USING VARX MODEL TO FORECAST THE EFFECT OF CULTIVATION OF MAIN CROPS ON WATER SECURITY IN SAUDI ARABIA

Othman S. Alnashwan, Mahmoud M. Alderiny*

Department of Agricultural Economics, College of Food and Agriculture Sciences, King Saud University, Riyadh 11451, Saudi Arabia

Received – March 14, 2017; Revision – May 28, 2017; Accepted – June 13, 2016

Available Online – August 31, 2017

DOI: http://dx.doi.org/10.18006/2017.5(Spl-1-SAFSAW).S126.S135

ABSTRACT

Augmented vector autoregressive model was applied to analyze the dynamic relationship between three-time series which includes the areas cultivated with dates, clover and fodder in Saudi Arabia. The application study relied on time series data from 1986 to 2013. The unit root test indicated that the areas cultivated with the three crops are stationary at the first differences. Information criteria showed the optimal lag and according to results of Wald test, the model can be used. The implementation of resolution No. 335 and the linear trend has significant effect on the first difference of the areas cultivated with clover as well as cultivated with fodder. Model was used for forecasting the areas from 2014 to 2018, and noted that the areas cultivated with dates will annually increase with rate 0.3% and the expected mean of water consumption is 2837.6 million m³ annually, the areas cultivated with clover will annually decrease with rate 1.6% and the expected mean of water consumption is 3708.4 million m³ annually, and the areas cultivated with fodder will annually decrease with rate 15% and the expected mean of water consumption is 3999.6 million m³ annually.

KEYWORDS

VARX Mod
Stationary
Unit root test
Criteria Information
Wald statistic
Forecasting
Mean Square Error

* Corresponding author
E-mail: malderiny@ksu.edu.sa (Mahmoud M. Alderiny)

Peer review under responsibility of Journal of Experimental Biology and Agricultural Sciences.
1 Introduction

The application of the decision to stop the cultivation of wheat in the Saudi Arabia to reduce water consumption, came inversely, where, led, according to experts in the agriculture sector to increase consumption by about 400 percent, due to the shift towards increasing the areas cultivated with green fodder (clover and fodder), which consumes large amounts of water (Economic Daily, 2010).

The aims of the research can be summarized in two points: the first, using the appropriate vector autoregressive (VAR) model, augmented by the exogenous, variables, labeled (VARX) model for studying the dynamics of areas cultivated with three main crops (dates, clover, and fodder), through the dependence of each crop area on all other crops areas at lag periods in addition the implementation of the decision No. (335) about stopping cultivation of wheat, and the second is using the selected (VARX) model for forecasting these areas through five years after last year in the time series under the study.

The augmented vector autoregressive (VARX) model is one of the most successful, flexible, and easy to use type of these models for the analysis of multivariate time series. The (VARX) model has proven to be useful especially for describing the dynamic behavior of economic and financial time series as well as forecasting. It often provides superior forecasts to those from univariate time series models and elaborate theory-based simultaneous equations models. Forecasts from (VARX) models are quite flexible because they can be made conditional on the potential future paths of specified variables in the model.

Applications of (VAR) models to financial data are given in Hamilton (1994), Cuthbertson (1996), Campbell et al. (1997), Mills (1999) and Tsay (2001). Athanasopoulos et al. (2011) studied the joint determination of the lag length, the dimension of the integrating space and the rank of the matrix of short-run parameters of a vector autoregressive (VAR) model using model selection criteria. They suggested a new two-step model selection procedure which is a hybrid of traditional criteria and criteria with the integrating space and the rank of the matrix of short-run parameters, determined dependent variables, the basic -lag augmented vector autoregressive (VARX) model written as:

\[ Y_t = \Psi Z_t + \Phi_1 Y_{t-1} + \Phi_2 Y_{t-2} + \ldots + \Phi_p Y_{t-p} + \varepsilon_t, \quad t = p+1, p+2, \ldots, T \]

(2.1)

Where \( Z_t \) is a \( l \times 1 \) vector of deterministic and/or exogenous variables, \( \{\Phi_i, i=1,2,\ldots,p\} \) and \( \Psi \) are \( n \times n \) and \( n \times l \) coefficient matrices, and \( \varepsilon_t \) is a sequence of \( n \times 1 \) independent white noise vectors with zero mean and nonsingular contemporaneous covariance matrix given by \( \Sigma_\varepsilon \). Hence, the VARX model is just a seemingly unrelated regression (SUR) model with lagged variables and deterministic and/or exogenous variables as common regressors. The VARX model (2.1) is based on some assumptions determinant Lütkepohl (1991) and Pesaran & Pesaran (1997):

**Assumption 1:** \( E(\varepsilon_t) = 0 \), \( E(\varepsilon_t\varepsilon_t') = \Sigma_\varepsilon \) for all \( t \), where \( \Sigma_\varepsilon = \sigma_\varepsilon^2 \delta_j, i, j = 1,2,\ldots, n \) is an positive definite matrix, \( E(\varepsilon_t\varepsilon_{t'}) = 0 \) for all \( t \neq t' \), and \( E(\varepsilon_t | Z_t) = 0 \).
Assumption 2: All the roots of \( |\Phi(L)|=|I_n - \sum_{j=1}^{p} \Phi_j L^j| = 0 \) fall outside the unit circle, equivalently, all eigenvalues of the \( np \times np \) companion matrix

\[
\Pi = \begin{bmatrix}
\Phi_1 & \Phi_2 & \ldots & \Phi_{p-1} & \Phi_p \\
I_n & 0 & \ldots & 0 & 0 \\
0 & I_n & \ldots & 0 & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & \ldots & I_n & 0
\end{bmatrix}
\]

have modulus less than one, and \( I_n \) is \( n \times n \) identity matrix.

Assumption 3: \( (Y_{t-1}, Y_{t-2}, \ldots, Y_{t-p}, Z_t), t = p+1, p+2, \ldots, T \) are not perfectly collinear.

Under Assumption 2, \( Y_t \) would be covariance-stationary, and equation (2.1) can be rewritten as the infinite moving average representation,

\[
Y_t = \sum_{j=0}^{\infty} \Lambda_j e_{t-j}, \quad t = p+1, p+2, \ldots, T
\]

(2.2)

Where, the \( n \times n \) coefficient matrices \( \Lambda_j \) can be obtained from solving the relation \( \Phi(L)^t(\Lambda(L)) = I_n \), and computed by using the following recursive relations:

\[
\Lambda_j = \Sigma \Lambda_{j-i} \Phi_j
\]

(2.3)

Where, \( \Lambda_0 = I_n \) and \( \Phi_j = 0 \) for \( j > p \), and \( \Gamma_j = \Lambda_j \Psi \).

2.1 Estimation

Consider the basic VARX(\( p \)) model (2.1). Under assumptions 2 and assuming that there are no restrictions on the parameters of the model, the general form of the multivariate linear model represented by

\[
Y_t = X_t \beta + E_t
\]

(2.4)

Where, \( Y_t = (Y_{t+p+1}, Y_{t+p+2}, \ldots, Y_{t+T})', \quad X_t = (X_{t+p+1}', \ldots, X_T'), \quad E_t = (e_{t+p+1}, e_{t+p+2}, \ldots, e_{T+t})', \quad m = T - p \), and \( k = np + l \). Then the conditional least squares (LS) estimator of \( \beta \) given by Johnson & Wichern (1992) as

\[
\hat{\beta}_L = (XX')^{-1}XY
\]

(2.5)

and the estimate of \( \Sigma_e \) is

\[
\hat{\Sigma}_e = \frac{\sum_{i=p+1}^{T} \tilde{e}_i e_i'}{m-k}
\]

(2.6)

Where, \( \tilde{e}_i \) is the residual vector. Let \( \hat{\beta}_L = \text{vec}( \hat{\beta}_L ) \) denotes the operator that stacks the columns of the \( (k \times n) \) matrix \( \hat{\beta}_L \) into a long \( (nk \times 1) \) vector, then under standard assumptions regarding the behavior of stationary and ergodic VARX(\( p \)) models \( \hat{\beta}_L \) is consistent and asymptotically normally distributed with asymptotic covariance matrix (Hamilton,1994) and (Lütkepohl, 1991).

\[
\text{as} \left[ \hat{\beta}_L - \beta \right] \rightarrow N \left( 0, \Sigma_{\epsilon} \otimes \mathbf{Q}^{-1} \right)
\]

(2.7)

Where, \( \mathbf{X} \mathbf{X}' / m \) converges in probability to \( \mathbf{Q} \), i.e.

\[
\sqrt{m} [ \hat{\beta}_L - \beta ] \rightarrow N \left( 0, \Sigma_{\epsilon} \otimes \mathbf{Q}^{-1} \right)
\]

(2.8)

2.2 Inference on Coefficients

Assume that \( \hat{\beta}_L \) denotes the \( i \)th element of \( \hat{\beta}_L \), it has asymptotically normally distributed with standard error given by the square root of \( i \)th diagonal element of \( \Sigma_{\epsilon} \). Hence, asymptotically valid \( t \)-tests on individual coefficient may be constructed in the usual way. More general linear hypotheses of the form given by

\[ H_0 : C\beta = r \]

(2.9)

involving coefficients across different equations of the VARX(\( p \)) may be tested using the Wald statistic:

\[
Wald = (\hat{C}_L' - r)' C (\hat{\Sigma}_{\epsilon}^{-1} C' \hat{C}_L' - r)
\]

(2.10)

Under the null hypotheses in (2.9), the Wald statistic (2.10) has a limiting \( \chi^2_q \) distribution, where \( q = \text{rank}(C) \) gives the number of linear restrictions.

2.3 Lag Length Selection

In practice, the lag order \( p \) in the VARX(\( p \)) model (2.1) is of course unknown. So, it may be determined using model selection criteria \( IC_p \). The general approach is to fit VARX(\( p \)) models with orders \( p = 1, 2, \ldots, p_{\text{max}} \) and choose the value of \( p \) which minimizes some model selection criteria, \( p = \text{arg}\min(IC_p) \), where the minimization is over \( p = 1, 2, \ldots, p_{\text{max}} \) and \( p_{\text{max}} \) is a given upper bound on the possible lag orders. In the case of co-integration tests, a useful rule of thumb for determining \( p_{\text{max}} \) suggested by Lütkepohl & Saikkonen (1999) which is;

\[ p_{\text{max}} = \text{integer}(T^{1/5}) \]

Model selection criteria for VARX(\( p \)) models generally denoted as

\[
IC_p = Ln[\hat{\Sigma}_e] + \mathbf{C} T \frac{m^2}{m - k}
\]

(2.11)

Where, \( \hat{\Sigma}_e = \frac{\sum_{i=p+1}^{T} \tilde{e}_i \tilde{e}_i'}{m} \) is the residual covariance matrix.
without a degrees of freedom correction from a VARX(p) model, $C_T$ is a sequence indexed by the sample size $n$. The three most common information criteria are the Akaike (AIC), Schwarz-Bayesian (SBC) and Hannan-Quinn. In applying Akaike’s information criterion (AIC) (Aka 1973) let $C_T = 2$, the Hannan-Quinn information criterion (HQIC) (Hannan & Quinn, 1979), $C_T = 2\ln(\lambda) \ln(n)$ and Schwarz’s Bayesian information criterion (SBC) (Schwarz, 1978), $C_T = \ln(n) \ln(p)$.

3 The Forecasting

Forecasting from a VARX(p) model is one of the main objectives of research and the following gives a brief description.

Consider first the problem of forecasting future values of $Y_t$ when the parameters ($\Phi_1, \Phi_2, ..., \Phi_p, \Psi$) of the VARX(p) process are assumed to be known and deterministic terms or exogenous variables with perfect predict. The best linear predictor, in terms of minimum mean squared error ($MSE$), of $Y_{t+1}$ or 1-step forecast based on information available at time $T$ is

$$Y_{T+1} = \Psi Z_{T+1} + \Phi_1 Y_T + \Phi_2 Y_{T-1} + ... + \Phi_p Y_{T-p+1}$$  (3.1)

Forecasts for longer horizons $h$ ($h$-step forecasts) may be obtained using the chain-rule of forecasting as

$$Y_{T+h} = \Psi Z_{T+h} + \Phi_1 Y_{T+h-1} + \Phi_2 Y_{T+h-2} + ... + \Phi_p Y_{T+h-p}$$  (3.2)

Where, $Y_{T+h} = Y_{T+j}$, for $j \leq 0$. The $h$-step forecast errors may be expressed as

$$Y_{T+h} - Y_{T+1} = \sum_{s=0}^{h-1} \Lambda_s e_{T+h-s}$$  (3.3)

Where, the matrices $\Lambda_s$ are determined recursive substitution as shown in equation (2.3). The forecasts are unbiased since all of the forecast errors have expectation zero, and the ($MSE$) matrix for $Y_{T+h} | T$ is

$$\Sigma(h) = MSE[Y_{T+h} - Y_{T+1}] = \sum_{s=0}^{h-1} \Lambda_s \Sigma e N_s$$  (3.4)

Now consider forecasting $Y_{T+h}$ when the parameters of the VARX(p) process are estimated using multivariate least squares. The best linear predictor of $Y_{T+h}$ is now

$$\hat{Y}_{T+h} = \hat{B}_h \hat{X}_{T+h}$$  (3.5)

Where, $\hat{X}_{T+h} = (Z_{T+h}, \hat{Y}_{T+h-1} | T, \hat{Y}_{T+h-2} | T, ..., \hat{Y}_{T+h-p} | T)$. Conditioned on the full set of information and on forecasts of exogenous variables, $Z_{T+h}$. The $h$-step forecast error is:

$$Y_{T+h} - \hat{Y}_{T+h} = \sum_{s=0}^{h-1} \Lambda_s e_{T+h-s} + (B - \hat{B}_h) \hat{X}_{T+h} | T$$  (3.6)

Where, $Y_{T+h} = Y_{T+h}, h = 0, 1, ..., Furthermore, $B_h$ and the term $(B - \hat{B}_h) \hat{X}_{T+h} | T$ captures the part of the forecast error due to estimating the parameters of the VARX(p). The ($MSE$) matrix of the $h$-step forecast is then (Green, 2003).

$$\Sigma(h) = MSE[\hat{Y}_{T+h} - Y_{T+h}] = \sum_{s=0}^{h-1} \Lambda_s \Sigma e N_s$$  (3.7)

Where, $MSE[\hat{Y}_{T+h} - Y_{T+h}]$ denoted as (Johnson & Wichern, 1992)

$$MSE[\hat{Y}_{T+h} - Y_{T+h}] = (X_{T+h} | T, XX)^{-1} (X_{T+h} | T, Y_{T+h} - Y_{T+h})$$ (3.8)

In practice, the second term $MSE[\hat{Y}_{T+h} - Y_{T+h}]$ is often ignored and $\Sigma(h)$ is computed using (3.4) as

$$\hat{\Sigma}(h) = \sum_{s=0}^{h-1} \Lambda_s \Sigma e N_s$$  (3.9)

with $\hat{\Lambda}_s = \sum_{j=1}^{h-1} \hat{\Lambda}_j - \hat{\Phi}_j$. Lütkepohl (1991) gives an approximation to $MSE[\hat{Y}_{T+h} - Y_{T+h}]$ which may be interpreted as a finite sample correction to (3.8).

Asymptotic $(1-\alpha)100\%$ confidence intervals for the individual elements of $Y_{T+h} | T$ are then computed as

$$Y_{L,T+h} = Z_{1-\alpha/2} \Sigma_e \hat{\sigma}_l h < Y_{T+h} < Y_{L,T+h} + Z_{1-\alpha/2} \Sigma_e \hat{\sigma}_l h$$  (3.10)

where $Z_{1-\alpha/2}$ is the $(1-\alpha/2)$ quartile of the standard normal distribution and $\hat{\sigma}_l h$ denotes the square root of the $k^h$ diagonal element of $\Sigma(h)$.

The Gaussianity of the forecast errors $e_{T+h} | T$ can also be used to obtain the following $(1-\alpha)$100% joint ellipsoid for all the variables in the system (Johnson & Wichern, 1992).

$$[Y_{T+h} - Y_{T+h}] \Sigma^{-1} [Y_{T+h} - Y_{T+h}] < \chi^2$$ (3.11)

Where, $\chi^2$ is the $\alpha$-quantile of the $\chi^2$ distribution with $n$ degrees of freedom. Constructing the ellipsoids in (3.11) can be quite demanding when $n$ is larger than two or three.

4 Methodology and Data description

4.1 Empirical VARX Model

The main purpose of this research is to use augmented vector autoregressive VARX model (2.1) for forecasting the areas cultivated with the crops that consume large amounts of water in Saudi Arabia which suffers from a scarcity of water sources. These crops determined in three important crops that are Dates ($Y_1$), Clover ($Y_2$) and Fodder ($Y_3$). Then VARX model was set up for those three series. The applied study depends on the time series data from 1986 to 2013. These data, collected from...
different numbers of the annual statistical book, were published by ministry of agriculture in KSA as a secondary source for data. So the total number of observations can be set \( T = 28 \). According to rule; \( p_{\text{max}} = \text{integer}(T^{1/3}) \), the upper bound of the possible lag orders can be approximated to 3, where, \( (p_{\text{max}} = \text{integer}(T^{1/3}) = \text{integer}(28^{1/3}) = 3) \). The \( p \) lag is 1,2,3, and the empirical VARX(\( p \)) model denoted as

\[
Y_t = \psi Z_t + \sum_{j=1}^{p} \Phi_j Y_{t-j} + \epsilon_t, \quad t = p+1, p+2, \ldots, 28, \quad p = 1,2,3
\]

(4.1)

Where, \( Y_t = \{\text{Dates } (Y_{1t}), \text{Clover } (Y_{2t}), \text{Fodder } (Y_{3t})\} \) is \( 3 \times 1 \) vector of areas cultivated with three crops at time \( t \), \( t = p+1, p+2, \ldots, 28 \), \( Z_t = \{1, \text{trend}, D_t\} \) is \( 3 \times 1 \) vector of deterministic components and \( D_t \) which denotes a dummy variable and it was defined as follows: \( D_t = 0 \) before the year 2007 (before the implementation of resolution No. 335) and \( D_t = 1 \) from year 2007 until the end of the series 2013, \( \psi = \{\psi_0, \psi_1, \psi_2\} \) is \( 3 \times 3 \) coefficients matrix of deterministic and dummy variable, \( \Phi_j, \quad j = 1,2,3, \) \( p \) are \( 3 \times 3 \) matrices of autoregressive coefficients and \( \epsilon_t \) is \( 3 \times 1 \) vector of disturbances.

In this paper unit root test was performed by using augmented Dickey-Fuller (ADF) normalized bias statistics that defined as

\[
\text{ADF}_n = n \tilde{\rho} \left( 1 - \phi_1 - \phi_2 - \ldots - \phi_p \right)
\]

(4.2)

Where, \( \tilde{\rho}, \phi_1, \phi_2, \ldots, \phi_p \) denote the (OLS) estimations for coefficient of the univariate autoregressive model which is defined as

\[
\Delta Y_t = \psi_0 + \psi_1 \text{trend}_t + \psi_2 D_t + \psi_3 Y_{t-1} + \sum_{j=1}^{p} \phi_j \Delta Y_{t-j} + \epsilon_t, \quad p = 1,2,3
\]

(4.3)

Dickey & Fuller (1979) critical values are used to reject or accept the null hypothesis \( H_0: \gamma = 0 (\text{series } Y_t \text{ has unit root}) \), or the time series \( Y_t \) is integrated from degree 1; \( f(1) \) against the alternative \( H_1: \gamma < 0 (\text{series } Y_t \text{ has not unit root}) \) or the time series \( Y_t \) is integrated from degree 0; \( f(0) \). if the value of test statistics (ADF\( \psi \)) is less than the critical value, then \( H_0 \) can be rejected and the null hypothesis can be accepted, this test will be re-run with the first difference \( \Delta Y_t \), in this case if the null hypothesis is rejected, VARX(\( p \)) model may be set up for the three series at the first difference; \( Y_t = \{\Delta Y_{1t}, \Delta Y_{2t}, \Delta Y_{3t}\} \).

Model selection criteria \( I^p \) in (2.11); (AIC, HQIC and SBC) will be applied to select the optimal lag period \( p = 1,2,3 \) for empirical VARX(\( p \)) in (4.1). Then the selected model will be estimated and used to forecast the area cultivated with the three crops under study during the period from 2014 to 2018 as well the amount of water that is expected to be consumed in the case of cultivation of these crops during that period.

### 4.2 Data description

Table 1 shows some descriptive statistics such as mean, standard deviation (St. Dev), coefficient of variation (C.V) \( \min \) \( (\text{Min}) \) and maximum \( (\text{Max}) \) for the areas cultivated with the important agricultural crops; Dates \( (Y_1) \), Clover \( (Y_2) \) and Fodder \( (Y_3) \) in two periods, the first period is from the year 1986 to 2006 before the implementation of resolution No. 335 and the second period from year 2007 to 2013 which it is after implementation of the resolution No. 335.

We note that, the mean of area cultivated with dates in the second period (after the implementation of the resolution) is greater than this mean in first period by 48.5%, the mean of area cultivated with clover in the second period great than this mean in first period by 18.3% and the mean of area cultivated with fodder in

<table>
<thead>
<tr>
<th>Series</th>
<th>Period</th>
<th>( n )</th>
<th>Mean</th>
<th>St. Dev</th>
<th>C.V</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>( Y_1 ) Dates</td>
<td>1st(1986-2006)</td>
<td>21</td>
<td>105776</td>
<td>33599</td>
<td>31.8</td>
<td>63033</td>
<td>152402</td>
</tr>
<tr>
<td></td>
<td>2nd(2007-2013)</td>
<td>7</td>
<td>157096</td>
<td>2267</td>
<td>1.4</td>
<td>155118</td>
<td>161975</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>28</td>
<td>118606</td>
<td>36735</td>
<td>31</td>
<td>63033</td>
<td>161975</td>
</tr>
<tr>
<td>( Y_2 ) Clover</td>
<td>1st(1986-2006)</td>
<td>21</td>
<td>96221</td>
<td>34362</td>
<td>35.7</td>
<td>46118</td>
<td>161140</td>
</tr>
<tr>
<td></td>
<td>2nd(2007-2013)</td>
<td>7</td>
<td>113833</td>
<td>13722</td>
<td>12.1</td>
<td>95589</td>
<td>126611</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>28</td>
<td>100624</td>
<td>31254</td>
<td>31.1</td>
<td>46118</td>
<td>161140</td>
</tr>
<tr>
<td>( Y_3 ) Fodder</td>
<td>1st(1986-2006)</td>
<td>21</td>
<td>319022</td>
<td>115295</td>
<td>36.1</td>
<td>92933</td>
<td>512786</td>
</tr>
<tr>
<td></td>
<td>2nd(2007-2013)</td>
<td>7</td>
<td>238858</td>
<td>23152</td>
<td>9.7</td>
<td>207013</td>
<td>264599</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>28</td>
<td>298981</td>
<td>105902</td>
<td>35.4</td>
<td>92933</td>
<td>512786</td>
</tr>
</tbody>
</table>

Source: Computed from time series data collected from the Annual Statistical Book, Ministry of Agriculture, KSA

http://www.jebas.org
second period less than this mean in first period by 25.1%. The mean of area cultivated with fodder greater than the mean of area cultivated with each of dates and clover. Area data for the three crops in the first period more dispersed from the data of area in the second period, then the second period more stable than the first period.

5 Empirical Strategy and Results Discussion

In this research SAS program version (9.2) was used for obtaining all the results concerning the application of the empirical \( \text{VAR}(p) \) model \((4.1)\) that represents the data of areas cultivated with three crops, the following is a presentation and discussion of these results.

5.1 The unit root test.

Augmented Dickey-Fuller \( (ADF_p) \) statistics test has been computed by using equation \((4.2)\) in order to test the stationarity for both normal time series \( (Y_1, Y_2, \text{ and } Y_3) \) as well as time series at the first difference of series \( (\Delta Y_1, \Delta Y_2, \text{ and } \Delta Y_3) \). Table 2 displays the values of \( (ADF_p) \) and \( \psi \text{value} \) in the case of \( \text{VAR}(p) \) model includes (zero mean, mean and trend).

From the values of \( (ADF_p) \) test statistics and \( \psi \text{value} \) in the table 2, we note that the time series for the normal areas cultivated with the crops under study \( (\text{dates } Y_1, \text{ clover } Y_2, \text{ and fodder } Y_3) \) in Saudi Arabia has a unit root at \( p \text{lag} = 1,2,3 \). This indicates the three normal time series are nonstationary. After taking the first difference, the \( \psi \text{value} \) show that each time series of the three \( (\Delta Y_1, \Delta Y_2, \text{ and } \Delta Y_3) \) does not have a unit root at \( 5\% \) significant level, and it means that the time series of the areas cultivated with this three crops are stationary at the first differences. So the empirical \( \text{VAR}(p) \) model that can be suggested for forecasting the areas cultivated with the three crops must be fit the data of areas at the first difference. Then \( \text{VAR}(p) \) model \((4.1)\) rewritten as

\[
\Delta Y_t = \psi Z_t + \sum_{i=1}^{\psi} \Delta Y_{t-i} + \varepsilon_t, \quad t = p+1, p+2, \ldots, 28, \quad p=1,2,3 \quad (5.1)
\]

Table 2 The results of \( (ADF_p) \) test statistics and \( \psi \text{value} \) for testing unit root in time series (wheat \( Y_1 \), clover \( Y_2 \), and fodder \( Y_3 \))

<table>
<thead>
<tr>
<th>Type of Model</th>
<th>Lag p</th>
<th>Dates</th>
<th>Clover</th>
<th>Fodder</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zero Mean</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>0.614</td>
<td>0.823</td>
<td>0.439</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.583</td>
<td>0.816</td>
<td>0.265</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0.546</td>
<td>0.807</td>
<td>0.231</td>
</tr>
<tr>
<td>Normal series (Y)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Single Mean</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>-1.199</td>
<td>0.856</td>
<td>-3.832</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>-1.215</td>
<td>0.854</td>
<td>-6.897</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>-1.466</td>
<td>0.827</td>
<td>-9.860</td>
</tr>
<tr>
<td>Trend</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>-3.616</td>
<td>0.895</td>
<td>-4.410</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>-3.077</td>
<td>0.923</td>
<td>-9.471</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>-3.501</td>
<td>0.900</td>
<td>-14.834</td>
</tr>
<tr>
<td>First difference</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zero Mean</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>-15.982</td>
<td>0.002</td>
<td>-17.320</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>-8.791</td>
<td>0.030</td>
<td>-15.088</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>-4.896</td>
<td>0.117</td>
<td>-16.678</td>
</tr>
<tr>
<td>Single Mean</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>-30.865</td>
<td>&lt;.0001</td>
<td>-19.581</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>-26.628</td>
<td>&lt;.0001</td>
<td>-18.552</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>-19.949</td>
<td>0.002</td>
<td>-22.851</td>
</tr>
<tr>
<td>Trend</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>-34.432</td>
<td>&lt;.0001</td>
<td>-20.974</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>-35.859</td>
<td>&lt;.0001</td>
<td>-21.370</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>-30.936</td>
<td>&lt;.0001</td>
<td>-34.850</td>
</tr>
</tbody>
</table>

Source: Computed from time series data collected from the Annual Statistical Book, published by the Ministry of Agriculture, KSA
Where, \( \Delta Y_t = (\Delta Y_{1t}, \Delta Y_{2t}, \Delta Y_{3t})' \) is a 3×1 vector of time series of areas at the first differences.

5.2 Information criteria \( IC_p \) for selecting optimal lag period

SAS output for the values of information criteria (\( AIC \), \( HQIC \) and \( SBIC \)) is summarized in Table 3 in order to select the optimal lag period for empirical \( VAR(2) \) in 26.

Table 3 The values of information criteria for empirical Model (5.1)

<table>
<thead>
<tr>
<th>Model Type</th>
<th>( p ) log</th>
<th>( AIC )</th>
<th>( HQIC )</th>
<th>( SBIC )</th>
</tr>
</thead>
<tbody>
<tr>
<td>No intercept</td>
<td>1</td>
<td>58.72</td>
<td>58.89</td>
<td>59.30</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>58.43</td>
<td>58.71</td>
<td>59.45</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>58.54</td>
<td>58.93</td>
<td>60.01</td>
</tr>
<tr>
<td>With intercept</td>
<td>1</td>
<td>58.89</td>
<td>59.10</td>
<td>59.61</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>58.62</td>
<td>58.95</td>
<td>59.79</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>58.75</td>
<td>59.18</td>
<td>60.37</td>
</tr>
<tr>
<td>Linear trend</td>
<td>1</td>
<td>58.46</td>
<td>58.71</td>
<td>59.33</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>58.23</td>
<td>58.59</td>
<td>59.54</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>58.41</td>
<td>58.88</td>
<td>60.18</td>
</tr>
</tbody>
</table>

Source: Computed from time series data collected from the Annual Statistical Book, Ministry of Agriculture, KSA

From Table 3, we note that most of the information criteria have small values in the case of whether the length of lag is equal 2, or the optimal \( log \) \( p \) is \( p = 2 \). So, we can suggest using \( VAR(2) \) model with linear trend to forecast the areas cultivated with crops (dates \( Y_{1t} \), clover \( Y_{2t} \) and fodder \( Y_{3t} \)) during the period (2014-2018) in Saudi Arabia. The suggested empirical model \( VAR(2) \) can be determined by setting \( p = 2 \) and \( Z_t = (1, t, D_t)’ \) into place in model (5.1) and rewritten as

\[
Y_t = \psi_0 + \psi_1 D_t + \psi_2 t + \phi_1 Y_{1t-1} + \phi_2 Y_{2t-1} + \phi_3 Y_{3t-1} + \phi_4 Y_{1t-2} + \phi_5 Y_{2t-2} + \phi_6 Y_{3t-2} + \epsilon_t
\]

Where, the system equations (5.2) contains three equations corresponding to the three crops mentioned above, and each single equation of the three includes 9 parameters.

5.3 Results of Using Wald Test for Some Restrictions

The values of Wald statistics for testing the null hypothesis that the addition of each independent variable to the model (5.2) already containing other variable does not significantly improve the prediction of areas computed by application equation (2.10) and set in Table 4.

From Table 4 we note that the addition of each constant term, linear trend term, dummy variable and the first difference of areas lagged one period to the model (5.2) already containing other variables will significantly improve the prediction of areas, but the addition of the first difference of areas lagged two periods does not significantly improve the prediction of areas. So we can remove the lagged difference \( \Delta Y_{t-2} \) and reduce the empirical \( VAR(2) \) model to become \( VAR(1) \) which written as:

\[
\begin{align*}
\Delta Y_{1t} &= \phi_{11} Y_{1t-1} + \phi_{12} Y_{1t-2} + \epsilon_{1t} \\
\Delta Y_{2t} &= \phi_{21} Y_{1t-1} + \phi_{22} Y_{1t-2} + \epsilon_{2t} \\
\Delta Y_{3t} &= \phi_{31} Y_{1t-1} + \phi_{32} Y_{2t-2} + \epsilon_{3t}
\end{align*}
\]

Table 5 displays the values of Wald test statistics under application reduced empirical model (5.3) as well as the value of \( Wald \) test statistics for testing the appropriateness of empirical model \( VAR(1) \) for forecasting.

From Table 5, we note that the addition of each constant term, linear trend term, dummy variable and the first difference of areas lagged one period to the model (5.3) already containing other variables will significantly improve the prediction of areas at significance level 5%. At the same time the reduced model \( VAR(1) \) is significantly suitable for forecasting the areas cultivated with the three crops.
5.4 The Results of Least Squares Estimates for $VARX(1)$ Model

Least squares estimates for coefficients of three equations, standard errors of these estimates, values of test statistics $I$ and the corresponding $p$-values computed and set in table 6.

From the results (Table 6), we note the following:

At 5% significance level we show that constant has positive significant effect on the first difference of the areas cultivated with dates and lagged one period $(\Delta Y_{1t-1})$ has negative significant effect on the first difference of the areas cultivated with fodder $(\Delta Y_{2t})$ as well as cultivated with clover $(\Delta Y_{2t})$, but at 15% significance level the first difference of the areas cultivated with dates and lagged one period $(\Delta Y_{1t-1})$ has negative significant effect on the first difference of the areas cultivated with dates $(\Delta Y_{1t})$, as well as the first difference of the areas cultivated with dates and lagged one period $(\Delta Y_{1t-1})$ has negative significant effect on the first difference of the areas cultivated with clover $(\Delta Y_{2t})$. 

Table 5 The values of Wald test statistics under application reduced empirical model $VARX(1)$

<table>
<thead>
<tr>
<th>Variable</th>
<th>Null Hypothesis</th>
<th>Wald</th>
<th>df</th>
<th>Pr &gt; Wald</th>
</tr>
</thead>
<tbody>
<tr>
<td>Const.</td>
<td>$H_0 : \psi_{10} = \psi_{20} = \psi_{30} = 0$</td>
<td>17.52</td>
<td>2</td>
<td>0.0002</td>
</tr>
<tr>
<td>Trend ($t$)</td>
<td>$H_0 : \psi_{11} = \psi_{21} = \psi_{31} = 0$</td>
<td>15.69</td>
<td>2</td>
<td>0.0004</td>
</tr>
<tr>
<td>Dummy ($D_t$)</td>
<td>$H_0 : \psi_{12} = \psi_{22} = \psi_{32} = 0$</td>
<td>15.82</td>
<td>2</td>
<td>0.0004</td>
</tr>
<tr>
<td>Log $(\Delta Y_{t-1})$</td>
<td>$H_0 : (\phi_1 \phi_2 \phi_3) Y = (\phi_1 \phi_2 \phi_3) = (\phi_1 \phi_2 \phi_3) Y = 0$</td>
<td>12.94</td>
<td>6</td>
<td>0.0441</td>
</tr>
<tr>
<td>Model</td>
<td>$H_0 : The VARX(1) model is not suitable to predict $Y_t$.</td>
<td>27.97</td>
<td>12</td>
<td>0.0056</td>
</tr>
</tbody>
</table>

Table 6 Results of the (LS) estimates for coefficients of $VARX(1)$ model

| Eq. | Coefficients $\beta$ | $\beta$ | $SE_\beta$ | $t$ | Pr $> |t|$ | Variable |
|-----|----------------------|---------|------------|-----|-----------|----------|
| $\Delta Y_{1t}$ Dates | $\beta_{10} = \psi_{10}$ | 9119.2 | 8209.1 | 1.11 | 0.280 | 1 |
| | $\beta_{11} = \psi_{11}$ | -346.8 | 330.4 | -1.05 | 0.306 | $t$ |
| | $\beta_{12} = \psi_{12}$ | 885.1 | 5455.9 | 0.16 | 0.873 | $D_t$ |
| | $\beta_{13} = \phi_{13}$ | -0.3742 | 0.2395 | -1.56 | 0.134 | $\Delta Y_{1t-1}$ |
| | $\beta_{14} = \phi_{14}$ | 0.1162 | 0.0978 | 1.19 | 0.249 | $\Delta Y_{2t-1}$ |
| | $\beta_{15} = \phi_{15}$ | -0.0805 | 0.0321 | -2.51 | 0.021 | $\Delta Y_{3t-1}$ |
| $\Delta Y_{2t}$ Clover | $\beta_{16} = \psi_{20}$ | 61234 | 18424 | 3.32 | 0.003 | 1 |
| | $\beta_{21} = \psi_{21}$ | -2264 | 741.5 | -3.05 | 0.006 | $t$ |
| | $\beta_{22} = \psi_{22}$ | -33180 | 12245 | -2.71 | 0.014 | $D_t$ |
| | $\beta_{23} = \phi_{23}$ | -0.3895 | 0.5376 | -0.72 | 0.477 | $\Delta Y_{1t-1}$ |
| | $\beta_{24} = \phi_{24}$ | -0.3314 | 0.2194 | -1.51 | 0.147 | $\Delta Y_{2t-1}$ |
| | $\beta_{25} = \phi_{25}$ | -0.1504 | 0.0700 | -2.15 | 0.044 | $\Delta Y_{3t-1}$ |
| $\Delta Y_{3t}$ Fodder | $\beta_{26} = \psi_{30}$ | 175644 | 57240 | 3.07 | 0.006 | 1 |
| | $\beta_{31} = \psi_{31}$ | -6997.1 | 2304 | -3.04 | 0.007 | $t$ |
| | $\beta_{32} = \psi_{32}$ | -98507 | 38043 | -2.59 | 0.018 | $D_t$ |
| | $\beta_{33} = \phi_{33}$ | -0.3852 | 1.6702 | -0.23 | 0.820 | $\Delta Y_{1t-1}$ |
| | $\beta_{34} = \phi_{34}$ | -0.4953 | 0.6818 | -0.73 | 0.476 | $\Delta Y_{2t-1}$ |
| | $\beta_{35} = \phi_{35}$ | -0.0822 | 0.2236 | -0.37 | 0.717 | $\Delta Y_{3t-2}$ |
From the values of (OLS) estimates for coefficients, the VARX(1) model (5.3) can be written as:

\[
\begin{bmatrix}
\Delta Y_{1t} \\
\Delta Y_{2t} \\
\Delta Y_{3t} \\
\Delta Y_{4t}
\end{bmatrix} = \begin{bmatrix}
9119.2 & -346.8 & 885.1 & 1 \\
61234 & -2264 & -33180 & 1 \\
175644 & -6997.1 & -9807 & 1 \\
\psi Z_t & \Phi Y_{t-1}
\end{bmatrix} + \begin{bmatrix}
-0.3742 & 0.1162 & -0.0805 & 1 \\
-0.3895 & -0.3314 & -0.1504 & 1 \\
-0.3852 & -0.4953 & -0.0822 & 1 \\
\end{bmatrix}\Delta Y_{1t-1} + \begin{bmatrix}
\Delta Y_{2t-1} \\
\Delta Y_{3t-1}
\end{bmatrix}
\]

(5.4)

5.5 Discussion of the Forecasting Results

The VARX(1) model (5.4) was used for forecasting the values of areas cultivated with the three crops (dates, clover, and fodder) by application equation (3.5) as well as 95% confidence interval by application equation (3.10) in the period (2014-2018). The results of prediction summarized in Table 7.

Table 7 The predicted values of area in period (2014, 2018) by using model (29)

<table>
<thead>
<tr>
<th>Crops</th>
<th>Years</th>
<th>Forecast</th>
<th>95% Confidence Limits</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Lower</td>
</tr>
<tr>
<td>Dates</td>
<td>2014</td>
<td>156691</td>
<td>143718</td>
</tr>
<tr>
<td></td>
<td>2015</td>
<td>157253</td>
<td>138436</td>
</tr>
<tr>
<td></td>
<td>2016</td>
<td>157638</td>
<td>134602</td>
</tr>
<tr>
<td></td>
<td>2017</td>
<td>158087</td>
<td>131416</td>
</tr>
<tr>
<td></td>
<td>2018</td>
<td>158542</td>
<td>128705</td>
</tr>
<tr>
<td>Grow rat r</td>
<td>2014</td>
<td>126287</td>
<td>97170</td>
</tr>
<tr>
<td></td>
<td>2015</td>
<td>126394</td>
<td>90324</td>
</tr>
<tr>
<td></td>
<td>2016</td>
<td>124558</td>
<td>81421</td>
</tr>
<tr>
<td></td>
<td>2017</td>
<td>122237</td>
<td>73668</td>
</tr>
<tr>
<td></td>
<td>2018</td>
<td>118598</td>
<td>64998</td>
</tr>
<tr>
<td>Grow rat r</td>
<td>2014</td>
<td>250786</td>
<td>160325</td>
</tr>
<tr>
<td></td>
<td>2015</td>
<td>231886</td>
<td>107108</td>
</tr>
<tr>
<td></td>
<td>2016</td>
<td>205897</td>
<td>49818</td>
</tr>
<tr>
<td></td>
<td>2017</td>
<td>174523</td>
<td>5294</td>
</tr>
<tr>
<td></td>
<td>2018</td>
<td>136810</td>
<td>-65055</td>
</tr>
</tbody>
</table>

From Table 8 we note that through the forecasting period (2014, 2018), it is expected that the mean of water consumption of dates 2837.6 million m$^3$ annually, the mean of water consumption of clover 3708.4 million m$^3$ annually and the mean of water consumption of fodder 3999.6 million m$^3$ annually.

Conclusions and the Recommendations

From the previous results we recommend the following:

i. The augmented vector autoregressive (VARX) model can be used in the analysis of multivariate time series when describing the dynamic behavior of economic and financial time series, so we recommend using this model for forecasting because it provides superior forecasts to those from univariate time series models and elaborate theory-based simultaneous equations models.

ii. The expected areas cultivated both the alfalfa and fodder represent large proportion during the forecast period, and then large amounts of water consumption although Saudi
Arabia suffer from water scarcity, so we recommend those interested in agricultural policy direction to reduce the areas of these two crops or curtailed and encourage the import or cultivation outside as external investment.

iii. From results we note that the area cultivated with dates stability during the forecast period, should be encouraged to export to achieve an increase in national income.

References


