Jour. Korean For. Soc. Vol. 95, No. 5, pp. 601~604 (2006)

JOURNAL OF KOREAN
FOREST SOCIETY

Development of the Lumber Demand Prediction Model

Dong-Jun Kim*

Assistant Professor, School of Forest Resources, Chungbuk National University, Cheongju 361-763, Korea

Abstract : This study compared the accuracy of partial multivariate and vector autoregressive models for lumber demand prediction in Korea. The partial multivariate model has three explanatory variables; own price, construction permit area and dummy. The dummy variable reflected the boom of lumber demand in 1988, and the abrupt decrease in 1998. The VAR model consists of two endogenous variables, lumber demand and construction permit area with one lag. On the other hand, the prediction accuracy was estimated by Root Mean Squared Error. The results showed that the estimation by partial multivariate and vector autoregressive model showed similar explanatory power, and the prediction accuracy was similar in the case of using partial multivariate and vector autoregressive model.

Key words : *lumber demand, partial multivariate model, vector autoregressive model, prediction accuracy*

Introduction

Lumber is mainly used as raw material of construction. Korean lumber consumption has steadily increased since 1975 except 1988 and 1998. The boom of lumber consumption in 1988 was from special demand due to Seoul Olympic, and lumber consumption decreased abruptly because of economic recession in 1998. If dividing lumber into softwood and hardwood by species, softwood lumber consumption increased and hardwood lumber consumption decreased. Most of lumber consumed is made of softwood. Korea depends most of softwood lumber consumption on domestic production.

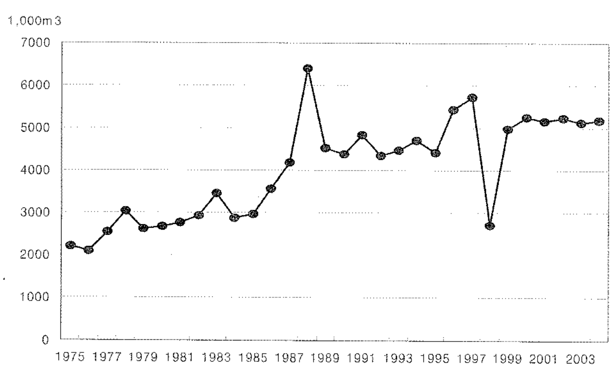

Figure 1. Lumber consumption trend.

*Corresponding author
E-mail: kdj63@chungbuk.ac.kr

This work was supported by chungbuk National University Grant in 2005.

On the other hand, most of hardwood lumber consumed are from foreign countries. The main import country of hardwood lumber is Indonesia. The import from Indonesia is about twenty five percent of total lumber import.

There are a few studies on lumber demand modeling using Korean data; Seok (1992) and Joo (1997) in the 1990's. No study was ever done using vector autoregressive model.

This article compared the prediction accuracy of partial multivariate and vector autoregressive model to develop the most efficient demand prediction model of lumber in Korea. The first objective of this article is to estimate lumber demand using partial multivariate and vector autoregressive model. And the second objective is to compare the prediction accuracy of partial multivariate and vector autoregressive model.

Research Methods

1. Data collection

Lumber consumption was calculated through the production plus import minus export of lumber. Lumber price and construction permit area were collected as variables influencing lumber demand.

Data for the period of thirty years from 1975 to 2004 were used. The data used are at the yearly level. This produced thirty observations.

2. Model specification and estimation

(1) Partial multivariate model

Since lumber demand is determined to a significant degree by final product demand, one should probably include variable describing the final product demand. Simplifying the model, we can derive the demand function for lumber.

$$QD_t = f(P_t, DX_t) \quad (1)$$

where QD_t is lumber demand, P_t is lumber price, and DX_t is the domestic usage of final product.

Consequently, we can express the empirical equation of Korean lumber demand as

$$QD_t = a_0 + a_1P_t + a_2CPA_t + a_3D + \varepsilon_t \quad (2)$$

where QD_t is lumber quantity, P_t is lumber price, CPA_t is construction permit area, D is dummy variable, and ε_t is error term.

Construction permit area was used as a rough proxy for the domestic usage of final product. Dummy variable was used to solve the problem of unrealistic assumption that lumber consumption propensity is the same all over the whole period. That is, it was used to reflect the abrupt change in socio-economic situation influencing on lumber consumption or the change in lumber consumption structure itself occurring over long time. Because lumber consumption structure connects with domestic construction activity closely, the assumption that lumber consumption structure maintains identically over long time is unrealistic. So, the usage of dummy variable is necessary (Hetemaki, 2001).

It was assumed that increase in construction permit area effects positively lumber demand, while an increase in own price decreases its demand, which implies that $a_1 \leq 0$ and $a_2 \geq 0$ in equation (2).

Equation (2) was estimated by ordinary least squares (OLS) method using software EViews 3.1.

(2) Vector autoregressive model

It was considered important to analyze whether the systems approaches, such as vector autoregressive and Vector error correction methods, could provide better forecasts than the partial approach.

The idea underlying the vector autoregressive model is first to summarize the dynamic correlation patterns among observed data series and then use this summary to explain and predict likely future values for each m series. Mathematically, a vector autoregressive expresses the current value of each of series as a weighted average of the recent past of all the series plus a term that contains all the other influences on the current values. A vector autoregressive representation of order p can be expressed as

$$Y_t = a_0 + a_1y_{t-1} + \ldots + a_py_{t-p} + \mu_t \quad (3)$$

where Y_t is the $m\times1$ vector of lumber quantity for year t, and μ_t is a vector of innovations.

The innovations measure the extent to which y_t cannot be determined exactly as a linear combination of y_t the past values of with weights given by the constant coefficients a_0 and a_t, $t = 1, \cdots, p$. It is assumed that μ_t is a random vector with zero mean and error covariance matrix positive definite, and that is uncorrelated with lagged values of y_t. The vector autoregressive can be rewritten as

$$\Delta y_t = \Pi y_{t-1} + \sum_{i=1}^{p-1} \Gamma_i \Delta y_{t-i} + bx_t + \varepsilon_t \quad (4)$$

where

$$\Pi = \sum_{i=1}^{p} a_i - I, \Gamma_i = -\sum_{j=i+1}^{p} a_j,$$

and x_t is the $m\times1$ vector of variable influencing lumber quantity for year .

In the general, vector autoregressive system with variables, if all the variables are stationary, using an unrestricted vector autoregressive in levels is appropriate. If these variables are all but no cointegration relation exists, then application of unrestricted vector autoregressive in first differences is appropriate. However, if the variables are cointegrated, then one can model the system as a vector error correction model (Hall, 1999).

3. Prediction accuracy

There are a number of different measures of prediction accuracy, none of which can be regarded unambiguously as the best. The most widely used measure is probably the Root Mean Squared Error (RMSE), given below.

$$RMSE = \sqrt{\frac{1}{h+1}\sum_{t=s}^{s+h}(f_t - a_t)^2} \quad (5)$$

where the prediction sample is $t=s$, $s+1, \ldots, s+h$, h being the number of prediction periods, and the actual and predicted values in period t are denoted as a_t and f_t, respectively.

The prediction error statistics depend on the scale of dependent variable, but since we are comparing predictions for the same series across different models, this is not a problem.

The RMSE can be decomposed to three components; the bias proportion, variance proportion, and covariance proportion. The bias proportion indicates how far the mean of the predictions is from the mean of the actual series, variance proportion indicates how far the variation of the predictions is from the variation of the actual series, and covariance proportion measures the remaining unsystematic prediction errors.

Table 1. Comparison of the estimation results of lumber demand function by models.

Model	F-ratio	P>F	Adj.R^2	DW
Partial multivariate	115.44***	0.00	0.95	1.87
Vector autoregressive	76.68***	0.00	0.93	1.76

***Reject the null hypothesis that all coefficients are zero at 1% significance level

Results and Discussion

1. Function estimation

(1) Partial multivariate model

Equation (6) is the result of lumber demand estimation by partial multivariate model. Lumber demand (QLD_t) was explained by own price (LP_t), construction permit area (CPA_t), and dummy variables (D_{88} and D_{98}). Dummy variables were used to reflect the boom of lumber demand in 1988 and the abrupt decrease in 1998. All variables are in logarithms, except dummies. T-statistics are in parentheses below the coefficients.

$$QLD_t = 6.04 - 0.18\ LP_t + 0.39\ CPA_t + 0.49D_{88} - 0.93D_{98} \quad (6)$$
$$(48.3)^{***}\ (-1.7)\quad (5.9)^{***}\quad (6.1)^{***}\ (-7.5)^{***}$$

The result showed that the coefficient signs of all variables were as expected. However, own price was not significant. The test results are shown in Table 1. The partial multivariate model has good explanatory power. And there is no autocorrelation.

(2) Vector autoregressive model

Equation (7) is the result of lumber demand estimation by vector autoregressive model. The tests for stationarity and cointegration indicated that all the variables were stationary in first differences and no cointegration relation existed. Also, the pairwise Granger causality tests indicated that a system with two endogenous variables, lumber demand and construction permit area, was most appropriate. According to the AIC and SC information criteria, the lag length was preferred one. Thus, lumber demand of any period (QLD_t) was explained by the demand in the previous period (QLD_{t-1}), the construction permit area in the previous period (CPA_{t-1}) and dummy variables (D_{88} and D_{98}). All variables are in logarithms, except dummies. T-statistics are in parentheses below the coefficients.

$$QLD_t = 5.75 + 0.08QLD_{t-1} + 0.25CPA_{t-1} + 0.53D_{88} - 1.21D_{98} \quad (7)$$
$$(8.26)^{***}(0.72)\quad (7.35)^{***}\ (5.79)^{***}\ (-9.49)^{***}$$

The result showed that the coefficient signs of all variables were as expected. However, lagged quantity was not significant. The test results are shown in Table 1. The vector autoregressive model has good explanatory power. And there is no autocorrelation.

Table 1 compared the estimation results of lumber demand function by models. The significance levels of F-statistics in both models rejected the hypothesis that all coefficients are zero. In the case of lumber demand function, the estimations by partial multivariate and vector autoregressive model showed similar explanatory power.

2. Prediction accuracy

We can compare the prediction accuracy of lumber demand function by models on the basis of the Root Mean Squared Error (RMSE) and its components. Table 2 is the result of comparing the prediction accuracy of lumber demand function by models. The lower the RMSE value, the better the prediction.

The RMSE values are similar for both partial multivariate and vector autoregressive models. For both partial multivariate and vector autoregressive models, less than one percent of the prediction error consisted of systematic bias. The variance proportions are similar for both partial multivariate and vector autoregressive models. And the covariance proportions are also similar for both models. In summary, based on the RMSE and its decomposition, the partial multivariate and vector autoregressive models appear to provide roughly equally good predictions. That is, the lumber demand can be equally accurately predicted by partial multivariate and vector autoregressive models in Korea. The estimated values from partial multivariate and vector autoregressive mod-

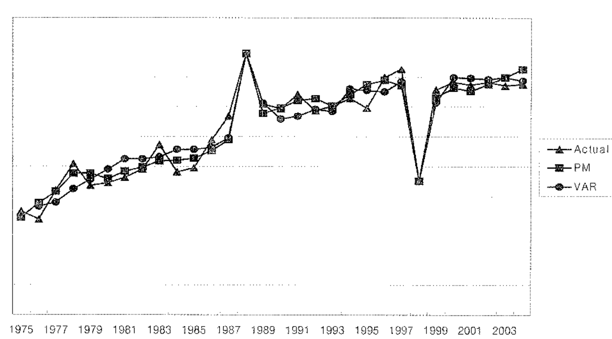

Figure 2. Comparison of the actual and estimated values of lumber consumption by models.

Table 2. Comparison of the prediction accuracy of lumber demand function by models.

Model	RMSE	Bias proportion	Variance proportion	Covariance proportion
Partial multivariate	0.07	0.00	0.01	0.99
Vector autoregressive	0.08	0.00	0.02	0.98

els and the actual values of lumber consumption are shown in Figure 2.

Conclusion

This study presented a first attempt to compare the lumber demand prediction accuracy of partial multivariate and vector autoregressive models in Korea. The partial multivariate model of lumber demand was specified with three explanatory variables; own price, construction permit area and dummy. The vector autoregressive model was specified with two endogenous variables, lumber demand and construction permit area with one lag. The dummy variable reflected the boom of lumber demand in 1988, and the abrupt decrease in 1998. On the other hand, the prediction accuracy was estimated on the basis of Root Mean Squared Error.

In the case of lumber demand, the estimation by partial multivariate and vector autoregressive model showed similar explanatory power. And the prediction accuracy of lumber demand was similar in the case of using partial multivariate and vector autoregressive model.

Literature Cited

1. FAO. 2006. FAOSTAT Forestry.
2. Hall, R., Lilien, D., Sueyoshi, G., Engle, R., Johnston, J., and Ellsworth, S. 1999. EViews manual. Quantitative Micro Software Inc.
3. Hetemaki, L., Hanninen, R. and Toppinen, A. 2001. A system for short-term forecasting of the Finnish forest sector. Finnish Forest Research Institute.
4. Joo, R., Lee, S. and Youn, Y. 1997. Market modeling of Forest products in Korea. Korea Forest Research Institute.
5. Korea Customs Service. 2005. Statistical Yearbook of Foreign Trade.
6. Korea Forest Service. 2005. Statistical Yearbook of Forestry.
7. Korea Forest Service. 2005. Statistics of Forest Products' Trade.
8. Korea National Statistical Office. 2005. Korean Statistical Information System.
9. Seok, H. 1992. The economic analysis of Forestry in Korea. Research paper of Korea Rural Economic Institute 264.

(Received August 22, 2006; Accepted October 18, 2006)