



## A Comparative Study of the Performance of Stepwise Multiple Regression, Ridge Regression, Box–Jenkins Method, and Holt’s Method in Forecasting Khao Dawk Mali 105 Rice Production.

Pheerach Rajapitaks<sup>1\*</sup>, Teeratas Sirisattayanupong<sup>1</sup>, and Pirapat Pumkong<sup>1</sup>

<sup>1</sup>Princess Chulabhorn Science High School Chonburi, Nong Chak, Ban Bueng, Chonburi, 20170 Thailand

\*E-mail: 6706390@pccchon.ac.th

### Abstract

This study investigated the relationships between climatic factors and Khao Dawk Mali 105 rice production, developed forecasting equations using stepwise multiple regression and ridge regression, constructed time series models using the Box–Jenkins and Holt’s methods, and compared their forecasting performance using the Mean Absolute Percentage Error (MAPE). Annual secondary data from Ubon Ratchathani Province covering 2006–2023 (18 years) were analyzed. The variables included average rainfall, average temperature, relative humidity, wind speed, and Khao Dawk Mali 105 rice production. Pearson’s correlation analysis was performed to examine the relationships among the variables.

The results of the study revealed that: (1) average temperature and average rainfall were moderately positively correlated with the production quantity of Khao Dawk Mali 105 rice, whereas average wind speed exhibited a moderately negative correlation and average relative humidity showed a weak negative correlation with the production quantity of Khao Dawk Mali 105 rice. The most appropriate forecasting equation was obtained from the stepwise multiple regression model, The resulting forecasting equation is as follows:

$$\hat{Y} = -2,227,142.4 + 100,489.2(temp) + 1945.1(rain) - 64,371.3(wind)$$

(2) The time series model developed for forecasting the production quantity of Khao Dawk Mali 105 rice using the Box–Jenkins method was

$$(1 - 0.2503B + 0.5627B^2)(1 - B)Y_t = 28375.913 + \varepsilon_t$$

while the forecasting model developed using Holt’s method was

$$\hat{Y}_{t+h} = 0.0083Y_t + 0.9917(L_{t-1} + T_{t-1}) + h[0.0001(L_t - L_{t-1}) + 0.9999T_{t-1}].$$

(3) The stepwise multiple regression and ridge regression models yielded MAPE values of 13.01% and 14.85%, respectively. Among the time series models, the Box–Jenkins ARIMA(2,1,0) with drift model achieved the lowest MAPE (4.30%), followed by Holt’s method (5.30%), indicating that the Box–Jenkins model was the most accurate.

**Keywords:** Stepwise Multiple Regression, Ridge Regression, Box–Jenkins Method, Holt’s Method