Functional-PCA Based Seasonal Adjustment Time Series Models with Single and Mixed Frequencies Data

Wei Lin\textsuperscript{1}, Jianhua Huang\textsuperscript{2}, and Tucker S. McElroy\textsuperscript{3}

\textsuperscript{1}International School of Economics and Management, Capital University of Economics and Business, Beijing, China

\textsuperscript{2}Department of Statistics, Texas A\&M University, College Station, TX

\textsuperscript{3}Center for Statistical Research and Methodology, U.S. Census Bureau, Washington, D.C.

September 1st, 2015

Department of Economics, Texas A\&M University
1 Introduction
   - Motivation/Literature/Contribution

2 Methodology
   - Assumptions/W.N./ARMA/ARIMA/ARIMA + Mixed Freq.

3 Simulation Study
   - Deterministic Seasonal Component + Different Noises

4 Two Empirical Applications
   - The Apple Production of N.Z. / Online Submission

5 Current Works
   - Some Pending Questions
Outline

1 Introduction
   - Motivation/Literature/Contribution

2 Methodology
   - Assumptions/W.N./ARMA/ARIMA/ARIMA + Mixed Freq.

3 Simulation Study
   - Deterministic Seasonal Component + Different Noises

4 Two Empirical Applications
   - The Apple Production of N.Z. / Online Submission

5 Current Works
   - Some Pending Questions
Motivation

Figure 1: Logarithm Real GDP of China (1992Q1 – 2014Q4)
Motivation

Figure 2: Apple Production of New Zealand (1988Q1 – 2004Q4)
Motivation

Traditional X13 method is not satisfactory in many aspects:

- It can only deal with seasonality in monthly and quarterly frequencies;
- It may fail to provide plausible results in some circumstances;
- It cannot deal with seasonal time series with multiple frequencies, which may emerge because
  - the sampling method changes;
  - government/support funding varies.
Generally speaking, there are two approaches on modeling seasonality:

- **stochastic**
  
  Seasonal ARIMA models, structural time series models, etc… (Box & Jenkins, 1970; Harrison & Stevens, 1976)

- **deterministic**
  
  Linear (or nonlinear) additive or multiplicative trends and seasonal components and an irregular noise component. (Shumway 1988; Brockwell & Davis 1991)

We take the *deterministic* approach.
The main contribution of this research work is to propose a series of new seasonal adjustment methods for single and mixed frequencies data based on Regularized Singular Value Decomposition (RSVD).

We expect that our proposed methods have the following advantages:

- Seasonal behavior is not restricted by specified models,
- Seasonality is fully data driven,
- more flexible and robust.
Outline

1. Introduction
   - Motivation/Literature/Contribution

2. Methodology
   - Assumptions/W.N./ARMA/ARIMA/ARIMA + Mixed Freq.

3. Simulation Study
   - Deterministic Seasonal Component + Different Noises

4. Two Empirical Applications
   - The Apple Production of N.Z. / Online Submission

5. Current Works
   - Some Pending Questions
For a seasonal time series $Y_t$, we assume that it comprises two components, a **deterministic seasonal component** $S_t$ and a **stochastic non-seasonal component** $E_t$, in additive form:

$$Y_t = S_t + E_t, \quad t = 1, \ldots, T.$$  

(1)
To rewrite the equation (1) into matrix form, we relabel the time index $t \equiv (i(t), j(t)) = (i(t) - 1)p + j(t)$ with index of season $j(t) \in \{1, \ldots, p\}$ and the number of season $i(t) \in \{1, \ldots, n\}$ with $n \equiv T/p$.

The seasonal component $S_t$ and time series $Y_t$ can be reshaped into $n \times p$ matrices:

$$Y = S + E = \begin{pmatrix} s_{1,1} & s_{1,2} & \cdots & s_{1,p} \\ s_{2,1} & s_{2,2} & \cdots & s_{2,p} \\ \vdots & \vdots & \ddots & \vdots \\ s_{i(t),j(t)} & \vdots & \cdots & s_{n,p} \\ \vdots & \vdots & \ddots & \vdots \\ s_{n,1} & s_{n,2} & \cdots & s_{n,p} \end{pmatrix} + \begin{pmatrix} e_{1,1} & e_{1,2} & \cdots & e_{1,p} \\ e_{2,1} & e_{2,2} & \cdots & e_{2,p} \\ \vdots & \vdots & \ddots & \vdots \\ e_{i(t),j(t)} & \vdots & \cdots & e_{n,p} \\ \vdots & \vdots & \ddots & \vdots \\ e_{n,1} & e_{n,2} & \cdots & e_{n,p} \end{pmatrix},$$

where we assume that the sum of each row of matrix $S$ is zero, i.e., $\sum_j s_{i,j} = 0$ for all $i = 1, \ldots, n$. 
Assumptions on Seasonality

For example, a simple time invariant seasonality is $s_t = v_j(t)$, then

$$
S = \begin{pmatrix}
1 \\
1 \\
\vdots \\
1
\end{pmatrix}_{n \times 1} \begin{pmatrix}
v_{0,1} & v_{0,2} & \cdots & v_{0,p}
v_{0,1} & v_{0,2} & \cdots & v_{0,p}
\vdots & \vdots & \ddots & \vdots 
v_{0,1} & v_{0,2} & \cdots & v_{0,p}
\end{pmatrix}_{1 \times p} = \begin{pmatrix}
v_{0,1} & v_{0,2} & \cdots & v_{0,p} \\
v_{0,1} & v_{0,2} & \cdots & v_{0,p} \\
\vdots & \vdots & \ddots & \vdots \\
v_{0,1} & v_{0,2} & \cdots & v_{0,p}
\end{pmatrix}.
$$

(2)
Assumptions on Seasonality

For another example, to allow seasonal effects changing over time but being representable by a linear combination of $r$ basis vectors, we have

$$ S = \begin{pmatrix} 1 & u_{1,1} & u_{1,2} & \cdots & u_{1,r} \\ 1 & \cdots & \cdots & \cdots \\ 1 & \cdots & \cdots & u_{i(t),k} & \cdots \\ \vdots & \cdots & \cdots & \cdots \\ 1 & u_{n,1} & u_{n,2} & \cdots & u_{n,r} \end{pmatrix} \begin{pmatrix} v_{0,1} & v_{0,2} & \cdots & v_{0,p} \\ v_{1,1} & v_{1,2} & \cdots & v_{1,p} \\ v_{2,1} & v_{2,2} & \cdots & v_{2,p} \\ \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots \\ v_{r,1} & v_{r,2} & \cdots & v_{r,p} \end{pmatrix} = U V^T. \quad (3) $$

For each $s_t$, we have

$$ s_t = \sum_{k=0}^{r} u_{i(t),k} v_{k,j(t)}. \quad (4) $$
Assumptions on Seasonality

In addition, we deem the seasonal matrix $S$ as two-way functional data whose row and column domains are both structured:

- The $i$-th row of $S$ is the seasonal behavior of series $Y_t$ during the $i$-th season span.
- Each column of $S$ characterizes the variation of seasonality across time.

We assume that the shape of seasonality should be relatively robust in the sense that:

- it only changes smoothly, or
- it changes smoothly but with some structural breaks across time.

It means that the elements in each column of matrix $U$ in (3) only changes smoothly or with some breaks across time.
Assumptions on Non-seasonal Component

With different assumptions on the stochastic non-seasonal component $E_t$: 

- White noise,
- ARMA noise,
- ARIMA noise,
- ARIMA noise with mixed frequencies,

the estimation procedures are different.
White Noise

The stochastic non-seasonal component $E_t$ is a white noise, i.e.,

$$Y_t = S_t + E_t, \quad E_t \sim W.N.(0, \sigma^2), \quad t = 1, \ldots, T.$$  (5)

Borrowing the idea of Huang, Shen, and Buja (2008) and Huang, Shen, and Buja (2009), we penalize the roughness of left singular vectors in the classical SVD problem,

$$\|Y - uv^\top\|_F^2 + \alpha u^\top \Omega u$$  (6)

where $u = (u_1, \cdots, u_n)^\top$, $v = (v_1, \cdots, v_p)^\top$ with $v^\top v = 1$ for identification, $\| \cdot \|_F$ is the Frobenius norm, $\Omega$ is a non-negative definite roughness penalty matrix, and $\alpha > 0$ a penalty parameter.
White Noise

Algorithm (NB)

Suppose there is no break in smoothness of \( u \):

**Step 0.** Column-wise demean for matrix \( Y \) (column means are the estimates \( \hat{v}_1 \)).

**Step 1.** Initialize \( u \).

**Step 2.** Repeat until convergence:

1. \( v \leftarrow \left( I_p - \frac{1}{p} \ell_p \ell_p^T \right) \frac{Y^T u}{u^T u} \),
2. \( v \leftarrow \frac{\|v\|}{\|v\|} \),
3. \( u \leftarrow (I + \alpha^* \Omega)^{-1} Y v \) with \( \alpha^* \) selected from the following generalized cross-validation,

\[
\alpha^* = \arg \min \alpha \frac{1}{n} \frac{\| (I - M(\alpha)) Y v \|^2}{\left( 1 - \frac{1}{n} \text{tr}\{M(\alpha)\} \right)^2}
\]  

where \( I \) is \( n \times n \) identity matrix, \( M(\alpha) = (I + \alpha \Omega)^{-1} \) is smoothing matrix.
White Noise

Algorithm (B)

Suppose there is a break in smoothness of $u$ after $n_1$ seasonal cycles:

Step 0. Column-wise demean for matrix $Y$ (column means are the estimates $\hat{v}_1$).

Step 1. Initialize $u$.

Step 2. Repeat until convergence:

1. $v \leftarrow \left( I_p - \frac{1}{p} p \ell_p \ell_p^T \right) \frac{Y^T u}{u^T u}$,
2. $v \leftarrow \frac{v}{\|v\|}$,
3. $u_1 \leftarrow (I_{n_1} + \alpha_1^* \Omega_1)^{-1} Y_1 v$ and $u_2 \leftarrow (I_{n_2} + \alpha_2^* \Omega_2)^{-1} Y_2 v$ with $\alpha_1^*$ and $\alpha_2^*$ selected separately from the similar generalized cross-validations in Algorithm (NB), where $n_1 + n_2 = n$ and $u = (u_1^T, u_2^T)^T$. 
White Noise

Apply Algorithm (NB) or (B) with $r$ times to matrix $Y$, we obtain $r$ pairs of singular vectors, concatenate them into the $n \times (r + 1)$ matrix $\hat{U} \equiv (1, \hat{u}_1, \ldots, \hat{u}_r)$ and the $p \times (r + 1)$ matrix $\hat{V} \equiv (\hat{v}_0, \hat{v}_1, \ldots, \hat{v}_r)$. Then, $\hat{S} \equiv \hat{U}\hat{V}^\top$ is the estimate of the seasonal matrix $S$. 
ARMA Noise

The non-seasonal component $E_t$ is allowed to have ARMA time dependence:

$$Y_t = S_t + E_t, \quad E_t \sim \text{ARMA}(p,q), \quad t = 1, \ldots, T.$$  \hspace{1cm} (8)

The estimate of singular values ($U$ and $V$) and ARMA parameter $\theta$ can be obtained by minimizing:

$$(\hat{U}, \hat{V}, \hat{\theta}) \equiv \arg \min_{U,V,\theta} \log |\Sigma(\theta)| + [Y_T - \text{Vec}(UV^\top)]^\top \Sigma(\theta)^{-1} \times [Y_T - \text{Vec}(UV^\top)]$$

$$+ \sum_{j=1}^{r} \alpha_j u_j^\top \Omega u_j,$$  \hspace{1cm} (9)
**ARMA Noise**

- ARMA dependences of $E_t$ introduces extra cyclical behaviors to $Y_t$. However, as long as $E_t$ is weakly stationary, Algorithm (B) and (NB) are still robust.

- We propose an alternative two-step estimation procedure that also consistently estimate the seasonality and ARMA parameters:
  
  1. **Step 1**: use Algorithm (NB) or (B) to obtain $\hat{U}$, $\hat{V}$, and the *pilot estimate* of seasonal component as $\hat{S} = \hat{U}\hat{V}^\top$.
  
  2. **Step 2**: obtain the estimate of nonseasonal component as $\hat{E} = Y - \hat{S}$ into vector $\hat{E}_t$, and fit it with ARMA model to obtain $\hat{\theta}$.

  3. **Step 3**: Given the estimates $\hat{U}$ and $\hat{\theta}$, a more efficient estimate $\tilde{V}$ can be obtained by solving the following minimization problem: 

$$
\tilde{V} \equiv \arg \min_V \left[ Y_T - \text{Vec}(\hat{U}\hat{V}^\top) \right]^\top \Sigma(\hat{\theta})^{-1} \left[ Y_T - \text{Vec}(\hat{U}\hat{V}^\top) \right]
$$

such that $I_p^\top V = 0_{1 \times (r+1)}$

with the *final estimate* of seasonal component as $\tilde{S} = \hat{U}\tilde{V}^\top$. 

ARIMA Noise

Similarly, we consider a nonstationary case in which the nonseasonal component follows an ARIMA process:

\[ Y_t = S_t + E_t, \quad E_t \sim \text{ARIMA}(p, 1, q), \quad t = 1, \ldots, T. \]  (10)

The estimate of singular values (\( \hat{U} \) and \( \hat{V} \)) and ARMA parameter \( \hat{\theta} \) can be obtained by minimizing:

\[
(\hat{U}, \hat{V}, \hat{\theta}) \equiv \arg\min_{U,V,\theta} \log |\Sigma_\Delta(\theta)| + [\Delta Y_T - \Delta \text{Vec}(UV^T)]^T \Sigma_\Delta(\theta)^{-1} \\
\times [\Delta Y_T - \Delta \text{Vec}(UV^T)] + \sum_{j=1}^{r} \alpha_j u_j^T \Omega u_j,
\]  (11)
ARIMA Noise

However, ARIMA dependence renders the estimation procedure for weakly stationary ARMA case invalid.

Suppose $E_t$ follows an $I(1)$ process. The matrix $E$ also has smooth nonstationary trends on each column. Consider the first column in matrix $E$, $(e_{1,1}, e_{2,1}, \ldots, e_{n,1})^T$, we have

\[
e_{1,1} = \varepsilon_{1,1},
\]
\[
e_{2,1} = e_{1,1} + \sum_{j=2}^{p} \varepsilon_{1,j} + \varepsilon_{2,1},
\]
\[\ldots\]
\[
e_{n,1} = e_{n-1,1} + \sum_{j=2}^{p} \varepsilon_{n-1,j} + \varepsilon_{n,1}.
\]

It means that the first column of matrix $E$ also follows an $I(1)$ process. Similarly, it can be shown that all the $p$ columns of the matrix have nonstationary trends.
An easy way to distinguish the smooth variation of seasonality $S$ from the nonstationary nonseasonal component $E$ is to take column-wise first difference of matrix $Y$.

\[
Y^\dagger \equiv Y\Delta_c \equiv \begin{pmatrix}
y_{1,2} - y_{1,1} & \cdots & y_{1,p} - y_{1,p-1} \\
y_{2,2} - y_{2,1} & \cdots & y_{2,p} - y_{2,p-1} \\
\vdots & \vdots & \vdots \\
y_{n,2} - y_{n,1} & \cdots & y_{n,p} - y_{n,p-1}
\end{pmatrix}
\]

\[
\begin{align*}
S\Delta_c + E\Delta_c & \equiv S^\dagger + E^\dagger \\
\begin{pmatrix}
s_{1,2} - s_{1,1} & \cdots & s_{1,p} - y_{1,p-1} \\
s_{2,2} - s_{2,1} & \cdots & s_{2,p} - y_{2,p-1} \\
\vdots & \vdots & \vdots \\
s_{n,2} - s_{n,1} & \cdots & s_{n,p} - y_{n,p-1}
\end{pmatrix} & + \begin{pmatrix}
e_{1,2} - e_{1,1} & \cdots & e_{1,p} - e_{1,p-1} \\
e_{2,2} - e_{2,1} & \cdots & e_{2,p} - e_{2,p-1} \\
\vdots & \vdots & \vdots \\
e_{n,2} - e_{n,1} & \cdots & e_{n,p} - e_{n,p-1}
\end{pmatrix}.
\end{align*}
\]
Note that $S = UV^\top$, and

$$S^\dagger = S\Delta_c = UV^\top\Delta_c = U(\Delta_c^TV)^\top \equiv UV^{*\top},$$

where

$$V^\top \equiv \begin{pmatrix} v_{1,1} & v_{1,2} & \ldots & v_{1,p} \\ v_{2,1} & v_{2,2} & \ldots & v_{2,p} \\ \vdots & \vdots & \ddots & \vdots \\ v_{r,1} & v_{r,2} & \ldots & v_{r,p} \end{pmatrix},$$

$$V^{*\top} \equiv V^\top\Delta_c = \begin{pmatrix} v_{1,2} - v_{1,1} & \ldots & v_{1,p} - v_{1,p-1} \\ v_{2,2} - v_{2,1} & \ldots & v_{2,p} - v_{2,p-1} \\ \vdots & \ddots & \vdots \\ v_{r,2} - v_{r,1} & \ldots & v_{r,p} - v_{r,p-1} \end{pmatrix}.$$
We propose the following estimation procedure:

- **Step 1:** Apply Algorithm (NB) or (B) to matrix $\mathbf{Y}^\dagger$, and obtain the estimated left and right singular values $\hat{\mathbf{U}}$ and $\hat{\mathbf{V}}^*$.  

- **Step 2:** solve the following minimization problem:

$$
\hat{\mathbf{V}} \equiv \mathbf{V} \arg \min \left[ \Delta \mathbf{Y}_T - \Delta \text{Vec}(\hat{\mathbf{U}}\mathbf{V}^\top) \right]^\top \left[ \Delta \mathbf{Y}_T - \Delta \text{Vec}(\hat{\mathbf{U}}\mathbf{V}^\top) \right] \\
\text{s.t. } \mathbf{V}^\top \mathbf{l}_p = \mathbf{0}_{1 \times (r+1)}, \tag{13}
$$

where $\Delta$ is the typical first order difference operator. Obtain the *pilot estimate* of seasonal component as $\hat{\mathbf{S}} = \hat{\mathbf{U}}\hat{\mathbf{V}}^\top$. 

Step 3: Obtain the estimate of nonseasonal component as $\hat{E} = Y - \hat{S}$, rewrite it into vector $\hat{E}_t$, and fit it with ARIMA model to obtain $\hat{\theta}$.

Step 4: using the estimate $\hat{\theta}$, we can obtain a more efficient estimate $\tilde{V}$ by solving the following generalized least squares problem,

$$\tilde{V} = \arg\min_{V} \left[ \Delta Y_T - \Delta \text{Vec}(\hat{U}V^\top) \right]^\top \Sigma_{\Delta}(\hat{\theta})^{-1} \left[ \Delta Y_T - \Delta \text{Vec}(\hat{U}V^\top) \right]$$

$$\text{s.t. } \iota_p^\top V = 0_{1 \times (r+1)}, \quad (14)$$

where $\Sigma_{\Delta}(\hat{\theta})$ is the estimated covariance matrix of differenced nonseasonal component $\Delta \hat{E}_t$, and the final estimate of seasonal component is $\tilde{S} = \hat{U}\hat{V}^\top$. 
ARIMA Noise with Mixed Frequencies

We consider the seasonal time series $Y_M$ with mixed frequencies, and the corresponding partially observable highest frequency time series is $Y_H$ with the following linear sample relationship:

$$Y_M = JY_H.$$  

The time series $Y_H$ has deterministic seasonal component $S_t$ and stochastic nonseasonal component $E_t$ following an nonstationary ARIMA process:

$$Y_t = S_t + E_t, \quad E_t \sim \text{ARIMA}(p, 1, q), \quad t = 1, \ldots, H. \quad (15)$$
ARIMA Noise with Mixed Frequencies

The singular values $\mathbf{U}$ and $\mathbf{V}$ for seasonal component and ARIMA parameter $\theta$ in (15) can be estimated by minimizing the following negative penalized quasi-likelihood function:

$$(\hat{\mathbf{U}}, \hat{\mathbf{V}}, \hat{\theta}) \equiv \arg \min_{\mathbf{U}, \mathbf{V}, \theta} \log |\mathbf{B} \Sigma_\Delta(\theta) \mathbf{B}^T| + [\mathbf{B} \Delta Y_H - \mathbf{B} \Delta \text{Vec}(\mathbf{U} \mathbf{V}^T)]^T$$

$$\times [\mathbf{B} \Sigma_\Delta(\theta) \mathbf{B}^T]^{-1} [\mathbf{B} \Delta Y_H - \mathbf{B} \Delta \text{Vec}(\mathbf{U} \mathbf{V}^T)]$$

$$+ \sum_{j=1}^{r} \alpha_j \mathbf{u}_j^T \Omega \mathbf{u}_j,$$  \hspace{1cm} (16)

Here we borrow the idea from McElroy and Monsell (2012) that, after proper linear transformation on the time series with mixed frequencies, the regression model can be estimated by quasi-MLE method.
ARIMA Noise with Mixed Frequencies

- When $Y_M = Y_H$, we can easily verify that the model degenerates to the ARIMA case in (10) and (11).

- When $Y_M \neq Y_H$, jointly estimating $U$, $V$, and $\theta$ is very difficult. However, if the left singular values $U$ were known, the model is reduced to the regression with mixed frequencies discussed in McElroy and Monsell (2012).

We only need to propose a estimation method for $U$ first, then we can directly use the estimation method and algorithm given by McElroy and Monsell (2012) to estimate the remaining $V$ and $\theta$. 
ARIMA Noise with Mixed Frequencies

Although the elements of matrix $\mathbf{Y}_H$ in highest frequency are not fully observable, we can always obtain its corresponding matrix $\mathbf{Y}_L$ in lowest frequency by right multiplying an aggregation matrix $\mathbf{A}$ to $\mathbf{Y}_H$:

$$\mathbf{Y}_L = \mathbf{Y}_H \mathbf{A},$$

where $\mathbf{A}$ is an known $p_H \times p_L$ matrix and all the elements in $\mathbf{Y}_L$ are either directly observable or obtainable by simple linear combination from the original mixed frequencies series $\mathbf{Y}_M$. 

The seasonal matrices in low and high frequencies have the same linear relationship, $S_L = S_H A$. A key fact is that the singular value decompositions for $S_L$ and $S_H$ share the same left singular vectors $U$, and their right singular vectors maintain the linear relationship, i.e.:

$$S_L = UV_L^T = UV_H^T A = S_H A \quad \Rightarrow \quad V_L^T = V_H^T A.$$

Therefore, we can convert the time series $Y_M$ with mixed frequencies $p_L$ and $p_H$ into the corresponding time series $Y_L$ in low frequency $p_L$, use the estimation method for the ARIMA case to obtain the estimate of left singular vectors $\hat{U}$ for $Y_L$. 

(CUEB-TAMU-USCB)
Now we propose the following estimation procedure for estimating seasonal component in time series with mixed frequencies:

- **Step 1:** obtain the time series $Y_L$ with low frequency $p_L$ from the original mixed frequencies series $Y_M$.

- **Step 2:** apply Algorithm (NB) or (B) to matrix $Y_L$, and obtain the estimated left singular vectors $\hat{U}$. 
ARIMA Noise with Mixed Frequencies

Step 3: solve the following minimization problem (McElroy and Monsell, 2012):

\[
(\tilde{V}_H, \tilde{\theta}) \equiv \arg \min_{V, \theta} \log |B \Sigma_{\Delta}(\theta)B^T| + \left[ B \Delta Y_H - B \Delta \text{Vec}(\hat{U} V^T) \right]^T \times [B \Sigma_{\Delta}(\theta)B^T]^{-1} \left[ B \Delta Y_H - B \Delta \text{Vec}(\hat{U} V^T) \right]
\]

such that \( u_{PH}^T V = 0_{1 \times (r+1)} \)

to obtain the estimate of right singular vectors \( \tilde{V}_H \) for the time series in high frequencies \( Y_H \). Then, the final estimate of seasonal component \( S_M \) with mixed frequencies is

\[
\hat{S}_M = J \text{Vec}(\hat{U} \tilde{V}_H^T)
\]

in vector form.
Outline

1 Introduction
   - Motivation/Literature/Contribution

2 Methodology
   - Assumptions/W.N./ARMA/ARIMA/ARIMA + Mixed Freq.

3 Simulation Study
   - Deterministic Seasonal Component + Different Noises

4 Two Empirical Applications
   - The Apple Production of N.Z. / Online Submission

5 Current Works
   - Some Pending Questions
Seasonal Component

We consider a deterministic monthly seasonal component

\[ S_t \equiv S_{i,m} = \left(1 + \frac{i}{10}\right) a_m \]

where \( i = 1, \ldots, 50 \) and \( m = 1, \ldots, 12 \) (therefore \( T = 600 \)) indicate year and month respectively, and the elements in vector \( a \equiv (a_1, \ldots, a_{12}) \) take following values,

\[
a = (-1.25, -2.25, -1.25, 0.75, -1.25, -0.25, 2.75, -0.25, 0.75, -0.25, 0.75, 1.75).
\]

The factor \( (1 + i/10) \) in \( S_t \) increases the magnitude of seasonal component slowly every year in linear fashion.
Simulated Seasonal Component

Figure 3: Simulated Seasonal Component
Then we consider three different data generating processes for stochastic non-seasonal component $E_t$:

(i) $E_t \sim i.i.d. N(0, \sigma^2)$, for all $t$,

(ii) $E_t \sim ARMA(1, 1)$, with $\phi = 0.8$ and $\psi = 0.1$ with innovation variance $\sigma^2 = 1$,

(iii) $E_t \sim ARIMA(1, 1, 1)$, with $\phi = 0.8$ and $\psi = 0.1$ with innovation variance $\sigma^2 = 0.04$. 
Signal-to-Noise Ratios

After we generate $E_t$, we scale the seasonal component by a constant $w$ such that

$$w = \frac{s(E_t)}{s(S_t)} \kappa$$

where $\kappa$ is the signal-to-noise ratio, and we set a series of values in each simulation.

- For the stationary cases (i) and (ii),
  $$\kappa = (0.2, 0.4, 0.6, 0.8, 1, 1.2, 1.4, 1.6, 1.8, 2);$$
- for the nonstationary case (iii) $\kappa = (0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1).$
Criterion Functions

We use two criteria to evaluate and compare our proposed seasonal adjustment methods with X13 method on the accuracy of estimation of the seasonal component. They are Mean Square Error (MSE) and Mean Percentage Error (MPE):

\[
MSE = E[(\hat{S}_t - S_t)^2],
\]

\[
MPE = E \left| \frac{\hat{S}_t - S_t}{S_t} \right| \times 100%,
\]

where \( \hat{S}_t \) is the estimate.
Criterion Functions

In the simulation, the sample values of these two criteria are obtained by

\[
\overline{MSE} = \frac{1}{B} \sum_{b=1}^{B} \left( \frac{1}{T} \sum_{t=1}^{T} \left( \hat{S}_{t}^{(b)} - S_{t} \right)^{2} \right),
\]

\[
\overline{MPE} = \frac{1}{B} \sum_{b=1}^{B} \left( \frac{1}{T} \sum_{t=1}^{T} \left| \frac{\hat{S}_{t}^{(b)} - S_{t}}{S_{t}} \right| \right) \times 100\%,
\]

where \( b \) is the \( b \)-th replication and \( B \) is the total number of replication, and \( B = 200 \) in our simulation exercise.
Results for White Noise

Figure 4: $E_t \sim \text{i.i.d. } N(0, \sigma^2)$, for all $t$. Black circle line is X13, red triangle line is our proposed method.
Figure 5: $E_t \sim ARMA(1, 1)$, with $\phi = 0.8$ and $\psi = 0.1$ with innovation variance $\sigma^2 = 1$. Black circle line is X13, red triangle line is our proposed method.
Results for ARIMA Noise

Figure 6: $E_t \sim ARIMA(1, 1, 1)$, with $\phi = 0.8$ and $\psi = 0.1$ with innovation variance $\sigma^2 = 0.01$. Black circle line is X13, red triangle line is our proposed method.
Results for ARIMA Noise

Figure 7: $E_t \sim ARIMA(1, 1, 1)$, with $\phi = 0.8$ and $\psi = 0.1$ with innovation variance $\sigma^2 = 0.01$. Black circle line is X13, red triangle line is our proposed method.
Outline

1 Introduction
   - Motivation/Literature/Contribution

2 Methodology
   - Assumptions/W.N./ARMA/ARIMA/ARIMA + Mixed Freq.

3 Simulation Study
   - Deterministic Seasonal Component + Different Noises

4 Two Empirical Applications
   - The Apple Production of N.Z. / Online Submission

5 Current Works
   - Some Pending Questions
The Apple Production of N.Z.

Figure 8: Solid black line - actual series, dashed black line - X13 method, blue line - proposed method
Figure 9: Left and right singular vectors and their corresponding break points.
Logarithm of Online Submission Time Series

Figure 10: Solid black line - actual series, blue line - proposed method
Logarithm of Online Submission Time Series

Figure 11: Left and right singular vectors and their corresponding break points.
Outline

1 Introduction
   - Motivation/Literature/Contribution

2 Methodology
   - Assumptions/W.N./ARMA/ARIMA/ARIMA + Mixed Freq.

3 Simulation Study
   - Deterministic Seasonal Component + Different Noises

4 Two Empirical Applications
   - The Apple Production of N.Z. / Online Submission

5 Current Works
   - Some Pending Questions
Some Pending Questions

- More efficient algorithm with weakly stationary noise.
- Other simulation exercises for different types of noises.
- Detection of outliers.
Thank you very much!
Questions and Comments.