0.37* | 0.65 | 0.72* | 0.06 | 0.13* | 0.36 | 0.37 | 0.10 | 0.30* | 0.46* | 0.87* |
|          | MSSA   | 0.28 | 0.58 | 0.67 | 1.31 | 0.05 | 0.24 | 0.24* | 0.46 | 0.19 | 0.57 | 0.69 | 1.32 |
| Georgia  | SD-MSSA | 0.10 | 0.31* | 0.56 | 0.67* | 0.10 | 0.33* | 0.69* | 0.80* | 0.07 | 0.21* | 0.22* | 0.65* |
|          | MSSA   | 0.19 | 0.47 | 0.56 | 1.10 | 0.18 | 0.59 | 0.86 | 1.25 | 0.05 | 0.41 | 0.42 | 0.81 |
| Iowa    | SD-MSSA | 0.23 | 0.35 | 0.32* | 0.56* | 0.07 | 0.17 | 0.32 | 0.30* | 0.21 | 0.59 | 0.89 | 1.05* |
|          | MSSA   | 0.23 | 0.36 | 0.47 | 0.74 | 0.07 | 0.10 | 0.32 | 0.47 | 0.20 | 0.66 | 0.90 | 1.45 |

* indicates statistical significance.
|                |  1    |  3    |  6    | 12   |  1    |  3    |  6    | 12   |  1    |  3    |  6    | 12   |
|----------------|-------|-------|-------|------|-------|-------|-------|------|-------|-------|-------|------|
| **SD-MSSA**    | 0.08  | 0.32  | 0.38  | 0.69 | 0.07  | 0.21  | 0.31  | 0.41 | 0.19  | 0.40  | 0.26  | 0.66 |
| **MSSA**       | 0.13  | 0.42  | 0.40  | 0.76 | 0.18  | 0.42  | 0.40  | 0.76 | 0.20  | 0.47  | 0.27  | 0.46 |
| **Illinois**   |       |       |       |      |       |       |       |      |       |       |       |      |
| **SD-MSSA**    | 0.11  | 0.47  | 0.41  | 0.70 | 0.10  | 0.41  | 0.40  | 0.75 | 0.07  | 0.29  | 0.30  | 0.58 |
| **MSSA**       | 0.18  | 0.54  | 0.66  | 1.20 | 0.09  | 0.35  | 0.54  | 1.12 | 0.12  | 0.56  | 0.58  | 1.12 |
| **Indiana**    |       |       |       |      |       |       |       |      |       |       |       |      |
| **SD-MSSA**    | 0.12  | 0.35  | 0.52  | 1.07 | 0.12  | 0.41  | 0.54  | 1.12 | 0.08  | 0.27  | 0.37  | 1.07 |
| **MSSA**       | 0.21  | 0.62  | 0.75  | 1.39 | 0.09  | 0.35  | 0.65  | 0.70 | 0.12  | 0.53  | 0.61  | 1.19 |
| **Kansas**     |       |       |       |      |       |       |       |      |       |       |       |      |
| **SD-MSSA**    | 0.12  | 0.32  | 0.53  | 0.71 | 0.12  | 0.41  | 0.54  | 0.88 | 0.10  | 0.45  | 0.46  | 0.83 |
| **MSSA**       | 0.15  | 0.46  | 0.58  | 0.81 | 0.09  | 0.35  | 0.65  | 0.70 | 0.11  | 0.49  | 0.47  | 0.83 |
| **Louisiana**  |       |       |       |      |       |       |       |      |       |       |       |      |
| **SD-MSSA**    | 0.81  | 1.24  | 1.00  | 1.02 | 0.09  | 0.44  | 0.31  | 0.67 | 0.07  | 0.17  | 0.30  | 0.75 |
| **MSSA**       | 0.81  | 1.00  | 1.34  | 1.53 | 0.11  | 0.45  | 0.53  | 0.96 | 0.10  | 0.44  | 0.46  | 0.96 |

**Note:** The table reports RMSE for two forecasting methods, SD-MSSA and MSSA.  
*RMSE results are statistically significant at $\alpha=0.05$ based on DM test.
forecasting accuracy is also assessed for four different horizons, one step ahead, three and six steps ahead and one year ahead (12-step).

The results of the forecasting competition are shown in Tables 4–6. These tables present the values of the RMSE obtained from recursive forecasts for 1, 3, 6 and 12 months. Table 4 shows RMSE values for IPI while Tables 5 and 6 summaries the RMSE values for UR series.

When analysing the forecast accuracy for IPI time series, MSSA models show lower RMSE values than autoregressive models, especially for longer horizons. In spite of showing the lowest RMSE values for Machinery and E.M. for Germany and France for $h = 1$, VAR models display the highest RMSE values for the UK. For a three month ahead horizon, the VECM returns the lowest RMSE only for B. M for the UK. Overall, in three countries, the lowest RMSE values between non-SSA models are returned by the BVARSV for the longer horizons $h = 3$ to 12.

While the out-of-sample forecast performance of MSSA models differs between horizons, the SD-MSSA model shows significantly lower forecasting errors compared to its counterpart, MSSA. In Table 7 we present the summary of the number of significant predictive accuracy based on the results of the DM test between each two models for one month to twelve month ahead forecasts. Among autoregressive models, BVARSV model shows a higher percentage of significant forecast accuracy for horizons greater than one month. MSSA performs better for one step ahead forecasts while its counterpart, SD-MSSA outperforms MSSA, BVARSV, VAR and VECM for horizons of 3,6 and 12 months. These differences are statistically significant in 75%, 79% and 83% of the cases respectively.

In order to better understand the forecasting gain the cumulative squared forecast of the out-of-sample errors (for all French Vehicle series) are shown in Figure 4. Most noticeably of all the SD-MSSA forecasting errors are significantly smaller than those of the MSSA, VAR and VECM models for both horizons, $h = 3$ and $h = 12$. In addition, the forecast errors estimated by SD-MSSA models are shown to be much smaller than their Bayesian counterparts.

Table 5 shows values of RMSE for MSSA and SD-MSSA for UR in Europe states. Not surprisingly, the values of RMSE of the SD-MSSA model are much smaller for horizons up to three month in comparison to the RMSE’s for MSSA. Nevertheless, MSSA performs better for forecasting Ireland and Spain for horizons up to 6. These are two out of five countries with significant growth in their unemployment rate. Comparisons between two forecasting models tested by DM test, Table 7, shows MSSA’s forecasts are statistically significant in 10% and 23% of cases for $h = 1$ and 3, respectively. On the other side, SD-MSSA is yet again a more accurate model for horizons 6 and 12 and the differences between two models are statistically significant in 63% and 93% of the cases.

Next, Table 6 tabulates the forecasting performance of MSSA and SD-MSSA for the USA unemployment time series. Once again, MSSA results in a lower RMSE for four series, Kentucky, Pennsylvania and Vermont for one month ahead forecasts. However, for longer term
horizons (3–12 months) SD-MSSA reports the lowest RMSE in nearly all cases. Table 7 also confirms that SD-MSSA returns higher percentage of forecast accuracy than MSSA.

The fact that SD-MSSA outperforms MSSA for forecasting IPI and UR time series is related to the fact that all of these time series are affected by the global economic and financial crisis in 2008. However, the impact of the shock varies between each country, and the strong correlation between these inter-related economic variables suggests multivariate time series analysis improves forecasting performance. Our results also suggest that there is a trade-off between computation cost (updating the LRF coefficients related to each series) and forecasting accuracy obtained with the SD-LRF.

5 | CONCLUSION AND SUMMARY

In this paper, we developed a state dependent model for the LRF in multivariate singular spectrum analysis. We proved theoretically that when a structural break happens to a time series the LRF, as an underlying assumption of MSSA (and SSA), does not coincide with a recurrent continuation of the series pre-break. Accordingly, in our model, called SD-MSSA, we propagate coefficients as a function of the state vector to involve their dependent movement in the LRF. The propagation coefficients can then mathematically explain their time evolution as well as their state position. Empirical evidence shows the proposed model to be superior to MSSA in terms of the model parameters recovering from a shock.

The performance of the SD-MSSA model is evaluated using both synthetic and real time series including a structural break. Of the five methods compared (SD-MSSA, MSSA, VAR, Bayesian VAR and VECM) SD-MSSA was found to be the most accurate method for forecasting horizons of up to a year. Our results showed that MSSA performs better for one step ahead forecasts while its counterpart, SD-MSSA outperforms MSSA, BVAR, VAR and VECM significantly for horizons of 3, 6 and 12 months. It was also shown that SD-MSSA has a higher percentage of forecast accuracy than MSSA.

The SDM model examined assumes an approximate Gaussian posterior for the coefficients of the LRF model (Underlying assumption of EKF) but further work involving a non-Gaussian posterior may be interesting such as using the unscented Kalman filter or the particle filter to estimate parameters. Also this might provide us with a means to reject noise in the time series which can have an adverse effect on forecasts. It would be also of interest to look at different kinds of non-linearity in the SDM (i.e., change the random walk model in Equation (10)). This would be interesting and in fact might align with the idea of using the unscented Kalman filter or particle filter.

DATA AVAILABILITY STATEMENT

Sample data and codes for Bootstrap MSSA and SD-MSSA may be found online (https://github.com/drahmani/SD-MS-SA).

ORCID

Donya Rahmani https://orcid.org/0000-0003-1044-1005

REFERENCES

Allan, M. R., & Smith, L. A. (1996). Monte Carlo SSA: Detecting irregular oscillations in the presence of colored noise. Journal of Climate, 9(12), 3373–3404.

Awichi, R. O., & Müller, W. G. (2013). Improving SSA predictions by inverse distance weighting. REVSTAT–Statistical Journal, 11(1), 105–119.

Babaie, T., Chawla, S., & Ardon, S. (2014). Network traffic decomposition for anomaly detection. arXiv preprint:1403.0157.

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

Ball, L., Jalles, J. T., & Loungani, P. (2015). Do forecasters believe in Okun’s law? An assessment of unemployment and output forecasts. International Journal of Forecasting, 31(1), 176–184.

Beaudry, P., Galizia, D., & Portier, F. (2015). Reviving the limit cycle view of macroeconomic fluctuations: NBER.

Broomhead, D. S., & King, G. P. (1986). Extracting qualitative dynamics from experimental data. Physica D: Nonlinear Phenomena, 20(2-3), 217–236.

Cao, L., Mees, A., & Judd, K. (1998). Dynamics from multivariate time series. Physica D: Nonlinear Phenomena, 121, 75–88.

Danilov, D., & Zhigljavsky, A. (1997). Principal components of time series: The caterpillar method: St.Petersburg Press, St. Petersburg (in Russian).

Diebold, F. X., & Mariano, R. S. (2002). Comparing predictive accuracy. Journal of Business & Economic Statistics, 20(1), 134–144.

Fendel, R., Lis, E. M., & Rülke, J. (2011). Do professional forecasters believe in the Phillips curve? Evidence from the G7 countries. Journal of Forecasting, 30(2), 268–287.

Fraedrich, K. (1986). Estimating dimensions of weather and climate attractors. Journal of the Atmospheric Sciences, 43, 419–432.

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2013). Bayesian data analysis: CRC press.

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

Golyandina, N., & Zhigljavsky, A. (2013). Singular spectrum analysis for time series: Springer.

Groth, A., & Ghil, M. (2011). Multivariate singular spectrum analysis and the road to phase synchronization. Physical Review E, 84(3), 36206.

Groth, A., Ghil, M., Halle, S., & Dumas, P. (2015). The role of oscillatory modes in US business cycles. Journal of Business Cycle Measurement and Analysis, 2015(1), 63–81.
Hassani, H., Heravi, S., & Zhigljavsky, A. (2012). Forecasting UK industrial production with MSSA. *Journal of Forecasting*, 408 (May), 1–6.

Hassani, H., Soofi, A. S., & Zhigljavsky, A. (2013). Predicting inflation dynamics with singular spectrum analysis. *Journal of the Royal Statistical Society A*, 176, 743–760.

Hendry, D. F., & Mizon, G. E. (2005). Forecasting in the presence of structural breaks and policy regime shifts. In D. W. Andrews & J. H. Stock (Eds.), *Identification and inference for econometric models: essays in honor of Thomas Rothenberg* (pp. 480–502). Cambridge, UK: Cambridge University Press.

Heravi, S., Osborn, D. R., & Birchenhall, C. R. (2004). Linear versus neural network forecasts for European industrial production series. *International Journal of Forecasting*, 20, 435–446.

Kantz, H., & Schreiber, T. (2004). *Nonlinear time series analysis*, Vol. 7: Cambridge university press.

Kapl, M., & Mueller, W. (2010). Prediction of steel prices: A comparison between a conventional regression model and MSSA. *Stat Interface*, 3(3), 369–375.

Kim, C.-J., Nelson, C. R., & Piger, J. (2004). The less-volatile US economy: A Bayesian investigation of timing, breadth, and potential explanations. *Journal of Business & Economic Statistics*, 22(1), 80–93.

Krüger, F. (2015). bvarsr: An R implementation of the primiceri model wind speed time series. https://cran.r-project.org/web/packages/bvarsr

Lahiani, A., & Scaillet, O. (2009). Testing for threshold effect in ARFIMA models: Application to US unemployment rate data. *International Journal of Forecasting*, 25(2), 418–428.

Lütkepohl, H. (2005). *New introduction to multiple time series analysis*: Springer Science and Business Media.

Menezes, M. L. D., Souza, R. C., & Pessanha, J. F. M. (2014). Combining singular spectrum analysis and PAR(p) structures to model wind speed time series. *Journal of Systems Science and Complexity*, 27(1), 29–46.

Milas, C., & Rothman, P. (2008). Out-of-sample forecasting of unemployment rates with pooled STVECM forecasts. *International Journal of Forecasting*, 24(1), 101–121.

Oropeza, V., & Sachchi, M. (2011). Simultaneous seismic data denoising and reconstruction via multichannel singular spectrum analysis. *Geophysics*, 76, 25–32.

Osborn, D. R., Heravi, S., & Birchenhall, C. R. (1999). Seasonal unit roots and forecasts of two-digit European industrial production. *International Journal of Forecasting*, 15(1), 27–47.

Patterson, K., Hassani, H., Heravi, S., & Zhigljavsky, A. (2011). Multivariate singular spectrum analysis for forecasting revisions to real-time data. *Journal of Applied Statistics*, 38(10), 2183–2211.

Pesaran, M. H., Pettenuzzo, D., & Timmermann, A. (2006). Forecasting time series subject to multiple structural breaks. *The Review of Economic Studies*, 73(4), 1057–1084.

Priestley, M. B. (1980). State dependent models: A general approach to non-linear time series analysis. *Journal of Time Series Analysis*, 1(1), 47–71.

Priestley, M. B. (1982). On the fitting of general non-linear time series models. In *4th international time series meeting, valencia, amsterdam, north holland pub.*, 1982, pp. 717–731.

Rahmani, D. (2017). Bayesian singular spectrum analysis with state dependent models. (Ph.D. Thesis), Bournemouth University.

Skare, M., & Buterin, V. (2015). Modelling and forecasting unemployment non-linear dynamics using spectral analysis. *Engineering Economics*, 26(9481), 373–383.

Takens, F. (1981). Detecting strange attractors in turbulence. *Dynamical systems and turbulence, warwick 1980*: Springer, pp. 366–381.

Trendle, B. (2002). Regional variation in Queensland’s unemployment rate. Australian and New Zealand Regional Science Association.

Vautard, R., Yiou, P., & Ghil, M. (1992). Singular-spectrum analysis: A toolkit for short, noisy chaotic signals. *Physica D: Nonlinear Phenomena*, 58(1-4), 95–126.

Viljoen, H., & Nel, D. G. (2010). Common singular spectrum analysis of several time series. *Journal of Statistical Planning and Inference*, 140(1), 260–267.

Yiou, P., Sornette, D., & Ghil, M. (2000). Data-adaptive wavelets and multi-scale singular spectrum analysis. *Physica*, 142, 254–290.

Zhang, Y., & Hui, X. F. (2012). Research on daily exchange rate forecasting with multivariate singular spectrum analysis. In *International Conference on Management Science and Engineering - Annual Conference Proceedings*, pp. 1365–1370.

Zhao, X., Shang, P., Jin, Q., & Qiuyue, J. (2011). Multifractal detrended cross-correlation analysis of Chinese stock markets based on time delay. *Fractals*, 19(03), 329–338.

**AUTHOR BIOGRAPHIES**

**Donya Rahmani** obtained her PhD in time series analysis from Bournemouth University in 2017. From 2005–2011 she obtained a degree in mathematics and a masters in mathematical statistics. She was also a research associate fellow in the department of mathematical science at university of Southampton working on statistical modelling, machine learning and topological data analysis. Donya’s interdisciplinary collaboration has ranged across applications in computational topological data analytics, econometrics forecasting, bioinformatics, genomics, chemistry and sport science, among other areas.

**Damien Fay** obtained his PhD from DCU in 2003 in time series analysis. From 2003-2007 he was a statistics lecturer at NUI Galway. He was a research associate in Cambridge, McGill and UCC working on applied graph theory and graph evolution time series of the internet. In 2017 joined the analytics group at Infor and he is currently the retail science lead. His interests lie in time series analysis, machine
learning and applied graph theory. He has published over 40 peer reviewed papers and was awarded a commercialisation award at UCC for his research and patents into time of arrival prediction techniques.

How to cite this article: Rahmani, D., & Fay, D. (2021). A state-dependent linear recurrent formula with application to time series with structural breaks. Journal of Forecasting, 1–21. https://doi.org/10.1002/for.2778

APPENDIX A

In this section, we show how the Extended Kalman Filter is used to estimate the state vector representation of the LRF:

$$\theta_{N+1}^{(m)} = F_N^* \theta_{N}^{(m)} + K_{N+1}^* (y_{N+1} - H_{N+1}^* F_N^* \theta_{N}^{(m)}),$$ (A1)

where $F_N^* \in \mathbb{R}^{2(L-1) \times 2(L-1)}$ is given by:

$$F_N^* = \begin{pmatrix}
\Delta y_N & 0 \\
\Delta y_{N-1} & \ddots \\
0 & \ddots & \Delta y_{N-L+2} \\
\end{pmatrix},$$

and $K_{N+1}^*$ is the “Kalman gain” matrix where $K_{N+1}^* = \Phi_{N+1}(H_{N+1}^*)^T \sigma_e^2$, and $\Phi_{N+1}$ is the variance-covariance matrix of the one-step prediction error of $\theta_{N+1}$, that is, $\Phi_{N+1} = E[(\theta_{N+1} - F_N^* \theta_N)(\theta_{N+1} - F_N^* \theta_N)^T]$. Moreover, $\sigma_e^2$ is the variance of the one-step ahead prediction error of $y_{N+1}$, i.e., $\sigma_e^2$ is the variance of $e_{N+1} = y_{N+1} - H_{N+1}^* F_N^* \theta_N$.

Let $C_{N+1}$ be the variance-covariance matrix of $(\theta_{N+1} - \theta_{N-1})$, then successive values of $\theta_{N+1}$ can be estimated by using the standard Kalman filter:

$$\Phi_{N+1} = F_N^* C_N(F_N^*)^T + \Sigma_W, \quad \Sigma_W = \begin{pmatrix}
0 & 0 \\
0 & \Sigma_V
\end{pmatrix},$$

$$K_{N+1}^* = \Phi_{N+1} (H_{N+1}^*)^T [H_{N+1}^* \Phi_{N+1} (H_{N+1}^*)^T + \sigma_e^2]^{-1},$$

$$C_{N+1} = \Phi_{N+1} - K_{N+1}^* [H_{N+1}^* \Phi_{N+1} (H_{N+1}^*)^T + \sigma_e^2] (K_{N+1}^*)^T.$$

(A2)

To provide initial values for the SD-LRF model, $\theta_{N+1}$, $\sigma_e^2$ and $C_{N+1}$, (Priestley, 1980) recommends fitting a standard AR model to the data, and provides a good starting points for SDM. In this paper, we use bootstrap SSA (Golyandina et al., 2001; Rahmani, 2017) and set $\theta_{N+1} = (\phi_{N+1 L-1}, \ldots, \phi_{N+1 L-1, 0}, 0, \ldots, 0), \sigma_e^2$ and $C_{N+1} = \begin{pmatrix}
\Omega_{\phi_1, \ldots, \phi_L} & 0 \\
0 & 0
\end{pmatrix}$, where $\Omega$ is the estimated variance-covariance matrix of $(\phi_{N+1 L-1, \ldots, \phi_{N+1 L-1}})$ obtained via bootstrapping. Note that as soon as the Kalman filtering recursion starts, the effect of the chosen starting value diminish quicky. As the Kalman filter is an optimal state estimator applied to a dynamical system perturbed with random noise then the estimates obtained within the sample period for the LRF are also optimal.