EKSAKTA
Journal of Sciences and Data Analysis

E-ISSN: 2720-9326
P-ISSN: 2716-0459
DOI: 10.20885/EKSAKTA.vol1.iss2.art11

Research Article

Utilization Copula in Determination of Shallot Insurance Premium Based on Regional Harvest Results

Hasna Afifah Rusyida[1,*], Achmad Zanbar Soleh[1], Lienda Noviyanti[1], Anna Chadidjah[1], Fajar Indrayatna[1]

[1] Department of Statistics, Universitas Padjadjaran, Sumedang Regency, 45363, Indonesia
[*] *Corresponding author: hasna.afifah@unpad.ac.id*

Received: 3 September 2020; Accepted: 30 September 2020; Published: 1 October 2020

Abstract: Shallot is one of the highest-yielding horticultural crops in Indonesia and has the tendency to increase the profits of farmers in Indonesia. But until now in Indonesia there is no insurance for horticultural crops other than corn, whereas the shallot farmers face various sources of risk such as weather changes, pest attacks, or other technical factors that ultimately lead to uncertainty of agricultural yields (revenue risk). To overcome this loss, insurance companies can make products based on shallot yields and shallot market prices. Therefore it is essential to grasp the distribution of risk variables (shallot yields and shallot market prices) that interact simultaneously, not separate from one another. Omitting dependencies among risk variables can cause biased risk estimation. Copula can model the non-linear dependencies and can identify the structure of the dependencies between variables. The suitable copula for modeling yield and price risk of shallot is simulated to compute the premium. Result show that clayton copula is suitable for dependence modelling between risk variables.

Keywords: Clayton Copula, horticulture crops insurance, non-linear dependencies

Introduction

In Indonesia, the role of agriculture sector is as a source of food for the community, the source of national income, opening job opportunities, source of investment, and producers of foreign exchange when the products of agricultural products are exported to other countries. On the other hand, agriculture is one of the most vulnerable sectors of the economy of the negative impacts of climate action, such as flood, aridity, and so on.

Less conducive climate change triggers the intensity of socioeconomic environmental factors as a source of risk and uncertainty. The occurrence of floods resulted in damaged transportation infrastructure so that the distribution of goods and services, including inputs and outputs of farming to be not smooth. Availability (quantity, quality, time, place) farm input on the market does not match the farmers' needs. The average price of these items also tends to be more expensive because transportation costs per unit increase. On the other hand, the price of farm output tends to decrease. Extremely high rainfall intensity (especially if it occurs at harvest) causes the quality and quantity of farm output to drop dramatically. This condition if accompanied also by the worsening condition of transportation infrastructure, cause the output price of farming became very low, even in certain cases farmers cannot market it. Generally, farmers face various sources of risk that ultimately lead to an uncertainty of farm produces or yield risk and marketing price or price risk of farm products (revenue risk). Because of that, agricultural insurance is needed to manage the risks faced by farmers.

Indonesia is one of the potential producers of horticultural plants. Horticultural plants have a high chance of increasing farmers' profits in Indonesia. But until now there is no insurance for horticulture other than corn, even though the Director of Vegetables and Medicinal Plants in ministry of agriculture has requested that agricultural insurance be applied to the horticultural commodity [8]. In addition, the farmers also requested that horticulture plants can be insured because the cost of cultivation is expensive so if something unexpected happens, such as market prices plummet that can be borne by the insurance company [1]. From BPS obtained data on seasonal vegetable production in Indonesia, during 2017 are as follows:

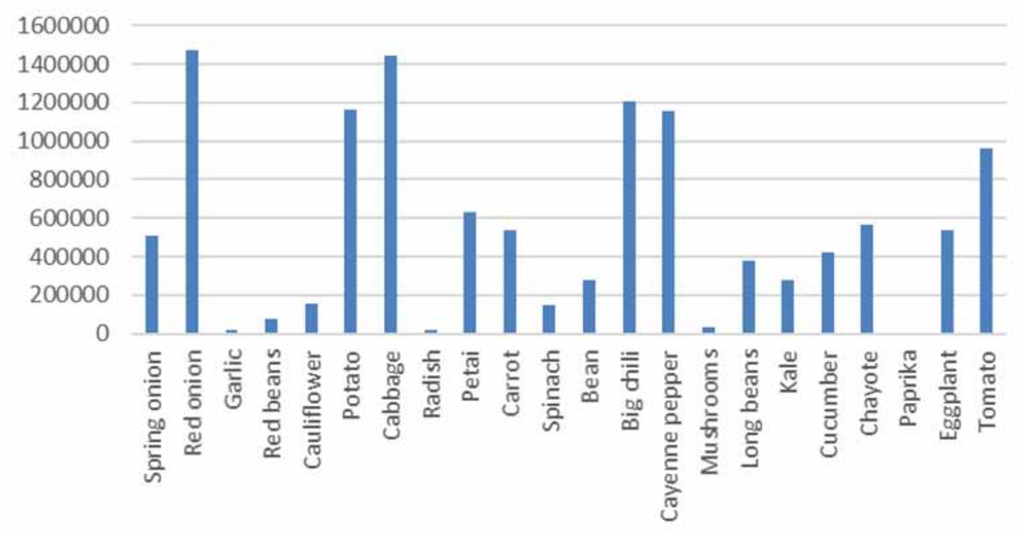

Figure 1. Seasonal Vegetable Production 2017 pattern

From these data, it can be seen that shallots are among the highest-yielding vegetable commodities in Indonesia. In addition, shallots have been exported. Based on data from the Central Statistics Agencies (BPS) in 2018, Indonesia exported 9000 tons of shallots to Thailand and in 2016 Indonesia exported 355 tons of red chili so that these commodities became a source of income and opportunity work for the community, and considered capable of contributing significantly to regional economic development, the government has also implemented various strategies to maintain the quantity of production of this commodity.

The principle models of risk management define that crop revenue insurance should be derived from the joint probability distribution of yields and prices of multiple crops. Based on the [11], it is known that some of the joint distribution may be decomposed into N marginal distributions and a Copula function. This function completely describes the dependency structure between N variables and dependency measurements among the more informative variables rather than linear correlations, when the joint distribution of variables is non-elliptical. When the correlation measures the dependencies between multiple variables with a single number, a Copula describes the complete dependencies between the various variables[6].

COPULA MODEL FOR SHALLOT INSURANCE

Copula functions

According to the previous explanation, it can be seen that essentially agricultural risk is yield shortfalls, price collapses and revenue losses. In designing and rating crop insurance contracts, it is important to understand the distributions of several risk variables interacting simultaneously, not isolation of one another. Ignoring dependencies among risk factors can lead to biased estimates of the risk. For example, in the case of the natural hedge in which revenue is stabilized because of the negative relationship between crop yields and prices. If the negative relationship between price and yield is ignored, it will overestimate the risk of the revenue insurance. Thus, it is important to be able to adequately model dependence and multivariate outcomes [13].

Copula is used to model the distributions of multivariate risk variables (yield and prices of shallot) simultaneously, the dependence structure between multivariate risk variables can capture by copula. Copula is a Latin which means relationship or bond. The word copula was first used in [11]. This theorem is used to describe a function that can combine several distribution functions from several random variables to a joint distribution function. Copulas are parametrically specified joint distributions generated from given marginal [12]. The joint distribution function of some random variables with various distributions can be identified by copula because copula can split the marginal behavior and the dependence structure of the variables from their joint distribution function.

The distribution function of randomized variables that have been combined to form a joint distribution function in the copula form is called the marginal function of the copula. This marginal function spreads uniformly (0,1) [9].

$$f(x_1,x_2,\cdots,x_d)=c(F_1(x_1),F_2(x_2),\cdots,F_n(x_d))\prod_{i=1}^{d} f_i(x_i). \tag{1}$$

If H is a distribution function having d dimension with a cumulative distribution function $F1$, $F2$, $\cdots$, Fd, then will come up a copula for all $\boldsymbol{x}$ in $\bar{\boldsymbol{R}}d$, the copula associated with F is a function distribution C: $[0,1]^d \to [0,1]$ with margin $U(0,1)$, as follows:

$$H(x_1,x_2,\cdots,x_d)=C(F_1(x_1),F_2(x_2),\cdots,F_d(x_d);\ \theta). \tag{2}$$

Where θ is a vector of parameters of the copula which are named the dependency parameters, which quantify dependence between the marginals. It is the principal focus of estimation.

The Equation (2) will be used in fitting copula-based models to the data in determining shallot insurance

Maximum Likelihood Estimation (MLE) method

MLE is one estimation method that can be used in estimating copula parameters. When estimating a copula's parameters several MLE methods exist. They include standard MLE where all of the (marginal and dependence) parameters are estimated simultaneously [5]. The likelihood function for copula is as follows:

$$\begin{aligned} l(\vartheta_1,\vartheta_2,\theta) &= \sum_{t=0}^{T} In\, c(F_X(x_t;\vartheta_1),F_Y(y_t;\vartheta_2);\theta) \\ &+ \sum_{t=0}^{T} In\, (f_X(x_t;\vartheta_1),f_Y(y_t;\vartheta_2);\theta), \end{aligned} \tag{3}$$

with sufficient conditions

$$\frac{\partial l}{\partial \vartheta_1}=\frac{\partial l}{\partial \vartheta_2}=\frac{\partial l}{\partial \theta}=0. \tag{4}$$

So that the parameter estimation is obtained in the MLE function i.e. $\hat{\vartheta}_1$, $\hat{\vartheta}_2$ and $\hat{\theta}$. The maximum likelihood estimator does not confidence intervals but is typically faster [4].

Empirical framework

This study uses secondary data from the Bandung District Agriculture Office, county-level Rancabali (An area in the district of Bandung, West Java Province, Indonesia) quarterly yields and price data for shallot.

Future Yields and Price Data

November 2018 shallot future yields and price data obtained through ARIMA Model. This model is a Box-Jenkins variant of the ARMA model where the model needs analysis of stationarity. Ordinarily, common ARIMA model (without seasonal patterns) have the following mathematical models.

$$\phi_p(B)(1-B)^d Z_t = \theta_0 + \theta_q(B)a_t, \tag{5}$$

where $\phi_p(B) = (1-\phi_1 B - \cdots - \phi_p B^p)$ and $\theta_q(B) = (1-\theta_1 B - \cdots - \phi_p B^q)$ and $(1-B)^d$ is a non-seasonal differencing in order d. This equation can also be written as ARIMA (p, d, q). When the data is stationary without differentiation process then the used model is ARMA (p, q). ARMA model have the following mathematical models.

$$X_t - a_1 X_{t-1} - \cdots - a_p X_{t-p} = \varepsilon_t + b_1 \varepsilon_{t-1} + \cdots + b_q \varepsilon_{t-q}, \tag{6}$$

where $a_1, a_2, \dots, a_p, b_1, b_2, \dots, b_p \in \mathbb{R}, \varepsilon_t \sim WN(0,\sigma^2)$

The ARIMA is used for predict future yields and prices on March for rice and corn. The ARIMA model used is in accordance with the conditions of the data.

Design of Multi Crops Insurance Using Copula

Copula will be applied to model the shallot insurance. Copula will model the price-yield dependence. Copula is broadly divided into several families such as Gaussian, t-Student, Archimedean, and others. the determination of the copula family that will be used is based on the structure of dependencies formed between the multi-dimensional variables of prices and yields, a Monte Carlo simulation method is applied to simulate the multi-dimensional variables of price and yields of shallots from the copula function that have been estimated by MLE method, which hold the rank correlations of these variables. After that the premium rates of the revenue insurance contract can get through the following calculations.

The indemnity scheme is

$$max[(\lambda R^e - R), 0], \qquad (7)$$

where $R = Y * P$ is the revenue, Re=E(R) is the expected revenue and λ ∈ (0, 1]) is the coverage level.Y and P are the non-negative random variables reflecting price and yield risks of shallot, with the joint probability distribution F in a certain copula function form F (P,Y) = C(F1(P), F2(Y); θ). If R < λRe, the insurer will pay (λRe − R) as an indemnity. An actuarially fair premium is equal to the expected loss of this contract. Suppose an insurance contract will insure some proportion λ of the mean crop revenue (Re), the expected loss for this insurance contract that guarantees λ × 100% of the predicted revenue (Re) takes the form of

$$\mathrm{EL}(R) = E[(\lambda R^e - R)I(R \leq \lambda R^e)] \qquad (8)$$

where R denotes the observed quarterly revenue of this insurance and Re represents the predicted revenue. The expected value of the revenue risk Re considered the marginal risk measures of price and yield risk. The marginal distribution of yield and price can be determined through the historical yield and price data. Copula is applied to build the joint distribution of yield and price risk.

The revenue loss and the premium rates of the revenue insurance contract can get at certain coverage levels. The scenario can use coverage level at 80 percent respectively. The mean and variance of expected revenue , the expected loss, and actuarially-fair premium rates for shallot insurance that assure a certain percent of the expected level of revenue will be computed.

Results and Discussion

To assign the bivariate model, the two models must be decided for the marginal distributions and one model that reflects the dependence structure between the marginals.

Modelling the marginal distribution

The univariate data (price and yield series) will be modeled by ARIMA model, the identification of the ARMA lags was done empirically by means of ACF and PACF plots for AR lags and MA lags respectively for the data. The optimal model that fit to the shallot yields is ARIMA (0,1,1) model. The model that allows the best fit to the shallot price is the ARIMA(0,2,1) model. The estimated parameters and their standard errors are informed in Table 1.

Table 1. ARIMA-t estimation

Parameters	Shallot yields		Shallot Price	
	Estimate	Error	Estimate	Error
φ_1	-	-	-	-
φ_2	-	-	-	-
θ_1	-1	0.1178	-1	0.1084
θ_2	-	-	-	-

The capability of the estimated models was also tested by Ljung-Box Q-test (LB). LB test is applied to the standardized residuals of marginal models. The LB test fails to reject the null hypothesis that indicates no significant serial correlation in the standardized residuals at the 5% significance level (pvalue = 0.9518 for shallot yields ARIMA Model and pvalue = 0.8517 for shallot price ARIMA Model.). Besides that, the Kolmogorov Smirnov test is applied to residual of marginal models. The Kolmogorov Smirnov test fails to reject the null hypothesis that indicates residual series are normally distributed at the 5% significance level (pvalue = 0.111 for residuals from shallot prices ARIMA model and pvalue = 0.2053 for residuals from shallot yields ARIMA model). Based on the results of the diagnostic check it can be concluded that each arima model is a good model for marginal data (shallot yields and shallot prices).

Modelling the dependence structre

After modeling each data marginally using the ARIMA model, the next step is to analyze the dependency structure between the data using the Copula. The step before analyze the dependence structure is determination the appropriate marginals. The Kolmogorov Smirnov goodness-of-fit statistics suggest that the yield of shallot marginal is uniform and the price of shallot marginal is three-parameter lognormal distribution. The marginal MLE of yields are estimated with shallot yields $y \sim U(167.49, 1290.9)$, shallot price $p \sim lognormal(14.03, 0.59687, 0.000036421)$. In the next step the dependence parameters for shallot prices and shallot yields are estimated given the marginal estimates from the step before. Akaike Information Criterion (AIC) is one of the criterions to select copula. [3] has selected the corresponding bivariate Archimedean Copula function which describes the data dependency structure through some criterias and the result shows that AIC is good for selecting bivariate copula. Besides that, the AIC tends to be superior in small samples [10].

The estimated parameter corresponding to the best fit copula and the AIC are reported in Table 2.

Table 2. Copula Parameter Estimation

Copula	Estimated Parameter	AIC
Clayton	0.56	-1.16
Gaussian	0.33	-0.19
t	0.34	1.67
Gumbel	1.27	-0.48
Frank	2.02	-0.64

Table 2 shows that the Clayton copula is the one that allows the best fit to the data since it has the lowest values for the AIC.

The Clayton copula is an asymmetric Archimedean copula, exhibiting greater dependence in the negative tail than in the positive. The Clayton copula is mostly used to study correlated risks because of their ability to capture lower tail dependence. However the level of dependence between the marginals depends on the value of the copula parameter. When the copula parameter value is equal to 0 then the marginals are independent, and when the parameter value goes to infinity the the Clayton copula approaches the Fréchet-Hoeffding upper bound. Due to the restriction on the dependence parameter, the Fréchet-Hoeffding lower bound cannot be attained by the Clayton Copula. This indicates that the Clayton copula cannot account for negative dependence[7].

The scatter plot of random sample simulated from this Copula were drawn in Figure 2.

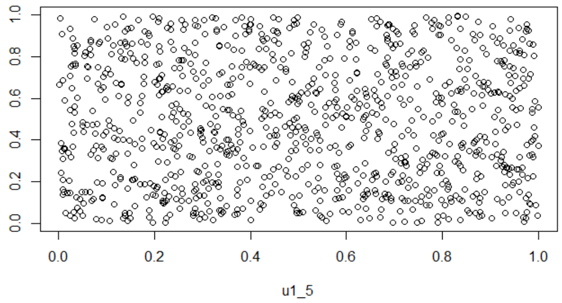

Figure 2. 1000 simulated points plot of the Clayton copula with parameter 0.56

The lower tail dependence is not visible because of the small value of θ.

Dependence Measures

The calculation of the copula based dependence measures are Kendall's tau, Spearman's rho, tail dependence is straightforward. Tail dependence measures the dependence between X and Y in the upper-right and lower-left quadrant of the joint distribution function [2].

The relation between the Clayton copula parameter and the dependence measures is Kendall's tau and the tail dependence are uncomplicated but the association between the copula parameter and the Spearman's rho is complex. [7] so that, we don't use Spearman's rho correlation. The dependence measure of shallot yields and shallot price are as follow:

Table 3. The Estimated Kendall's tau, tail dependence based on the estimated copula

Kendall's tau	Upper tail	Lower tail
0.22	0.00	0.29

The value of the upper tail dependence is equal to zero. This is because the Clayton Copula cannot capture the upper tail dependence.

Policy Implication

The dependence structure is imposed by the Clayton copula as estimated before. 1000 revenue series were drawn from the Clayton copula which hold the dependence structure between prices and yields shallot as well as the uniform yields and lognormal three parameter price. The simulations are used in counting the expected loss and premium rates for the revenue insurance contract at 80% coverage level at November 2018. Based on simulated predicted revenue and restriction data (if the simulation of revenue > Rp.3.567.040.000 (the maximum losses from historical data),- and indemnity ≤ 0 data is not used), There are 720 data < Rp.3.567.040.000 and 494 data > 0 so obtained that expected loss = $\frac{494}{720} \times 797.810.000 =$ 547.386.306 (797.810.000 is the non zero expectation of indemnity that is uniformly distributed). So that premium = Rp. 547.386.306 for whole Rancabali area. The total area planted with shallots in November 2018 in Rancabali is 8 hectares, so the amount of the premium rate is 68.423,29 per m2.

Conclusions

This study has looked at the horticultural insurance in Indonesia especially shallots which is the highest-yielding vegetable commodities in Indonesia but until now there is no insurance for it. To make this insurance product, copula is involved in calculating the premium. Copula is applied to modelling dependencies between price and yield of shallot in order to get an unbiased risk estimate.

The Clayton copula with a parameter value equal to 0.56 is the one that allows the best fit for our data. Furthermore, the dependency modeled by this copula is highly discussed. The results of this analysis indicate that the joint distribution of yield and shallot prices shows on higher lower tail dependences. It is imply that that large losses from yields and prices of shallots are more likely to occur simultaneously than large gains.

From the simulation obtained that premium for shallot insurance is Rp. 547.386.306 for whole Rancabali area. so the amount of the premium rate is 68.423,29 per m2.

Acknowledgment

We would like to thank READI Project for funding this research in the academic year of 2019.

References

[1] Choirul Anam, Petani Minta Tanaman Hortikultura juga Diasuransikan. Retrieved from https://ekonomi.bisnis.com/read/20150805/99/459627/petani-minta-tanaman-hortikultura-juga-diasuransikan, (2015).
[2] Cherubini, U., Luciano, E., & Vecchiato, W., *Copula methods in finance*. John Wiley & Sons., 2004.
[3] Gülöksüz, Ç. T., Comparison of some selection criteria for selecting bivariate archimedean copulas. *Afyon Kocatepe Üniversitesi Fen Ve Mühendislik Bilimleri Dergisi*, *16*(2), (2016), 250-255.
[4] Hofert, M., Mächler, M., and McNeil, A. J., Likelihood inference for Archimedean copulas in high dimensions under known margins. Journal of Multivariate Analysis 110, (2012), 133–150.
[5] Jane, R., Dalla Valle, L., Simmonds, D., & Raby, A., A copula-based approach for the estimation of wave height records through spatial correlation. *Coastal Engineering*, *117*, (2016), 1-18.
[6] Kortekaas, M.P., *A Vine Copula-GARCH Option Pricing model*. Master's Thesis in Financial Econometrics, Faculty of Economics and Business, University of Amsterdam, 2013.
[7] Mahfoud, M., & Michael, M., Bivariate Archimedean copulas: an application to two stock market indices. *BMI Paper*, *1517333*, 2012.
[8] Nur Aini, Kementrian ingin asuransi pertanian bias untuk hortikultura. Retrieved from https://www.republika.co.id/berita/ekonomi/keuangan/18/02/27/p4t4sb382-kementan-ingin-asuransi-pertanian-bisa-untuk-hortikultura, 2018.
[9] Nelsen RB., *An Introduction to Copulas*. Portland (US): Springer, 2005.
[10] Shumway, Robert H. and Stoffer, David S., *Time Series Analysis and Its Applications.* New York: Springer, 2011.
[11] Sklar, A., Fonctions de r´epartition `a n dimensions et leurs marges. *Publications de l'Institut de Statistique de l'Universit´e de Paris 8*, (1959), 229–231

[12] Trivedi, P. K., & Zimmer, D. M., Copula modeling: an introduction for practitioners. *Foundations and Trends® in Econometrics*, *1*(1), (2007), 1-111.

[13] Zhu, Y., Ghosh, S. K., & Goodwin, B. K., *Modeling dependence in the design of whole farm---a copula-based model approach* (No. 382-2016-22619), 2008.