Fitting Replicated Multiple Time Series Models

Ross S. Bowden and Brenton R. Clarke*

Mathematics and Statistics,
School of Engineering and Information Technology,
Murdoch University, Murdoch, Australia, 6150

September 12, 2014

Abstract

This paper shows that the interleaving of replicated multiple time series allows the estimation methods available in standard multiple time series packages to be applied simultaneously to each of the replicated series without loss of information. The methodology employs a non-trivial multivariate extension of an earlier univariate result involving interleaving. The interleaving approach is used to model more

*Corresponding Author: Brenton R. Clarke, Mathematics and Statistics, School of Engineering and Information Technology, Murdoch University, Murdoch, Australia, 6150, email: B.Clarke@murdoch.edu.au.
than sixty years of daily maximum and minimum temperatures for Perth, Western Australia.

1 Introduction

Time series model fitting is typically undertaken for only one realisation of the time series process. However there are many situations where multiple independent replications of the same process are available for use in simultaneous modelling and estimation such as in chemical reactions and longitudinal sampling on multiple subjects. In both these examples the data is typically a multivariate vector at each recording point and is likely to be autocorrelated. This naturally suggests simultaneous multiple time series models, in general, and simultaneous vector autoregressive moving average (VARMA) models, in particular.

To facilitate replicated multiple time series modelling, this paper presents a method of representing repeated independent realisations of a VARMA process as one VARMA time series with the same dimension as each of the replicated series. The method uses a multivariate extension of the interleaving method introduced for ARMA processes in Bowden and Clarke (2012) and, conveniently, can be implemented using existing VARMA estimation software. The univariate interleaving method has been cited and employed by several researchers (see Guenther and Bradley (2013) as well as ElBakry et al. (2013), Luo et al. (2012) and Tan et al. (2012)).
In Section 2 this paper begins with a brief overview of VARMAX models which are the multivariate extension of the better known univariate (Box-Jenkins) ARMAX models. Section 3 explores the available literature on fitting one time series model to replicated realisations of the same multivariate process. Section 4 proves the multivariate extension of the theorem on interleaving in Bowden and Clarke (2012) and discusses its implications for model fitting. Section 5 applies the multivariate interleaving technique to over sixty years of daily maximum and minimum temperatures for Perth, Western Australia. Finally Section 6 presents recommendations for further research into estimation of replicated time series models.

In this paper the terms VARMA and VARMAX (VARMA with extraneous inputs) are used interchangeably. The term VARMA is employed in general to refer to multivariate extensions of autoregressive moving average models. However, when considered necessary to be explicit concerning extraneous inputs, the term VARMAX will be employed.

2 VARMAX Models

A Vector Autoregressive Moving Average process of order $p$ and $q$ with extraneous inputs (VARMAX($p,q$)), $\{x_t\}_{t=1}^n$, is a $k$-dimensional multiple time series generated by the model,

$$\phi(B)(x_t - \mu_{x_t}(z_t)) = \theta(B)a_t$$  \hspace{1cm} (1)
where \( \{a_t\}_{t=1}^n \) is a series of independent identically distributed random error vectors with constant variance matrix, \( \Sigma_a \), \( E(a_t) = 0 \) for all \( t \) and \( E(a_t a'_u) = 0 \) for all \( t \neq u \). Also \( \phi(B) \) and \( \theta(B) \) are matrix polynomials in \( B \), the backshift operator, of order \( p \) and \( q \) respectively. The roots of \( \det(\phi(B)) = 0 \) and \( \det(\theta(B)) = 0 \) all lie outside the unit circle ensuring stationarity and invertibility respectively.

Typically the zeroth order matrix in the polynomial, \( \phi(B) \), is an identity matrix and similarly for \( \theta(B) \). In this case, \( \Sigma_a \) is of general symmetric positive-definite form and this specification results in a canonical formulation for the VARMAX model which allows for unique identification. It is often assumed that the innovations are multivariate normal (but this is not required for Theorem 1 in Section 4).

We will assume that \( \mu_{xt}(z_t)(= E(x_t|z_t)) = \psi z_t \) where \( \{z_t\}_{t=1}^n \) is a series of explanatory (input) vectors and \( \psi \) is the matrix of regression parameters.

The above is one form of standard VARMAX specification. We will now present an alternative standard specification partially dictated by the R package, dse, used in Section 5 but also to provide an arguably more informative model.

The VARMAX process could be described as being generated by a seemingly-unrelated regression model (see Johnston and Dinardo (1996)) with VARMA errors in that the series mean vector corrects the series mean level to zero before application of the VARMA filter. However if the extraneous variables are introduced on the right hand-side of (1) their influence on the time
series vector can only be assessed with knowledge of the VAR filter. This alternative expression of the mean level is,

\[ \phi(B)x_t = \Upsilon z_t + \theta(B)a_t, \]  

(2)

where \( \Upsilon = \phi(1)\psi \), that is,

\[ \psi = \phi(1)^{-1}\Upsilon. \]  

(3)

The software used in this paper to fit the replicated VARMAX model to daily temperatures (the R package, dse) uses the specification (2).

An alternative specification of \( \phi(B) \), \( \theta(B) \) and \( \Sigma_a \) is possible which allows unique identification (See Lutkepohl (2005) p. 447ff). This uses the unique Cholesky LDL decomposition of the innovations matrix, that is, \( \Sigma_a = LDL' \) where \( L \) is upper triangular with a unit diagonal (so-called "unitriangular") and \( D \) is a diagonal matrix.

This alternate specification is,

\[ (L^{-1} + L^{-1}\phi_1B + L^{-1}\phi_2B^2 + ... + L^{-1}\phi_pB^p)(x_t - \mu_x(z_t)) \]
\[ = (I + L^{-1}\theta_1LB + L^{-1}\theta_2LB^2 + ... + L^{-1}\theta_qLB^q)u_t. \]  

(4)
where \( \mathbf{u}_t = \mathbf{L}^{-1} \mathbf{a}_t \) and hence \( \mathbf{V}(\mathbf{u}_t) = \mathbf{D} \).

This now provides a representation with a diagonal innovations covariance matrix but where the zeroth order MA matrix is an identity matrix and the zeroth order AR matrix is upper unitriangular because \( \mathbf{L}^{-1} \) is upper unitriangular. This AR formulation explicitly makes the first element of \( \mathbf{x}_t \) (that is, \( x_{1t} \)) a linear function of elements \( (x_{2t}, \ldots, x_{kt}) \) as well as other elements of \( \mathbf{x}_t \) at non-zero lags. Similarly \( x_{2t} \) is a linear function of \( (x_{3t}, \ldots, x_{kt}) \) as well as other elements of \( \mathbf{x}_t \) at non-zero lags, and so on.

As an example, if we have a zero-mean VAR(2) process \((k = 2)\) then this can be expressed as,

\[
\begin{bmatrix}
1 & \phi'_{2,0} \\
0 & 1 \\
\end{bmatrix}
\begin{bmatrix}
x_{1t} \\
x_{2t} \\
\end{bmatrix}
+ \begin{bmatrix}
\phi'_{11,1} & \phi'_{12,1} \\
\phi'_{21,1} & \phi'_{22,1} \\
\end{bmatrix}
\begin{bmatrix}
x_{1t-1} \\
x_{2t-1} \\
\end{bmatrix}
+ \begin{bmatrix}
\phi'_{11,2} & \phi'_{12,2} \\
\phi'_{21,2} & \phi'_{22,2} \\
\end{bmatrix}
\begin{bmatrix}
x_{1t-2} \\
x_{2t-2} \\
\end{bmatrix}
= \begin{bmatrix}
u_{1t} \\
u_{2t} \\
\end{bmatrix}. \tag{5}
\]

So,

\[
x_{1t} = -\phi'_{12,0}x_{2t} - \phi'_{11,1}x_{1t-1} - \phi'_{12,1}x_{2t-1} - \phi'_{11,2}x_{1t-2} - \phi'_{12,2}x_{2t-2} + u_{1t} \tag{6}
\]
\[ x_{2t} = -\phi'_{21,1}x_{1t-1} - \phi'_{21,2}x_{2t-2} - \phi'_{22,1}x_{1t-2} - \phi'_{22,2}x_{2t-2} + u_{2t}. \] (7)

In the context of the application in Section 5, this formulation appears more natural as the \( x_{1t} \) and \( x_{2t} \) are recorded sequentially on the same day and is similar to that of a periodically correlated ARMA model (see Parzen and Pagano (1979)). The results presented in this paper use model (4).

### 3 Review of the Literature

A review of the limited literature on the analysis of replicated univariate time series was undertaken in Bowden and Clarke (2012). The paucity of published work extends into the multivariate domain.

Bowne and Nesselroade (2005) discuss the modelling of replicated VARMA processes from a psychometric perspective. The authors initially use a generalised mean and covariance matrix for the moments of the time series vector. This is then reduced to a restricted set of parameters using a VARMA model and a result from du Toit and Browne (2007). In the latter the covariance matrix of the vector time series is represented as a closed-form function of the autoregressive matrices, of the variance of the initial (unobserved) system state and of the covariance matrix of the one-step ahead forecasting errors. This then facilitates estimation of the model parameters using maxi-
The mean of the process is modelled using Gompertz trend curves.

Feder (2001) modelled repeated multivariate time series data from sample surveys. The author employs conventional state space models and incorporates sample survey error (See also Beck (2001) and Abrahams and Vijayan (1992)).

Models of simultaneous multivariate time series processes are also employed in the analysis of econometric panel data. In this, data is collected on "panels" of participants (which are similar to longitudinal cohorts) and regression models are typically employed to extract the effects on the output variables, such as gross domestic product, of various input variables such as price and economic interventions (see Baltagi (2005), Croissant and Millo (2008) and Wooldridge (2010)).

Given that panel data is commonly of a time series nature, autocorrelation of the error term can be present (see Banerjee et al. (2010), Eberhardt (2011), Gorenflo (2013) and Kocenda and Cerny (2014)). These models, when including autoregressive terms, are typically fitted using fixed effects OLS and GLS methods which do not result in fully efficient estimation and required purpose-built software. Moving average models of the first order are typically fitted using iterative OLS algorithms which again do not result in fully efficient estimation and employ specially-written programs.
4 The Interleaving Method

In this section, we prove that replicated independent VARMAX processes can be represented as a single VARMAX process with the same dimension as each of the replicated series. This method is a convenient tool for model fitting and can be accommodated in existing VARMAX software.

4.1 Replicated VARMAX Process

Let the \(i^{th}\) replicated \(k\)-dimensional vector series over the time span, \(t = 1, \ldots, n\), be \(\{x_{i,t}\}_{t=1}^n, i = 1, \ldots, m\) and assume each series is generated by the following VARMAX\((p,q)\) model,

\[
\phi(B)(x_{i,t} - \mu_{x_{i,t}}(z_{i,t})) = \theta(B)a_{i,t}
\]

(8)

where \(\{a_{i,t}\}_{i=1}^n\) is a series of independent zero-mean identically-distributed random error vectors with \(E(a_{i,t}a'_{j,u}) = \Sigma_a\) for \(i = j\) and \(t = u\) and 0 otherwise. Hence the error vectors have a variance (matrix) of general form but otherwise the vectors are assumed to be independent between realisations at all lags and within realisations at all non-zero lags. Also \(\phi(B)\) and \(\theta(B)\) are matrix polynomials in \(B\), the backshift operator, of order \(p\) and \(q\) respectively and the (conditional) mean of \(x_{i,t}\) is,

\[
E(x_{i,t}|z_{i,t}) = \mu_{x_{i,t}}(z_{i,t}) = \psi z_{i,t},
\]
where \( \{z_{i,t}\}_{t=1}^{n}, i=1,\ldots,m \), are \( m \) series of explanatory vectors.

It is possible to effectively have a unique parameterisation of \( \psi \) (say \( \psi_i \)) for each realisation, \( i \), through simply reconfiguring (and expanding) the dimension of the input vector, \( z_{i,t} \), by a factor of \( m \). A binary (intervention) variable for each realisation is introduced in addition to the interactions of the binary variables with each of the original elements of \( \psi \).

We will call the time series (8) a RVARMA (replicated VARMA) process, that is, RVARMA(\( p,q,m \)). It has a mean which can vary with each series realisation but it otherwise maintains a consistent generating mechanism between realisations. In fact the mean can be any linear combination of the extraneous vector variables, \( z_{i,t} \) (and, in general, can be a non-linear function of the \( z_{i,t} \)).

### 4.2 Equivalent Replicated VARMAX Representation

We now state and prove a theorem that reduces the apparent dimensionality of the replicated process by a factor of \( m \).

**Theorem 1.** Let the replicated \( k \)-dimensional series \( \{x_{i,t}\}_{t=1}^{n}, i=1,\ldots,m \), be generated by the above RVARMA(\( p,q,m \)) process (see (8)), and let,

\[
\begin{align*}
y_{m(t-1)+i} &= x_{i,t}, \\
w_{m(t-1)+i} &= z_{i,t} \quad \text{and} \\
c_{m(t-1)+i} &= a_{i,t}.
\end{align*}
\]
Then,
\[
\phi(B^m)(y_s - \mu_{y_s}(w_s)) = \theta(B^m)e_s
\]  \hspace{1cm} (9)
where \(E(e_s) = 0\), \(V(e_s) = \Sigma_a\) and \(E(e_s' e_r') = 0\), \(s \neq r\). That is, the interleaved series, \(\{y_s\}_{s=1}^{mn}\), is a \(k\)-dimensional VARMA process of order \((mp,mq)\).

**Proof.** Consider the autoregressive moving average difference equation, (9) and select all \(s\) such that \(s|m\) (that is, \(s \mod m\)) equals some constant \(i\), that is, \(s = m(t-1) + i\). We then have \(y_s = x_{i,t}, w_s = z_{i,t}\), and \(e_s = a_{i,t}\), \(t = 1, \ldots, n\), and, setting \(D = B^m\),
\[
\phi(D)(x_{i,t} - \mu_{x_{i,t}}(z_{i,t})) = \theta(D)a_{i,t}
\]  \hspace{1cm} (10)
where \(D\) is equivalent to a one-lag backshift operator on \(x_{i,t}\) i.e. \(Dx_{i,t} = x_{i,(t-1)}\). We note that, solely from the specification of the interleaved model (9), \(E(e_s) = E(a_{i,t}) = 0\), and \(V(e_s) = V(a_{i,t}) = \Sigma_a\). If we now choose \(r\) \((= m(u-1) + j, j = r|m)\) such that \(r \neq s\) (that is, where \(i \neq j\) and/or \(t \neq u\)), then \(E(e_s e_r') = E(a_{i,t}a_{j,u}') = 0\) where \(i \neq j\) and/or \(t \neq u\). This completes the proof of equivalence between representations (9) and (1).

To paraphrase Theorem 1 (herein called the Multivariate Interleaving Theorem), any \(m\) replicated \(k\)-dimensional VARMA\((p,q)\) time series each of length \(n\) can be represented by one \(k\)-dimensional VARMA\((mp,mq)\) process of length \(mn\). This equivalence is achieved by interleaving the \(m\) series and by ensuring that AR and MA parameters are only non-zero at orders that are multiples of \(m\). The equivalence uses an interleaving which is illustrated
in Figure 1 for two artificial bivariate series, each of length seven.

The Multivariate Interleaving Theorem implies that, say, 10 realisations of a VARMA bivariate process of order \((p,q)\), each of length, say, 50, can be represented by a single VARMA bivariate process of order \((10p,10q)\) and length 500 with non-zero AR and MA parameter matrices for orders at multiples of 10 only. Using the notation in (1), the VARMA model to be fitted to the interleaved series, \(\{y_s\}_{s=1}^{500}\), is,

Figure 1: Multivariate interleaving of an artificial bivariate series.
\[
(I + \phi_1 B^{10} + \phi_2 B^{20} + ... + \phi_p B^{10p})(y_s - \mu_{y_s}(w_s))
\]
\[
= (I + \theta_1 B^{10} + \theta_2 B^{20} + ... + \theta_q B^{10q})e_s. \tag{11}
\]

In fitting (11), the AR and MA matrices of the following orders are set to zero,

\[
(1, ..., 9, 11, ..., 19, ..., 10(p - 1) + 1, ..., 10(p - 1) + 9). \tag{12}
\]

That is, only the matrices corresponding to lags, \((10, 20, ..., 10(p-1), 10p)\), are non-zero.

By using VARMA software such as R’s dse package (R Development Core Team (2010)), Scilab’s Grocer (Scilab (2010)) and Gauss’s Time Series MT (GAUSS (2010)) which all allow VAR and VMA matrix parameters to be set to zero the interleaving method can be employed in RVARMA model fitting without preparing purpose-built computer programs.

It may be suspected that the stability of the RVARMA estimation routine could be in question for large \(m(\sim 60)\). However it is the experience of the authors that this is not the case for the maximum likelihood approach. Of course, the interleaving method can also be applied to other estimation approaches including robust methods, least squares and the method of moments.
Likelihood ratio testing of unique parameters per series can be undertaken by using maximum likelihood to fit the single $RVARMA(p,q,m)$ model to the interleaved data and then fitting a separate $VARMA(p,q)$ model to each series. The likelihood ratio test then uses the $RVARMA$ likelihood versus the product of the likelihoods for each of the separate estimations.

5 Application to Daily Maximum and Minimum Temperatures for Perth, Australia

This section uses the dse package in R to fit the $RVARMA$ model with maximum likelihood estimation to sixty-seven\footnote{In Bowden and Clarke (2012) it was incorrectly reported that sixty-six years of data were analysed whereas sixty-seven years were actually modelled.} years of daily maximum and minimum temperature readings for Perth, Western Australia (to 2009). Amongst other uses, it provides an understanding of the relationship between maximum and minimum temperatures which is informative for forecasting hourly electricity and gas demand.

Figure 2 plots the daily maximum and minimum values for three years and it is clear that there is a strong relationship between values on the same day. The expected seasonal cycle is also evident as is an increase in the variability of the maximum temperatures over summer.

Given the increased variability in summer and the known changes in weather patterns between summer and winter it is likely that VARMA mod-
els of the bivariate daily temperature data vary over the year. Hence the modelling in Sections 5.1 to 5.3 was undertaken separately for each week (in the year) of daily data but simultaneously for all years. This follows the univariate analysis carried out previously for daily maximum temperatures in Bowden and Clarke (2012), again for each week in the year and simultaneously for all years.

The current modelling will explore the relationship between the daily maximum and minimum temperatures and the effects of the intervention terms for change of location and for trend. As with the original analysis the maximum and minimum temperatures were modelled in part to assess the effect of changes in location and of drift due to climate change. The five locations where the temperature data was collected are listed in Table 1.

Table 1: Change of Location for Perth’s Temperature Recording Device

<table>
<thead>
<tr>
<th>Location</th>
<th>Last Recording Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>King’s Park</td>
<td>August 1963</td>
</tr>
<tr>
<td>Old Hale School</td>
<td>June 1967</td>
</tr>
<tr>
<td>Wellington St</td>
<td>May 1992</td>
</tr>
<tr>
<td>Perth Airport</td>
<td>November 1993</td>
</tr>
<tr>
<td>Mt Lawley</td>
<td>December 2009</td>
</tr>
</tbody>
</table>

5.1 Model Identification with Interleaving

To begin the RVARMA model fitting, an interleaved bivariate series was created using the 67 years ($m = 67$) of daily maximum and minimum temperatures available for each week-in-the-year. The RVARMA model order
was deduced by examining the raw sample cross-correlations (see Chatfield (2004)), the prewhitened sample cross-correlation function (see Granger and Newbold (1986) p. 237ff and Jenkins and Alavi (1981)) and the sample partial lag autocorrelation function (Wei (2006)) after correction for all intervention effects (herein called detrending) and adjusting for interleaving. For the first week of interleaved data ($n = 67 \times 7$) these correlations are plotted in Figures 3, 4 and 5 respectively with the 95 percent confidence intervals shown as dashed lines. The full (and hollow - see below) points indicate the interleaved lags ($-3 \times 67, ..., 3 \times 67$) that correspond to lags -3 to 3 in the original series.

Detrending involved fitting a univariate AR(2) model separately to both daily maximum and minimum temperatures with associated intervention terms for the change in location and for the long-term trend in temperature (The order was suggested from univariate RARMA modelling - see Bowden and Clarke (2012)). The resulting intervention terms (without the autoregressive filters) were then used to correct the two univariate time series for the change in location and for the long-term increase in temperatures. The resultant zero-mean series were then prewhitened.

Prewhitening applied the AR(2) filter from the AR detrending model for daily minimum temperatures to the detrended daily maximum temperatures. The sample cross-correlations were then calculated for the two filtered series. This prewhitening applied the constraints imposed by interleaving, that is, only parameters for AR orders corresponding to multiples of the number of
Figure 2: Three years of daily maximum and minimum temperatures (in degrees Celsius). The x-axis values are the week-in-the-year.

replicated series (years) were non-zero.

As mentioned above, the sample partial lag correlation matrix was also employed in model selection (Wei (2006)). The sample partial lag correlation matrix is the sample cross correlation matrix at a lag of $k$ time intervals after removing the (linear) influence of the intervening lags. For an AR($p$) process the correlations cut off at lag $p$ as with the multivariate partial autoregression matrix.
The sample partial lag correlation matrix differs from the partial autoregression matrix (Tiao and Box (1981)) in that, as with the cross-correlation function, the absolute value of the former's elements can't exceed one. The original series is corrected for future values of the time series up $k - 1$ steps ahead and similarly the time series vector at lead $k$ is corrected for past data back to lag $k - 1$. The partial lag correlations, $\hat{\rho}_{x,y}(k)$, between say univariate elements, $x_{t+k}$ and $y_t$, of the vector series shares the property with the cross-correlation function of $\hat{\rho}_{x,y}(k) = \hat{\rho}_{y,x}(-k)$.

To accommodate the constraints impose by interleaving, the partial lag autocorrelations were derived by only fitting autoregressions to multiples of the number of replicated series (that is, of years). In the approach of Wei (2006) for estimating the partial lag autocorrelations, this implies fixing the (detrended) sample cross-correlations to zero for the other lags before using them in the estimation routine.

5.2 Identification Results

In Figure 3, the sample detrended cross-correlations between maximum and minimum temperatures for week one (as indicated by the black dots at lags that are multiples of 67) show a range of significant values. However the sample prewhitened detrended cross-correlations (in Figure 4) show significant values at lags 0 and 1 only (that is, lags 0 and 67 with interleaving). This demonstrates the ability of prewhitening to substantially simplify the model selection process within an interleaving paradigm.
Figure 3: Detrended cross-correlations using interleaving (week 1) with 95% confidence intervals. The black dots indicate the sample correlations at lags -3 to 3 in the original series.

Figure 4: Prewhitened detrended cross-correlations using interleaving (week 1) with 95% confidence intervals. The black dots indicate the sample correlations at lags -3 to 3 in the original series.

Figure 5: Wei's partial lag detrended autocorrelations using interleaving (week 1) with 95% confidence intervals. The black dots indicate the sample correlations at lags -3 to 3 in the original series. The circles are the same values uncorrected for assumed zero correlations at lags that are not multiples of the number of replicated series (67 in this case).
Sample partial lag autocorrelations are plotted in Figure 5. As with cross-correlations the values corresponding to lags $-3$ to $3$ in the original replicated series are marked by full dots. The values reveal significant correlations at lag one and arguably at lag two. The hollow dots indicate partial lag autocorrelations calculated without setting relevant intermediate correlations to zero.

Given that these correlation results were similar for all weeks, it was decided to fit a RVAR model of order two (that is, a RVAR(2,0,67) model) for each week. This was undertaken using the interleaving method from Section 4 (with $m=67$, that is, a total sample per week-in-the-year of $67 \times 7 = 469$) and employing the method of maximum likelihood via the R package, dse.

The dse package uses the VARMAX representation (2) but the results in this paper employ the representation (4), derived by applying the Cholesky LDL transformation from Section 2 and the transformation of the process mean from (3). Hence the VAR(2) models in this paper utilise an intervention vector that is the (conditional) mean of the process, a zeroth order VAR matrix that is upper unitriangular and a variance matrix of the innovations vector that is diagonal.

This permits what is arguably a simpler interpretation whereby the mean correction due to the intervention terms is applied directly to the vector time series before application of the AR filter, the daily maximum temperature is related to the (earlier) minimum temperature on the same day and the elements of the innovation vector are independent.
The standard errors of the sample parameter estimates for model (4) were not immediately available from the model fits and had to be derived after transformation from the fitted model (2) using simulation (10,000 simulations). To achieve this, the estimated innovations variance matrix for (2) is assumed to be independent of the other sample parameter estimates for (2). Repeated realisations of the sample variance matrix of the innovations from the fitted model (2) were simulated using bootstrapping on the model residuals vectors. The other parameter estimates from (2) were simulated using an assumption of multivariate normality where the mean vector is the VAR and intervention parameter estimates (in (2)) and the variance matrix is the associated Hessian-derived variance matrix.

Each set of combined simulated parameter estimates for (2) was transformed using the LDL transformation from Section 2 used in model (4) and the mean transformation (3). The empirical distribution of the resulting simulated parameters was then used to calculate the sample variance matrix of the parameter estimates for (4). The square roots of the diagonal elements of the matrix were used as standard errors of the transformed parameters estimates shown in Figures 6, 7, 8 and 9.

5.3 Estimation Results

The VAR parameters are plotted by week in Figure 6. Note that these are the negative of the VAR parameters from representation (4) to reflect their use in a predictive formulation.
It is clear that the parameters change smoothly and substantially over the year with the strongest relationship between maximum temperatures and past maximum and minimum temperatures in summer with little relationship in winter. The minimum temperatures show a much weaker set of relationships although the VAR parameters are now generally strongest in winter. Maximum temperatures exhibit positive AR parameters with respect to their own lagged values but negative parameters with past minimum temperatures. Minimum temperatures show a less well-defined mix of positive and negative AR parameters.

![Graphs of VAR(2) parameters by week with 95% confidence intervals.](image)

Figure 6: VAR(2) parameters by week with 95% confidence intervals.

The standard deviations of the (independent) innovations by week are
shown in Figure 7 and indicate that the variation of the maximum temperatures changes substantially over the year with the greatest standard deviation in summer. The minimum temperatures show a relatively unchanged standard deviation.

Figure 7: Innovation standard deviation by week with 95% confidence intervals.

The top left plot in Figure 8 shows the mean daily maximum temperatures by week for the base site used to 1963 and exhibit the expected seasonal cycle. The next four plots show the mean difference by week between the data recorded at the site for the period in question and the base site. The sixth plot shows the annual trend in maximum temperatures (by week-in-
the-year) over the 67 years to 2009.

The results suggest colder temperatures in summer for the Wellington St site (1967-92) and the current Mt Lawley site (1993-99). This implies that maximum temperatures recorded at Mt Lawley in summer are likely to underestimate the true value compared to the historical record from King’s Park (pre-1963). The results for the intervention terms for daily maximum temperatures are very similar to the univariate outcomes from Bowden and Clarke (2012)\(^2\).

The estimated annual (positive) trend in maximum temperatures from combining the fifty-two weeks’ results is 0.0147°C (±0.0096). This equates to 0.98°C (±0.64) as the total increase over 67 years.

For minimum temperatures (see Figure 9) the localised results are similar to that for maximum temperatures although there is additional evidence of lower minimum temperatures over the whole year at the current Mt Lawley site. Over the fifty-two weeks the mean difference compared to the King’s

\(^2\)It is also informative to compare the form of the temperature model’s VAR filter to the univariate ARMA filter from Bowden and Clarke (2012). Corollary 11.1.1 from Lutkepohl (2005) can be used to show that the univariate (marginal) time series from a multivariate VAR(2) process are ARMA(4,2). To demonstrate this, in the current bivariate case we define a vector, \(F = [1, 0]\), and hence \(M = 1\) and \(K = 2\) as used in Lutkepohl (2005). So the (univariate) process, \(u_t = F \epsilon_t\), is an ARMA(\(\hat{p}, \hat{q}\)) process where \(\hat{p} \leq (K-M+1)p = 2p\) and \(\hat{q} \leq (K-M)p + q = (2-1)p + q = p + q\) (The inequalities accommodate the potential cancellation of AR and MA terms). Hence with \(p = 2\) and \(q = 0\) (from our VARMAX(2,0) models) the marginal ARMA models are of order at most \((4, 2)\), that is, ARMA(4,2). This compares to the AR(2) models identified and fitted in Bowden and Clarke (2012).

Although not undertaken here it is possible to calculate the parameters of the univariate marginal process from the VARMA parameters by equating the autocovariances of the marginal ARMA models to the diagonal elements of the cross-covariance matrix of the VARMA process.
The Park site is $-1.84 \, ^\circ C \, (\pm 0.35)$. The (positive) annual trend in minimum temperatures is $0.0179 \, ^\circ C \, (\pm 0.0069)$ or $1.20 \, ^\circ C \, (\pm 0.46)$ as a total over 67 years.

It is proposed in future research to extend the multivariate modelling of the daily maximum and minimum temperatures to consider periodically-correlated RVARMA models (see Parzen and Pagano (1979) for periodically-correlated univariate models). These models describe situations where the parameters of the VARMA representation of the process vary deterministically according to some cycle.

In this, it is planned to vary the parameters by days-of-the-week to mimic the known cycling of other weather parameters (see Cerveny and Balling (1998) for results regarding the mean of daily rainfall). The models will be fitted by month-in-the-year (rather than by week) using the same 67 years of weather data for Perth. This will increase the sample size to compensate for the larger dimension of the parameter space.

6 Conclusion

In this paper it has been shown that the interleaving method allows replicated realisations of the same VARMA process to be modelled as a single VARMA process of the same dimension as each of the original series but with extended length. This has advantages in parameter estimation in that existing VARMA-fitting software can be used. The multivariate interleaving
Figure 8: Estimates by week for maximum temperatures of the effect of change of location and of trend over time (with average annual trend) with 95% confidence intervals. The plot of annual trend also shows the overall mean annual change as a horizontal full line.

approach was demonstrated on daily maximum and minimum daily temperatures for Perth, Western Australia.

It is proposed to undertake further research in the area of replicated time series by modelling replicated series that vary in length and that have differences in the scale of their innovation variance matrix. This will involve multivariate models similar to the univariate models employed for spectral estimation by Quinn (2006).

Acknowledgement: The authors would like to acknowledge the work
Figure 9: Estimates by week for minimum temperatures of the effect of change of location and of trend over time (with average annual trend) with 95% confidence intervals. The plot of annual trend also shows the overall mean annual change as a horizontal full line.

of Dr Paul Gilbert, the original author of the R multiple time series package, dse, used in the final model fitting. We also acknowledge the assistance of Mr Éric Dubois, the author of the Scilab econometrics package, Grocer, used at an early stage of the current research. The authors are indebted to Associate Professor Robin Milne and Professor Christine Müller for conversations leading to the discussion of econometric panel data in this paper.
References


Cerveny, R. and Balling, R. (1998). Weekly cycles of air pollutants, precip-


