FORECASTING EXPORTS OF INDUSTRIAL GOODS FROM PUNJAB - AN APPLICATION OF UNIVARIATE ARIMA MODEL

GULSHAN KUMAR, SANJEEV GUPTA *

ABSTRACT: The present study is an attempt to build a Univariate time series model to forecast the exports of industrial goods from Punjab for ensuing decade till 2020. The study employs Box-Jenkin’s methodology of building ARIMA (Autoregressive Integrated Moving Average) model to achieve various objectives of study. Annual time series data for exports of industrial products have been culled from Directorate of Industries, Punjab for the period 1974-75 to 2007-08. Different selected models were tested by various diagnostic tests to ensure the accuracy of obtained results. The results revealed that during the days to come, exports of industrial products from Punjab are going to experience a sharp decline in growth as compared to past three decades in which growth maintained two digit level. In light of the forecasts, concerted efforts on the part of Government, entrepreneurs, industrialists, farmers and producers are the need of the hour to establish a healthy state economy and its export sector.

KEY WORDS: ARIMA; Forecasting; Box-Jenkin Method; Akaike Information Criteria; Schwarz Bayesian Information Criteria

JEL CLASSIFICATION: C53

1. INTRODUCTION

The development literature abounds with the models of export-led growth, and it is claimed both historically and in the contemporary world economy that growth of exports can be taken as engine of growth (Thirlwall, 1994, p.364). In the present era of globalisation and liberalised policy regime, export promotion strategies occupy the central stage. Although, the trade policy is the jurisdiction of central government, the role of state governments is no less significant because export production is a local activity which is highly influenced by the activities and policies of state government.

* Senior Lecturer, Ph.D., D.A.V. College, Hoshiarpur, Punjab, India, jsgsass40@rediffmail.com
Ph.D., Dr BR Ambedkar National Institute of Technology, Jalandhar, India, guptas@nitj.ac.in
Thus, concerted and planned efforts are required not only at country level as a whole but also at State levels to push up the growth of exports (Nanda & Raikhy, 2000, p.266). Punjab is performing well (even more than all India growth of exports) as export of industrial goods is concerned, still potential is there to perform better (Kumar & Gupta, 2007, p.77). The export of Industrial products from Punjab was amounted just Rs. 62.95 crore in 1974-75 but soared high to Rs. 11267.04 crore by 2007-08 (see Table 1).

Table 1. Growth of Exports

<table>
<thead>
<tr>
<th>Year</th>
<th>Exports (Rs. Crore)</th>
<th>Year</th>
<th>Exports (Rs. Crore)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1974-75</td>
<td>62.95</td>
<td>91-92</td>
<td>900.81</td>
</tr>
<tr>
<td>75-76</td>
<td>75.63</td>
<td>92-93</td>
<td>1214.77</td>
</tr>
<tr>
<td>76-77</td>
<td>90.13</td>
<td>93-94</td>
<td>1815.47</td>
</tr>
<tr>
<td>77-78</td>
<td>93.85</td>
<td>94-95</td>
<td>2082.3</td>
</tr>
<tr>
<td>78-79</td>
<td>106.2</td>
<td>95-96</td>
<td>2564.61</td>
</tr>
<tr>
<td>79-80</td>
<td>125.45</td>
<td>96-97</td>
<td>3641.01</td>
</tr>
<tr>
<td>80-81</td>
<td>162.13</td>
<td>97-98</td>
<td>4204.78</td>
</tr>
<tr>
<td>81-82</td>
<td>234.76</td>
<td>98-99</td>
<td>3629.13</td>
</tr>
<tr>
<td>82-83</td>
<td>228.71</td>
<td>99-2000</td>
<td>4062.62</td>
</tr>
<tr>
<td>83-84</td>
<td>197.19</td>
<td>2000-01</td>
<td>4014.96</td>
</tr>
<tr>
<td>84-85</td>
<td>203.57</td>
<td>2001-02</td>
<td>4407.9</td>
</tr>
<tr>
<td>85-86</td>
<td>245.2</td>
<td>2002-03</td>
<td>7013.51</td>
</tr>
<tr>
<td>86-87</td>
<td>274.83</td>
<td>2003-04</td>
<td>8933.31</td>
</tr>
<tr>
<td>87-88</td>
<td>341.66</td>
<td>2004-05</td>
<td>7914.35</td>
</tr>
<tr>
<td>88-89</td>
<td>466</td>
<td>2005-06</td>
<td>9655.91</td>
</tr>
<tr>
<td>89-90</td>
<td>647.65</td>
<td>2006-07</td>
<td>11797.68</td>
</tr>
<tr>
<td>90-91</td>
<td>769.2</td>
<td>2007-08</td>
<td>11267.04</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Compond Annual Growth Rates</th>
</tr>
</thead>
<tbody>
<tr>
<td>1974-75 to 91-92</td>
</tr>
<tr>
<td>1991-92 to 2007-08</td>
</tr>
<tr>
<td>1974-75 to 2007-08</td>
</tr>
</tbody>
</table>

Source: 1. Directorate of Industries, Punjab. 2. Author’s calculations based on the data supplied by Directorate of Industries, Punjab.

Note: * Significant at 5 percent level of significance.

The growth of exports is also very impressive as 18.72 percent during 1974-2008 as exhibited in this table. Even then, future is uncertain and nothing can be stated confidently about the ensuing years. Keeping in view all this, there arises an urgent need to make concerted efforts to unveil the future and take timely decisions. Hence, the present study is an attempt to build a Univariate time series model to forecast the exports of industrial products from Punjab for the ensuing decade i.e. till 2020, by applying Box-Jenkin’s methodology of building ARIMA (Autoregressive Integrated Moving Average) model.
2. OBJECTIVES OF THE STUDY

Present study has been conducted keeping in mind the following objectives:
1. To generate forecasts of exports of industrial goods from Punjab
2. To suggest suitable forecasting ARIMA model for generation of forecasts.
3. To examine the growth of exports from Punjab.

3. DATA BASE AND ANALYTICAL FRAMEWORK

Present study is based on secondary data for the period 1974-75 to 2007-08. The aggregate data relating to the exports of industrial goods from Punjab (A State in India) were culled from Directorate of Industries, Punjab (A govt. department engaged in collection of data). The forecasts of the above mentioned variable for a lead time of 12 years were generated by applying Box-Jenkins’ ARIMA method with the help of statistical software SPSS (version 7.5). One of the advantages of Box-Jenkins over other forecasting models is that this modelling is not based on economic theory and capable of capturing slightest variation in the data. Box-Jenkins methodology rests on the simplifying assumption that the process which has generated a single time series, is the stationary process but unfortunately most time series encountered are rarely stationary, still it is possible to transform them to stationary by the appropriate level of differencing (maximum up to second level). The degree of differencing transforms a non-stationary series into a stationary one. If non-stationary is added to a mixed ARIMA model, then the general ARIMA (p, d, q) is obtained, it has the form as under:

\[ \Phi_P(B) (1-B)^d Y_t = C + \theta_q (B) \epsilon_t \] \hspace{1cm} (1.1)

or

\[ \Phi_P(B) W_t = C + \theta_q (B) \epsilon_t \] \hspace{1cm} (1.2)

which will be non-stationary unless d=0.

The model is said to be of the order (p, d, q), where p, d and q are usually 0, 1 or 2 (Makridakis, 1998, p.345). Having tentatively identified one or more models that seem likely to provide parsimonious and statistically adequate representation of available data, the next step is to estimate the values of the parameters. Sum of squares of the residuals were computed by using maximum likelihood estimation method given the respective initial estimates of the parameters, optimum values of the parameters were searched by improving the initial estimates iteratively by supplementing them with the information contained in the time series. For a given model involving k parameters, the iterative procedure was continued till the difference between successive values of sum of squared residuals became so small that could be ignored for practical considerations (Box, Jenkins & Reinsll, 1994, p.225).

In order to make an assessment of the validity of the estimated models for the given time series, following certain diagnostic measures were worked out:

(a) Autocorrelations and partial autocorrelation of residuals: The autocorrelation coefficient was worked out by applying formula:
\[ r_k(e) = \frac{\sum_{t=1}^{n-k} e_t e_{tk}}{\sum_{t=1}^{n} e_t^2}; k=1,2,\ldots,n \]

(2.1)

The major concern of ACF of residuals was that whether the residuals were systematically distributed across the series or they contain some serial dependency. Acceptance of the hypotheses of serial dependency concludes that the estimated ARIMA model is inadequate.

And partial autocorrelation were worked out as under:

The sample partial autocorrelation coefficients \( \Phi_{11}, \Phi_{22}, \ldots, \Phi_{kk} \) were estimated with the help of following equation:

\[
\Lambda \\
\Phi_{11} = r_1 \\
k-1 \Lambda \\
r_k - \sum_{j=1}^{k-1} \Phi_{k-1,j} r_{k-j} \\
\Lambda \\
\Phi_{kk} = \frac{1 - \sum_{j=1}^{k-1} \Phi_{k-1,j} r_j}{k-1} \Lambda \quad (k = 2,3,\ldots) \\
\Lambda \\
\Phi_{kj} = \Phi_{k-1,j} - \Phi_{kk} \Phi_{k-1,k-j} \\
(k = 3,4,\ldots; j = 1,2,\ldots, k-1) \\
\Lambda \\
\text{Again, the } \Phi_{kk} \text{ verses the lag } k \text{ could be displayed graphically so as to get sample partial auto correlation coefficient (also known as partial autocorrelogram) (0’ Donovan, (1983), p.82: Gupta, 2003, p.44) to compute 95 percent confidence interval for partial autocorrelogram.}

(b) Portmanteau Test: Ljung-Box Q statistics was computed from the model’s residuals by using

\[
Q = n(n+2) \sum_{k=1}^{\ell} r_k(e)^2 (n-k)^{-1} 
\]

(3)

Non-significance of portmanteau test was taken to imply the generated residuals could be considered a white noise, thereby indicating the adequacy of estimated model.
(c) **Sum of Squares of Error (SSE):** Sum of squares of the errors of fitted models was computed. We selected that model adequate, in case of which sum of square of errors was minimum.

(d) **Akaike Information Criteria (AIC):** AIC was computed to determine both how well the model fits the observed series, and the number of parameter used in the fit. We compared the value AIC with other fitted model to the same data set and we selected that fitted model adequate in case of which AIC was minimum. The AIC is computed as under:

\[
AIC = n \log (SSE) + 2k
\]  

where

\[
k = \text{Number of parameters that are fitted in the model}
\]

\[
\log = \text{Natural logarithm}
\]

\[
n = \text{Number of observations in the series}
\]

\[
SSE = \text{Sum of Squared Errors}
\]

(e) **Schwarz Bayesian Information Criteria (SBC):** SBC is a modification to AIC; it is based on Bayesian consideration. Like AIC it was computed to determine how well the model fits amongst the competing models, and we selected that model adequate in case of SBC was minimum. The SBC is as under:

\[
SBC = n \log (SSE) + k \log (n)
\]

(f) **Karl Pearson’s correlation coefficient:** The bivariate correlation tells the degree of correlation between the trend values generated by various fitted ARIMA models with the original data for the variable and the significance of correlation coefficients is tested by applying t-test.

**Forecasting:** For making forecasts equation (1.2) was unscrambled to express \( Y_t \) and \( e_t \) by using the relation \( W_t = (1-B)^d Y_t \). Given the data up to time \( t \) the optimal forecasts of \( Y_t + \ell \) [designated by \( Y_t (\ell) \)] made a time \( t \) was taken as conditional expectation of \( Y_t + \ell \), where \( t \) is the forecast origin and \( \ell \) is the forecast lead-time. Error term \( e_t \) completely disappeared once we made forecasts more than \( q \) period ahead. Thus for \( \ell > q \), then \( \ell \) period ahead forecast was made as under:

\[
Y_t + \ell = C + \Phi_1 Y_{t-1} + \ldots + \Phi_p Y_{t-p}
\]

(Gupta, 2003, p.47-49; Bashier & Bataineh, 2007, p.2)

4. RESULTS AND DISCUSSION

The results have been discussed in brief under the following sub-heads:

4.1. Stationarity

The application of Box-Jenkins methodology in building an ARIMA model requires that the series is stationary. Therefore, the process starts with testing the series
for stationarity using the plot diagrams, correlogram etc. In order to confirm the mean stationarity and to calculate appropriate level of differencing, the graphical presentation of original series, test of correlogram and Ljung Box Q-statistics were exercised (figures and results for the original series are not shown here for the cause of simplicity and briefness). All the tests confirmed that after the first differencing, (d=1) the variable achieved stationarity.

4.2. Model Identification

Present study has adopted all possible eight basic ARIMA models depending on the values of p,d,q as p and q can adopt any value out of 0,1,2. Models are: \{(1,d,0), (2,d,0), (0,d,1), (1,d,1), (2,d,2), (0,d,2), (1,d,2), (2,d,1)\}. Here, for all the eight models the value of ‘d’ has remained 1.

Table 2. Initial estimate of the Parameters

<table>
<thead>
<tr>
<th>Model</th>
<th>Parameters</th>
<th>Exports</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARIMA (1,d,0)</td>
<td>C</td>
<td>391.3845</td>
</tr>
<tr>
<td></td>
<td>AR1</td>
<td>0.2250</td>
</tr>
<tr>
<td>ARIMA (0,d,1)</td>
<td>C</td>
<td>381.4163</td>
</tr>
<tr>
<td></td>
<td>MA1</td>
<td>-0.5920</td>
</tr>
<tr>
<td>ARIMA (1,d,1)</td>
<td>C</td>
<td>441.7689</td>
</tr>
<tr>
<td></td>
<td>AR1</td>
<td>0.8988</td>
</tr>
<tr>
<td></td>
<td>MA1</td>
<td>0.7479</td>
</tr>
<tr>
<td>ARIMA (2,d,2)</td>
<td>C</td>
<td>404.3551</td>
</tr>
<tr>
<td></td>
<td>AR1</td>
<td>0.4725</td>
</tr>
<tr>
<td></td>
<td>AR2</td>
<td>0.4053</td>
</tr>
<tr>
<td></td>
<td>MA1</td>
<td>0.0273</td>
</tr>
<tr>
<td></td>
<td>MA2</td>
<td>0.6825</td>
</tr>
<tr>
<td>ARIMA (0,d,2)</td>
<td>C</td>
<td>364.5756</td>
</tr>
<tr>
<td></td>
<td>MA1</td>
<td>-0.4288</td>
</tr>
<tr>
<td></td>
<td>MA2</td>
<td>0.2541</td>
</tr>
<tr>
<td>ARIMA (1,d,2)</td>
<td>C</td>
<td>366.4387</td>
</tr>
<tr>
<td></td>
<td>AR1</td>
<td>-0.1819</td>
</tr>
<tr>
<td></td>
<td>MA1</td>
<td>-0.5786</td>
</tr>
<tr>
<td></td>
<td>MA2</td>
<td>0.1732</td>
</tr>
<tr>
<td>ARIMA (2,d,0)</td>
<td>C</td>
<td>370.1460</td>
</tr>
<tr>
<td></td>
<td>AR1</td>
<td>0.2409</td>
</tr>
<tr>
<td></td>
<td>AR2</td>
<td>-0.2480</td>
</tr>
<tr>
<td>ARIMA (2,d,1)</td>
<td>C</td>
<td>351.4024</td>
</tr>
<tr>
<td></td>
<td>AR1</td>
<td>-0.3181</td>
</tr>
<tr>
<td></td>
<td>AR2</td>
<td>-0.3746</td>
</tr>
<tr>
<td></td>
<td>MA1</td>
<td>-0.6787</td>
</tr>
</tbody>
</table>

Source: Author’s calculations based on the data supplied by Directorate of Industries, Punjab.
Note: In all cases, d=1

Estimation of different Ordered ARIMA models: As discussed earlier, in order to make choice for suitable forecasting models, ARIMA process of the order (1,2,0), (2,2,0), (0,2,1), (1,2,1), (2,2,2), (0,2,2), (1,2,2), (2,2,1) were estimated. For
estimating parameters of selected models, we have started with some initial values of C, AR1, AR2, MA1, MA2 for different ordered models as exhibited in Table 2.

4.3 Diagnostic Checking

Once the ARIMA models are identified, then the tests of suitability of the selected model are used and accordingly the analysis of each model is carried out. Table 3 shows comparative results from various models. A model is selected on the basis if it possesses minimum sum of squares of residuals, minimum value of standard error, minimum AIC value, minimum value of SBC, and minimum value of non-significant Box-Ljung Q statistics and high correlation coefficient. Alternative models for each variable were examined comparing the values of these parameters. Only that model has been selected which satisfied maximum number of above mentioned criterion.

Values of the above mentioned criteria computed from the different ordered ARIMA models have been presented in Table 3. Almost in all the cases for different order ARIMA models, correlogram of residuals showed no serial dependency (All correlogram for residuals are not shown here as the number of figures was large, only correlogram for the optimal model has been presented in Fig. 1 and Fig. 2).

Table 3. Comparative Results from Various Models

<table>
<thead>
<tr>
<th>Variable</th>
<th>Estimate</th>
<th>ARIMA (0,d,1)</th>
<th>ARIMA (0,d,2)</th>
<th>ARIMA (1,d,1)</th>
<th>ARIMA (1,d,2)</th>
<th>ARIMA (2,d,0)</th>
<th>ARIMA (2,d,1)</th>
<th>ARIMA (2,d,2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exports</td>
<td>Sum of Squares</td>
<td>16986009</td>
<td>15498725</td>
<td>16564090</td>
<td>14359045.5</td>
<td>14318561</td>
<td>14265615.8</td>
<td>14265615.8</td>
</tr>
<tr>
<td></td>
<td>Standard error</td>
<td>764.68545</td>
<td>725.9833</td>
<td>766.18721</td>
<td>736.68404</td>
<td>708.96391</td>
<td>720.66766</td>
<td>757.56583</td>
</tr>
<tr>
<td></td>
<td>SBC</td>
<td>504.5376</td>
<td>501.67529</td>
<td>507.26075</td>
<td>509.93635</td>
<td>502.71932</td>
<td>506.14876</td>
<td>506.49239</td>
</tr>
<tr>
<td></td>
<td>r</td>
<td>0.858</td>
<td>0.85</td>
<td>0.894</td>
<td>0.894</td>
<td>0.847</td>
<td>0.847</td>
<td>0.851</td>
</tr>
</tbody>
</table>

Note: In all cases, $d=1$

Source: Author's calculations based on the data supplied by Directorate of Industries, Punjab.
Figure 1. Exports (ACF), (Q = 7.564)

Figure 2. Exports (PACF)

Table 3 depicts the values of all the parameters for the variable - exports. Examination of Table 3 has revealed the basis for optimum model (based on satisfaction of maximum number of criterion by a particular model). It is observed that the model (2,d,1) is optimal for the variable exports. This model contains least SSE (13400204.3), minimum standard error (694.49034), minimum AIC (498.43423),
lowest Q (7.564). Moreover r is very high in this case although not highest. Moreover Fig.1 clearly indicates that autocorrelation coefficients individually as well as a group showed no pattern and were lying within lower and upper limits for residuals indicating insignificance of residuals. Same is the case with partial autocorrelation coefficients shown in Fig.2. So this model satisfied maximum no. of criterion hence can be treated optimal.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Optimum Model</th>
<th>Exports</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ARIMA(2,d,1)</td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>351.40239</td>
<td></td>
</tr>
<tr>
<td>AR1</td>
<td>-0.31808</td>
<td></td>
</tr>
<tr>
<td>AR2</td>
<td>-0.37458</td>
<td></td>
</tr>
<tr>
<td>MA1</td>
<td>-0.67872</td>
<td></td>
</tr>
<tr>
<td>MA2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AIC</td>
<td>498.43423</td>
<td></td>
</tr>
<tr>
<td>SBC</td>
<td>504.17018</td>
<td></td>
</tr>
<tr>
<td>Q</td>
<td>7.564</td>
<td></td>
</tr>
<tr>
<td>r</td>
<td>0.845</td>
<td></td>
</tr>
</tbody>
</table>

Note: Here, d=1
Source: Author's calculations based on the data supplied by Directorate of Industries, Punjab.

Perusal of table 4 clearly indicates that ARIMA model (2,d,1) which keeps constant c (351.40239), autoregressive AR1 and AR2 as -0.31808 and -0.37458 respectively. Moving average MA1 (-0.67872) and AIC equal to 498.43423 and SBC being 504.17018, is the optimal one because of satisfying maximum no. of criterion.

5. FORECASTS

Future is highly uncertain but most people view the future as consisting of a large number of alternatives. Future research or forecasting is the best way of examining the different alternatives, identifying the most probable ones and thus reducing the uncertainty to the least. Forecasting is the best designed tool to help decision making and planning in the present (Walonick; 1993). After extracting the optimum model for generation of forecasts, the next step is to prepare forecasts of exports of industrial goods from Punjab. Table 5 keeps all the details of forecasted figures.

Perusal of Table 5 revealed that in the year 2008-09, the forecasted figure of exports is 11228.36 Rs. crore expected to rise to Rs. crore 12567.87 in 2011-12 and to Rs. crore 14032.62 in 2015-16 and finally expected to be Rs. crore 15448.53 by the year 2019-20. As far as compound annual growth rate (CAGR) is concerned, it is expected to be just 2.73. As this growth figure is much less than the growth of exports depicted in table 1 for the previous decades. Situation is quite alarming and timely efforts are need of the hour.
Table 5. Forecasts on the basis of Optimum Model

<table>
<thead>
<tr>
<th>Year</th>
<th>Exports(Rs. Crore)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2008-09</td>
<td>11228.36</td>
</tr>
<tr>
<td>2009-10</td>
<td>12034.24</td>
</tr>
<tr>
<td>2010-11</td>
<td>12387.20</td>
</tr>
<tr>
<td>2011-12</td>
<td>12567.87</td>
</tr>
<tr>
<td>2012-13</td>
<td>12972.99</td>
</tr>
<tr>
<td>2013-14</td>
<td>13371.26</td>
</tr>
<tr>
<td>2014-15</td>
<td>13687.63</td>
</tr>
<tr>
<td>2015-16</td>
<td>14032.62</td>
</tr>
<tr>
<td>2016-17</td>
<td>14399.19</td>
</tr>
<tr>
<td>2017-18</td>
<td>14748.17</td>
</tr>
<tr>
<td>2018-19</td>
<td>15094.66</td>
</tr>
<tr>
<td>2019-20</td>
<td>15448.53</td>
</tr>
<tr>
<td><strong>CAGR:</strong></td>
<td><strong>2.73</strong>*</td>
</tr>
</tbody>
</table>

Source: Author’s calculations based on the data supplied by Directorate of Industries, Punjab.
Note: * Significant at 5 percent level of significance

6. CONCLUSION

Punjab is basically an agricultural state but it keeps proud place in industrial map of India especially the small scale industrial sector. The Auto Regressive Integrated Moving Average (ARIMA) model through Box-Jenkins approach has been used to generate forecasts regarding export of industrial goods from Punjab. The forecasts have depicted not a very bright picture ahead. These forecasts can provide Government and policy makers a direction to design policies accordingly to push up growth in this sector. The government (both state and union) must design supportive industrial policies for export oriented units in state. The emphasis should be given on R&D and quality improvement. The existing export oriented units should be strengthened. Block level export training and information centres should be established. Concerted efforts on the part of Government, entrepreneurs, industrialists, farmers and producers are the need of the hour to establish a healthy state economy and its export sector.

REFERENCES: