Modeling multivariate time series with copulas: Implications for pricing revenue insurance

Gislaine Vieira Duarte ⓘ
Escola Superior de Agricultura Luiz de Queiroz, Universidade de São Paulo, Brasil

Vitor A. Ozaki ⓘ
Escola Superior de Agricultura Luiz de Queiroz, Universidade de São Paulo, Brasil

Much of the soybean produced in Brazil is exported and, consequently, the domestic soybean price (R$) is greatly influenced by the price traded at the Chicago Mercantile Exchange Group (CME Group) (US$). Therefore, to model the dependency structure between soybean yield and price, the exchange rate must be incorporated into the modeling. This study aims to model the dependency structure between these three variables using the Copula methodology, calculate the crop revenue insurance rates, and compare with the rates offered in the insurance market. The rates applied by the Brazilian insurance market are overpriced when compared to the methodology presented in this study with the incorporation of the dollar rate in the modeling, which could increase the problem of adverse selection exchange and hamper massification of agricultural insurance in the Brazilian territory.

Keywords. Three-dimensional parametric copulas, Agricultural risk management, Revenue insurance.

JEL classification. Q1, C10.

1. Introduction

In the work of Duarte et al. (2017), we analyzed the dependency structure between yield and monthly nominal average prices (in Reais[1]) received by soybean producers in the state of Paraná. This spot price is collected from regional centers and weighted by the relationship between production in the region and state, and then the simple arithmetic average is used to calculate the monthly average price.

The purpose of a business in the spot (or physical) market is to make an immediate purchase and/or sale. Thus, the delivery and payment of the product occur concomitantly. This type of market is typically sporadic and has a high degree of uncertainty regarding price behavior, supply regularity, and product quality. Therefore, a company that adopts this mechanism assumes a great risk, which may lead to failure in its operations due to uncertainties related to product demand and supply.

Gislaine Vieira Duarte : gislainevduarte@gmail.com
Vitor A. Ozaki : vitorozaki@usp.br

[1]Brazilian currency (R$)

One way to reduce those risks is using future markets, known as hedging, which is a price-locking operation in a future date. In most contracts, financial settlement occurs prior to the delivery date and does not require physical delivery to the agreed location. For example, soybean produced in Brazil is usually traded at CBOT/CME; however, the physical delivery of the commodity is in China or the European Union.

Cooperatives, or trading companies, that trade Brazilian soybean use the futures market to price the commodity for future delivery. Thus, since much of the soybean produced in Paraná is exported, the US dollar volatility has a great influence on prices received by farmers and thus impacts directly in their income.

Therefore, in the income insurance context it is important to study not only the effect of the dependency structure between agricultural yields and price but also when the exchange rate variable is incorporated in the analysis.

In Brazil, two insurance companies offer revenue insurance to farmers, from now on called insurer A and insurer B. Insurer A uses the annual productivity series in its pricing for each municipality and the daily physical price series (CEPEA/ESALQ indicator). This indicator represents the average trading value of Brazilian export-grade bulk soybeans, referenced at the Paranaguá Port, Paraná, Brazil.

On the other hand, insurer B uses CME future price series transformed into Reais. In this case, the correlation between future price (CME price) and the dollar rate (EXG rate) is not taken into account and may lead to a reduction in the impact of price (CME).

In order to understand the variation and dependency of the exchange rate (EXG rate) on revenue modeling, this variable was incorporated into the model. Therefore, this paper presents an alternative approach with the incorporation of the EXG rate into the Copula model (three-dimensional approach). The study of the relationship of these three variables was not found in the literature, in the context of agricultural risk and the comprehension of the relationship between these three variables is the main contribution of the study. Recent articles in the agricultural risk context only consider the modeling of two variables, price and yield (Duarte and Ozaki, 2019).

This study modeled the dependency structure between soybean yield, CME price, and EXG rate through Copula functions.

In order to compare the inclusion and the impact of the third variable the analysis was performed under the two-dimensional and the three-dimensional approaches. In addition, in both approaches, we calculated according to Duarte et al. (2017) and estimated insurance premium revenue rates and their comparison with the rates applied by the Brazilian insurance market.

In the next section, we present the way of calculating the revenue insurance premium, under the three-dimensional approach methodology to calculate premium rate. Section 3 presents the description of the data used in the study. Section 4 presents the Copula methodology for calculating the three-dimensional distribution and the models used for modeling marginal distributions. Section 5 presents the study results. Section 6 discusses the results from both approaches. Finally, the conclusions are presented in Section 7.

2. Revenue Insurance Premium Rate

In this study, we consider revenue F as a function of three variables, soybean yield (X), USD price (Y), and EXG rate (Z) in the expression $F = XYZ$. Guaranteed yield is defined by $X_g = \lambda X_e$, where $0 < \lambda < 1$ is the coverage level (CL) chosen by the producer and X_e is the expected yield, usually calculated by the average of the last five harvests. Guaranteed price is defined by $Y_g = \lambda Y_e$, where $0 < \lambda < 1$ is the CL and Y_e is the expected price, usually calculated by the average of the last 15 days of the futures price traded at CME on the day of contracting the insurance.

Similarly, we define the guaranteed EXG rate ($Z_g = \lambda Z_e$).

In this type of insurance, we assume that the compensation paid per unit of area to the rural producer is given by: $I = \max\{(x_g y_g z_g - xyz; 0\}$.

Loss probability is given by $p = \Pr(X < x; Y < y; Z < z \mid xyz < x_g y_g z_g)$. This conditional probability expresses the cumulative joint probability of an event (loss) as a function of the random variables X, Y and Z, since the choice of XYZ values is restricted to the condition. Thus, the premium rate is given by the expected loss, according to the equation:

$$
\begin{aligned}
\mathrm{E}(Loss) &= \mathrm{E}(\max\{(x_g y_g z_g - xyz); 0\}) \\
&= \Pr(X < x; Y < y; Z < z \mid xyz < x_g y_g z_g)[x_g y_g z_g - \mathrm{E}(XYZ \mid xyz < x_g y_g z_g)].
\end{aligned}
$$

Therefore, the optimum prize rate is given by

$$
\pi = \frac{\Pr(X < x; Y < y; Z < z \mid xyz < x_g y_g z_g)[x_g y_g z_g - \mathrm{E}(XYZ \mid xyz < x_g y_g z_g)]}{x_g y_g z_g}.
$$

Note that the expected revenue is related to risks in marginal distributions of price, yield, and USD. The marginal distribution of yield and price is inferred from historical yield and price data using empirical and parametric methods. In turn, the construction of the joint distribution uses a parametric copula function. The theoretical concepts related to determining the premium rate of the two-dimensional revenue insurance were described by Miqueleto (2011) and Brisolara (2013).

3. Data Description

For yield modeling, we analyzed the annual grain yield series (kg/ha) of the municipalities of Cascavel and Toledo in the state of Paraná, Brazil, from 1980 to 2016, totaling 36 observations, and available by Ipardes (2017). For price modeling, we used the daily price series of soybean futures contract traded at Chicago Board of Trade (CBOT) of CME Group due in March 2017 (code ZSH17). These data are available in Barchart (2017).

Futures contract prices, Y_{CME}, are available in cents/bushels; therefore, the transformation factor

$$
FT = \left(\frac{\dfrac{Y_{CME}}{100}}{27{,}216} \times 60 \right)
$$

was used to convert them into USD for 60-kg bags. The observation period was from Apr 05, 2015 to March 14, 2017, totaling 455 observations.

The daily rate of trade quote for the same period was taken from the Macroeconomic and Regional Database of the Institute of Applied Economic Research Ipea (2017), on Feb 05, 2017.

To obtain the same periodicity of price and EXG rate series, we simulated the soybean yield reaching 455 observations; thus, allowing the use of the Copula methodology for the three variables with the same periodicity (455 observations each). In addition, a second analysis was performed simulating 20,000 observations of the three series from the adjusted Copula. Finally, the commercial rates charged by insurance coverage applied by insurer A are available from MAPA (2017).

4. Methodology

The methodological steps of this work consist of:

- **Step 1)** Yield modeling.
  Agricultural yield over the years has shown an upward trend, due to the great advance in technologies used in crops, such as planting techniques, the use of inputs, machinery, among others. Therefore, the deterministic procedure suggested by Gallagher (1987) and Tejeda and Goodwin (2008) is used for the yield detrend.

- **Step 2)** CME Price and exchange modeling.
  Soybean prices are less predictable than yield, due to the large international soybean market. Price and EXG rate series have a much more complex dynamic as a number of external factors. Domestic supply and demand issues also affect both variables.

  Two characteristics of these series are the trend and non-constant variance over time. Therefore, it was decided to work with the differentiated series, transforming it into a series of prices adjusted at the end of the trading day, removing the trend of these series. In addition, to extract what is important for price risk, modeling price volatility of marginal univariate series is incorporated into the final Copula model.

- **Step 3)** Copula modeling.
  To study the dependence structure of the variables in question, Copula functions are used to construct the multivariate function. In this work, two approaches were used: two-dimensional and three-dimensional.

  In the two-dimensional approach, the CME price and EXG rate daily price series were multiplied and transformed into a single price series in Reais.

  In this case, a null correlation was considered between these variables. This approach is similar to that used by insurer B when pricing agricultural revenue insurance.

  In turn, the three-dimensional approach takes into account the correlation between the CME and EXG rate price variables. To understand the risk of Brazilian soybean price, influenced by exports and, consequently, international prices, it was decided to incorporate the exogenous exchange rate variable into the revenue insurance pricing model.

4.1 *Yield Modeling*

As previously described, soybean yield has increased over the years, mainly due to the technology employed in the field.

Therefore, in order to model the yield risk we detrend the series, through the deterministic procedure suggested by Gallagher (1987) and Tejeda and Goodwin (2008). For soybean yield modeling in Cascavel and Toledo, the Normal, Skew Normal, Odd Log Logistic Normal, and Skew T distributions were used, as presented in Duarte et al. (2017).

The well-known Normal models were used as a first approach to modelling yields. The description of the methodology of the Normal and Skew Normal distribution follows Botts and Boles (1958),Just and Weninger (1999),Azzalini (1985),Ozaki and Silva (2009). In this context, we proceed detailing the OLLN and Skew T distribuitions.

Now we presented the Skew-t Student distribution, propose by Azzalini and Capitanio (2003). The X variable is called "Skew-t Student", with location parameters ($\mu \in \mathbb{R}$), scale parameter ($\sigma \in \mathbb{R}_+^*$), asymmetry parameter ($v \in \mathbb{R}$) and kurtosis parameter ($\tau \in \mathbb{R}_+^*$), then the probability density is defined by:

$$f(x;\mu,\sigma,\lambda) = \frac{2}{\sigma} t\left(\frac{x-\mu}{\sigma}\right) T\left[v\sqrt{\lambda}\left(\frac{x-\mu}{\sigma}\right)\right],$$

where $z = \frac{x-\mu}{\sigma}$, $\omega = v\lambda^{1/2}z$, $\lambda = \frac{\tau+1}{\tau+z^2}$, t_{Z_1} e T_{Z_1} are respectively, the standard t-Student (pdf) with τ degrees of freedom and the standard t-Student (pdf) with $\tau+1$ degrees of freedom. As a notation, we adopt $Z \sim ST(\mu,\sigma,v,\tau)$. Note thet when $v = 0$, the distribution is equivalent to t-Student symmetric.

Figure 1 illustrates the behavior of the ST's pdf. Note in the graphs 1a and 1b the contribution of the parameter v in the form of the distribution, with μ, σ and τ fixed. Positive values of v indicate positive asymmetry and negative values of v indicate negative asymmetry. Figure 1c shows the effect of the kurtosis parameter on the distribution when the other parameters are fixed.

Another distribution used in this work is the Odd log-logistic-Normal(OLLN) distribution. The new OLLN distribution family allows a greater flexibility of the distribution's tails. The cumulative probability function (cdf) with a shape parameter $\alpha > 0$, is defined by:

$$F(x;\mu,\sigma,\alpha) = \int_0^{\frac{\Phi(x;\mu,\sigma)}{\overline{\Phi}(x;\mu,\sigma)}} \frac{\alpha t^{\alpha-1}}{(1+t^\alpha)^2} \, dt = \frac{\Phi^\alpha\left(\frac{x-\mu}{\sigma}\right)}{\Phi^\alpha\left(\frac{x-\mu}{\sigma}\right) + \left[1-\Phi\left(\frac{x-\mu}{\sigma}\right)\right]^\alpha}, \tag{1}$$

where $\overline{\Phi}(x;\mu,\sigma) = 1-\Phi(x;\mu,\sigma)$. The OLLN density is given by

$$f(x;\mu,\sigma,\alpha) = \frac{\alpha\phi\left(\frac{x-\mu}{\sigma}\right)\Phi\left(\frac{x-\mu}{\sigma}\right)\left[1-\Phi\left(\frac{x-\mu}{\sigma}\right)\right]^{\alpha-1}}{\sigma\left\{\Phi^\alpha\left(\frac{x-\mu}{\sigma}\right) + \left[1-\Phi\left(\frac{x-\mu}{\sigma}\right)\right]^\alpha\right\}^2}, \tag{2}$$
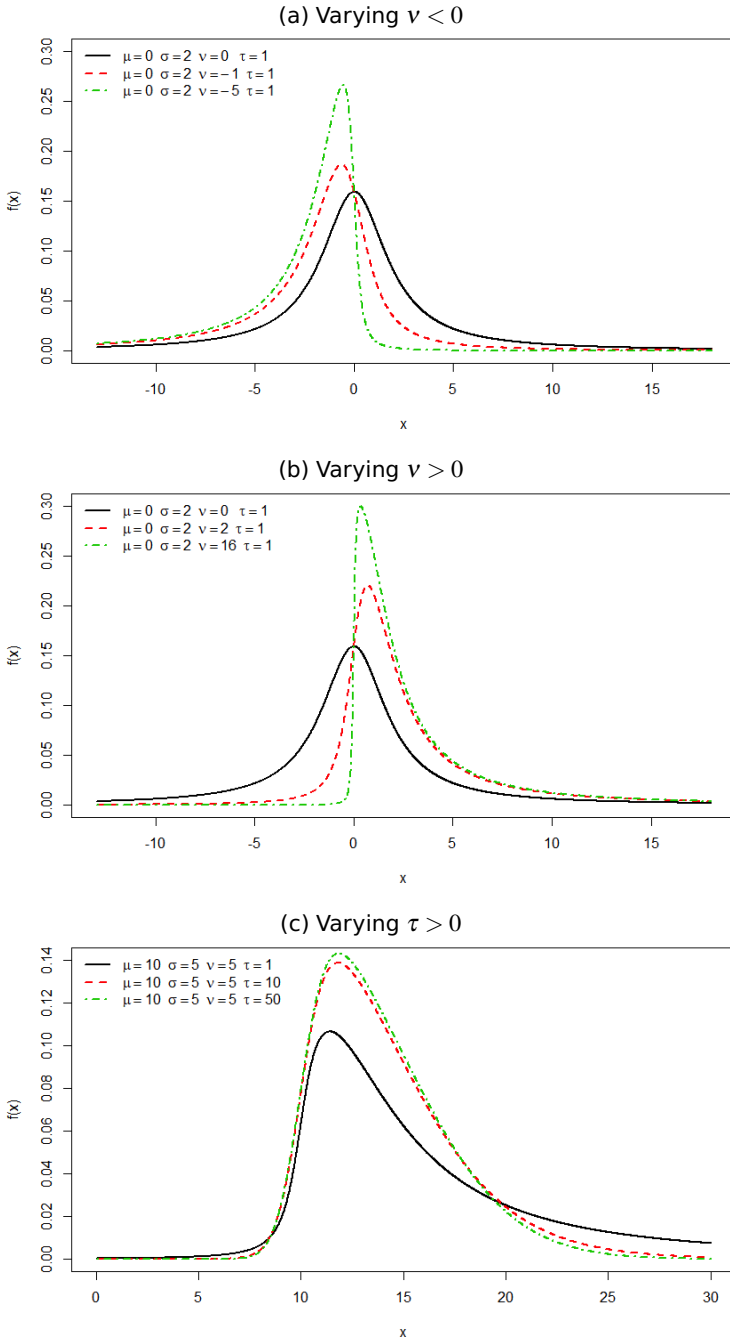
Figure 1. Plots of the Skew-*t* density function for different parameter values.

respectively. Note that $\alpha > 0$ is a shape parameter. Henceforth, a random variable with density function (2) as above is denoted by $X \sim \text{OLLN}(\alpha, \mu, \sigma)$. For $\mu = 0$ and $\sigma = 1$, we obtain the standard OLLN distribution. Further, the OLLN distribution with $\alpha = 1$ reduces to the Normal distribution.

Figure 2 shows the density of the OLLN distribution for some values of the parameters μ, σ and α. It should be noted in the plots 2a and 2b the contribution of the parameter α on the unimodality and bimodality of the distribution, when μ and σ are fixed. When the parameter α approaches zero, the pdf presents bimodality. On the other hand, when the value of α increases the function presents unimodality. It is observed that when μ varies, the plots is translated in the x-axis, regardless of the form Figure 2c.

4.2 *Price and EXG rate Modeling*

One of the goals in finance is the risk assessment of a portfolio of financial assets. In this study, the objective was to assess the risks of trading soybean commodity at the CME exchange. Risk is often measured in terms of asset price changes.

Suppose Y_t is the commodity price at time t, usually a trading day at CME or the EXG rate quote (Z_t). The price change between t and $t-1$ is given by the difference $\Delta Y_t = Y_t - Y_{t-1} = d_t$, adding it to daily price adjustments.

In finance, relative price variation is widely used, called simple net return ($R_t = \frac{\Delta P_t}{P_{t-1}}$) or log return $r_t = \log \frac{P_t}{P_{t-1}}$. According to Morettin (2008), it is preferable to work with returns, as they are free of scales and have statistical properties that are more interesting (such as stationarity and ergodicity).

The ARMA, ARCH, and GARCH family models are used to model returns. Returns usually have interesting features: returns are self-correlating, return squares are self-correlating, have volatility groupings over time, and the returns distribution have heavier tails than the normal distribution. In addition, distribution is generally leptocurtic, although approximately symmetrical. These facts are known as stylized facts concerning returns (Morettin, 2008).

Therefore, deterministic volatility could be calculated by the ARCH family models, which were introduced by Engle (1982), with the purpose of estimating the inflation variance. The idea is that returns are serially uncorrelated; however, the conditional variance (volatility) depends on returns applied to a quadratic function (Morettin, 2008).

An ARCH model of order m, denoted by ARCH (m) is defined by:

$$r_t = \sqrt{h_t}\varepsilon_t$$
$$h_t = \alpha_0 + \alpha_1 r_{t-1}^2 + \cdots + \alpha_m r_{t-m}^2$$

where ε_t are independent and identically distributed (iid) with zero mean, $\alpha_0 > 0$, $\alpha_m > 0$ and $\alpha_i \geq 0$, $i = 1, \ldots, m-1$.

A drawback of this model is that positive and negative returns are treated similarly because the square of returns is present in the volatility formula. Generally, volatility reacts differently to positive and negative returns. In addition, because the returns are for the square, some large isolated values may lead to overestimated forecasts. Thus, ARCH models tend to overestimate future volatility, incorporating past extreme event outcomes (Morettin, 2008).
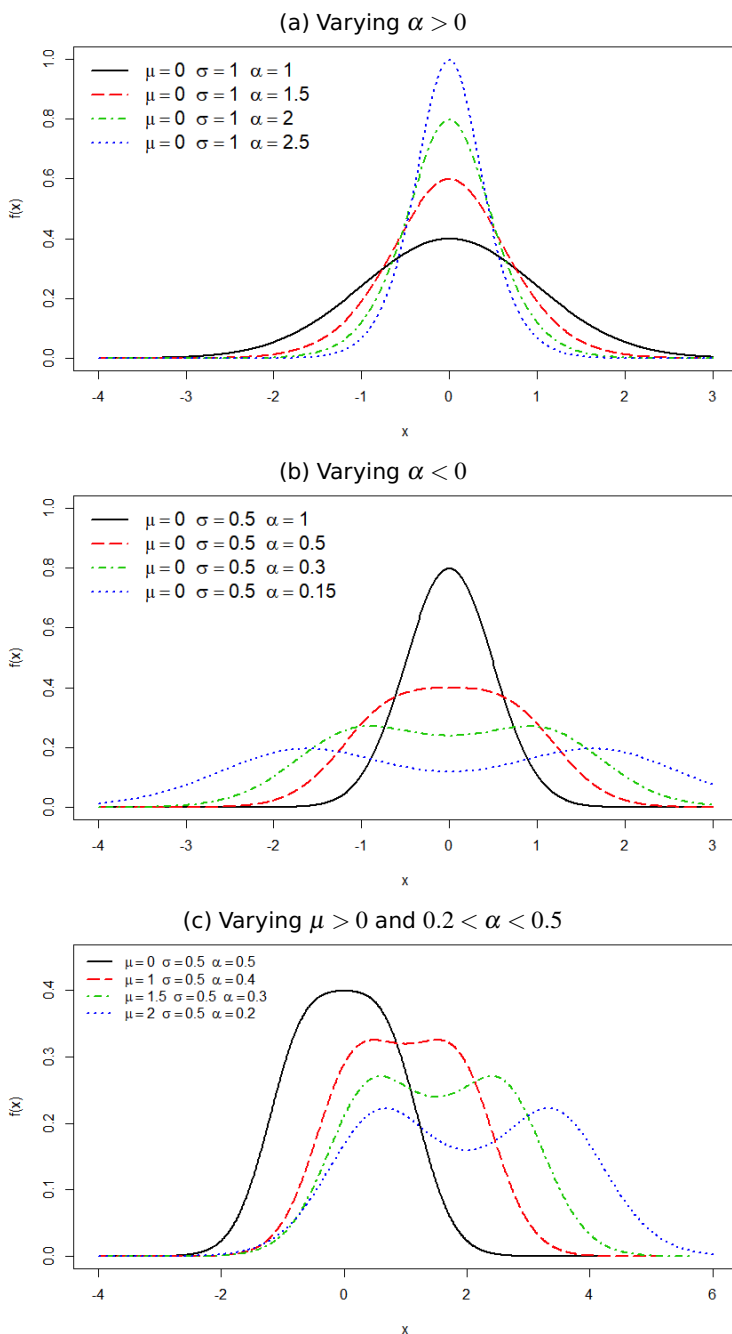
Figure 2. Plots of the OLLN density function for different values of the parameters.

4.2.1 *GARCH Models*   A generalization of the ARCH models was proposed by Bollerslev (1986), called the Generalized ARCH (GARCH) model. These models may be thriftier

than the ARCH models. An GARCH (m,n) model is defined by:

$$r_t = \sqrt{h_t}\,\varepsilon_t$$

$$h_t = \alpha_0 + \sum_{i=1}^{m} \alpha_i r_{t-i}^2 + \sum_{j=1}^{n} \beta_j h_{t-j}$$

where: ε_t are iid with zero mean, $\alpha_0 > 0$, $\alpha_m > 0$, $\alpha_i \geq 0$, $i = 1,\ldots,m-1$, $\beta_j \geq 0$, $j = 1,\ldots,n-1$, $\beta_n > 0$, $\sum_{i=1}^{q}(\alpha_i + \beta_i) < 1$, $q = \max(m,n)$.

It is generally assumed that ε_t errors are usually distributed or follow a t-Student distribution or a generalized error distribution (ged), described in Nelson (1991). According to Morettin (2008), it is recommended in practice to define low order models, for example, GARCH (1,1), (1,2) or (2,1).

4.3 Copulas

According to Joe (2014), the copula is a multivariate distribution in which all marginal one-dimensional distributions are uniform U (0,1). Consequently, if C is a Copula, then the distribution of a vector of dependent random variables is U (0,1). The result below is an extension of Sklar theorem, described in Sklar (1959).

Theorem 1. *For a d-dimensional distribution $F \in \mathfrak{F}(F_1, F_2, \ldots, F_d)$, where F_j is the jth marginal univariate distribution, the Copula associated with F is a distribution function $C : [0,1]^d \to [0,1]$,*

$$F(\mathbf{y}) = C(F_1(y_1), F_2(y_2), \ldots, F_d(y_d)), \ \mathbf{y} \in \mathbb{R}^d, \tag{3}$$

with marginals $U(0,1)$ that satisfy:

(a) *If F is a continuous d-dimensional distribution with uniform margins $F_1, \ldots, F_d$ and quantile functions $F_1^{-1}, \ldots, F_d^{-1}$, then:*

$$C(\mathbf{u}) = F(F_1^{-1}(u_1), \ldots, F_d^{-1}(u_d)), \ \mathbf{u} \in [0,1]^d$$

is unique.

(b) *If F is a d-dimensional distribution of discrete random variables, then the Copula is unique only in the set $Dom(F_1) \times \cdots \times Dom(F_d)$*

The proof of this theorem can be found on page 8 of Joe (2014).

Copula C is parameterized by a δ parameter vector, called the multivariate dependency parameter.

Copula families are generally given as a function of accumulated distribution (for likelihood inference) and Copula density is obtained by differentiation. If $C(\mathbf{u})$ is an absolutely continuous cumulative distribution then its density function is given by:

$$c(\mathbf{u}) = c(u_1, \ldots, u_d) = \frac{\partial^d C(\mathbf{u})}{\partial u_1 \cdots \partial u_d}, \ \mathbf{u} \in [0,1]^d. \tag{4}$$

There are different multidimensional Copulas, including the Gaussian, t-Student and Archimedean families (Cherubini et al., 2004).

4.3.1 *Inference of Multivariate Copulas via Pseudo-Likelihood* In the study of Copula dependence, there are several possibilities for choosing the marginal distribution and the Copula family. For continuous variables, the estimated copula parameter can be compared using the pseudo maximum-likelihood approximation, also known as the canonical maximum likelihood, proposed by Genest et al. (1995).

This is a semi-parametric method using the empirical distributions $\hat{F}_1, \ldots, \hat{F}_d$ and maximizing:

$$L_{pseudo}(\delta) = \sum_{i=1}^{n} \log c(\hat{F}_1(y_{i1}), \ldots, \hat{F}_d(y_{id}); \delta),$$

where δ is the Copula parameter and $\hat{F}_j(y) = n^{-1}\{\sum_{i=1}^{n} \mathbb{I}(y_{ij} \leq y) - 0.5\}$.

According to Joe (2014), the pseudo-maximum likelihood with empirical distribution for marginal functions can only be implemented if the variables are continuous and there are no covariates or censored observations.

If the Copula model is specified correctly, then Genest et al. (1995) shows that δ_{pseudo} is consistent and asymptotically normal.

In addition, Chen and Fan (2006) proved other asymptotic properties when using semi-parametric inference in the time series context and developing estimation methods for one-dimensional copula and marginal functions. Copula selection, when applicable to the pseudo-likelihood approach is commonly performed by comparing AIC Akaike (1974) or BIC Schwarz et al. (1978), based on pseudo-likelihood logs.

In this work, we chose this method for the estimation of Copula parameters.

4.4 *Graphical Tool to Detect Dependency - Chi-plot*

In addition to the scatter plot, the Chi-plot is another graphical tool used in the literature to detect dependency between two variables, proposed by Fisher and Switzer (2001).

The description of this section is based on the works of Genest and Favre (2007) and Tursunalieva and Silvapulle (2007). The Chi-plot graph is based on the data ranks. The Chi-plot is a scatter plot between pairs (λ_i, χ_i) which are transformations of n pairs (X_i, Y_i), $i = 1, \ldots, n$.

Let H be a continuous bivariate distribution function and F_n and G_n the marginal distributions of X and Y, respectively, which can be estimated for each pair (X_i, Y_i) as follows:

$$H_i = \frac{1}{n-1} \sum_{j \neq i} \mathbb{I}(X_j \leq x_i, Y_j \leq y_i) \tag{5}$$

$$F_i = \frac{1}{n-1} \sum_{j \neq i} \mathbb{I}(X_j \leq x_i) \tag{6}$$

$$G_i = \frac{1}{n-1} \sum_{j \neq i} \mathbb{I}(Y_j \leq y_i), \tag{7}$$

where $\mathbb{I}(E)$ denotes the indicator function of the event E.

Fisher and Switzer (2001) proposes to plot the graph of (λ_i, χ_i) pairs, defined by:

$$\chi_i = \frac{H_i F_i G_i}{\sqrt{F_i(1 - F_i)G_i(1 - G_i)}} \tag{8}$$

$$\lambda_i = 4S_i \max\{(F_i - 0.5)^2, (G_i - 0.5)^2\} \tag{9}$$

where $S_i = sign\{(F_i - 0.5)^2, (G_i - 0.5)^2\}$.

The Chi-plot graph has confidence bands (dashed lines on the graph) drawn in $\pm\frac{c_p}{\sqrt{n}}$ (approximate value for c_p and with a significance level of 95% is 1.78).

The pairs (λ_i, χ_i) of independent and continuous marginal factors tend to be located within the bands. For negatively correlated marginal functions, the pairs of (λ_i, χ_i) tend to spread below the lower range. On the other hand, if the marginal functions are positively correlated, the pairs of (λ_i, χ_i) tend to spread above the upper range. Finally, pairs close to zero $\chi_i = 0$ and between ranges indicate independence between variables.

4.5 *Multivariate Independence Test*

For the construction of Copulas, the studied variables $Y_1, Y_2, \ldots, Y_n$ must be a random sample from the joint distribution function $H(Y_1, Y_2, \ldots, Y_n)$. According to Miqueleto (2011), to ensure that the sample is random, the variables must be mutually independent. In other words, if the null hypothesis from which the variables are independent is rejected, this provides evidence that $Y_1, Y_2, \ldots, Y_n$ were not randomly generated. The description of this section is based on the works of Kojadinovic and Yan (2011) and Genest and Rémillard (2004).

It is known that the dependency structure between the variables is completely summarized by the C Copula and the mutual independence between them occurs if, and only if, $C(u_1, u_2, \ldots, u_n) = \Pi_{i=1}^n u_i$, a test of the mutual independence of the components of $Y_i, i = 1, \ldots, n$, can be statistic I_n, as proposed by Genest and Rémillard (2004):

$$I_n = \int_{[0,1]^d} n(C_n(u) - \Pi_{i=1}^n u_i)^2 \, du.$$

Under the hypothesis of independence, $Y_1, Y_2, \ldots, Y_n$ the empirical process $\sqrt{n}(C_n - \Pi_{i=1}^n u_i)$ can be decomposed using the Mobius transform into a collection of $2^d - d - 1$ sub-processes $\sqrt{n}M_A(C_n)$ where $A \subseteq \{1, \ldots, d\}$, $|A| > 1$ converges together to a mutually independent Gaussian process.

Genest and Rémillard (2004) proposes that $Y_1, Y_2, \ldots, Y_n$ mutually independent is equivalent to having $M_A(C(u)) = 0, \forall u \in [0,1]^d$, $A \subseteq \{1, \ldots, d\}$ for $|A| > 1$.

Rather than simple I_n statistics, the authors suggest considering $2^d - d - 1$ test statistics as:

$$M_{A,n} = \int_{[0,1]^d} n(M_A(C_n(u))^d \, du,$$

that are asymptotically independent under the hypothesis of independence.

To view test results, when based on all $2^d - d - 1$ statistics $M_{A,n}$, a graphical representation called a dependogram can be used. For each subset $A \subseteq \{1, \ldots, d\}, |A| > 1$, a vertical bar will be drawn whose height is proportional to the value of $M_{A,n}$.

The critical values of $M_{A,n}$ are represent by black balls.

Subassemblies for which the bar exceeds the critical value can be considered to be dependent variables.

5. Results

In this section, we present the modeling of soybean yield series, CME futures price, and EXG rate quotes. In addition, copula estimates are presented under the two-dimensional and the three-dimensional approaches.

5.1 Yield

Table 1 presents the descriptive statistics for the trend-adjusted yield series for Cascavel and Toledo. We observed an average yield of 3405.587 Kg/ha, standard deviation 329.91, and asymmetry measurement -0.329 for the municipality of Cascavel, indicating a slight left asymmetry (from the normal distribution). In addition, the coefficient of variation in Toledo (15.82%) indicates a higher yield risk than in Cascavel (9.687%).

Table 1. Descriptive Statistics for the yield series adjusted for Cascavel and Toledo (n = 37 observations)

Series	Mean	Median	Std Dev	Asymmetry	Kurtosis	Max.Value	Min.Value	CV(%)
Cascavel	3405.6	3447.4	329.9	−0.3	−0.7	4021.8	2707.0	9.7
Toledo	3290.0	3442.9	520.7	−1.3	1.9	4005.1	1505.8	15.8

The annual yield series for the municipalities of Cascavel and Toledo were modeled using the OLLN distribution as described in Duarte et al. (2017).

Table 2 presents the estimated parameters related to the OLLN distribution for the yield series, with standard errors in parentheses. Note that the parameters are very similar for both municipalities, indicating similar yield characteristics. The result of these estimates allows simulating the required number of observations and their respective cumulative probabilities that are useful for matching price and USD probabilities. Thus, as the series have the same number of observations, the copula estimates can be obtained.

5.2 Bivariate Analysis for Revenue Contract

Consider X the simulated yield for the municipalities for Toledo or Cascavel (kg/ha), Y the price of CBOT/CME (US$ per 60-kg bag) and Z the USD quote.

For the calculation of revenue, let us consider the price $P = Y * Z$, in this case, we do not take into account the exchange variability. The revenue F will be considered as a

Table 2. Estimated parameters of the OLLN(μ,σ,α) distribution simulation, with standard errors in parentheses

Municipalities	μ	σ	α
Toledo	3289.6	637.9	1.04
	(117.1)	(338.4)	(0.68)
Cascavel	3415.3	680.9	1.63
	(72.6)	(797.1)	(2.09)

function of two variables $F = X * P$. This procedure does not take into the dependency structure between price and USD variables. It is equivalent to assume that the correlation between the variables is nonexistent. In this case, we are only assuming that there is a dependency structure between yield and price in Reais.

Table 3 presents the descriptive statistics for the simulated yield series (20,000 observations) of the OLLN distribution for Cascavel, Toledo and price in Reais. An average yield of 3402.604 Kg/ha, standard deviation (Std Dev) 425.01, and asymmetry measurement 0.0015 were observed for Cascavel.

In addition, the statistics are close to each other, suggesting a very similar premium rate for both municipalities. Moreover, the statistics presented in Table 3 are close to the original ones presented in Table 1, demonstrating that the simulations were properly implemented. The average price in Reais is 73,504 per 60-kg bag, with a coefficient of variation of 6.80 %.

Table 3. Descriptive Statistics for simulated yield series for Cascavel, Toledo and price

Series	Mean	Median	Std Dev	Asymmetry	Kurtosis	Max.Value	Min.Value	CV(%)
Cascavel	3402.6	3406.7	425.0	0.002	−0.20	4752.3	2085.2	12.5
Toledo	3279.7	3285.4	593.1	0.001	−0.19	5167.5	1437.1	18.1
Price(R$)	73.5	73.8	5.0	−0.127	−0.76	84.9	62.8	6.8

There is a trend in the soybean futures price series between Apr 5, 2015 to March 3, 2017, presented by the Dickey-Fuller test (p-value = 0.4537). Therefore, we opted for differentiating it, turning it into daily adjustments. Series differentiation does not preclude review of a revenue insurance contract, because the difference in guaranteed price p_g and the price that occurred will be the adjustment.

The graphs in Figures 3a and 3b show the Histogram and Normal QQ-Plot for the adjustment price series in Reais (R$), respectively. Note that the adjusted price series fluctuates around zero, appearing to be stationary, tested by the Dickey-Fuller test (p-value < 0.01, rejecting the hypothesis of unit root); however, with time-dependent variability (volatility).

Moreover, by the normal QQ-plot graph, the adjusted series has a tail heavier than normal. The main concern in the actuarial area is with the left tail of the distribution, as it is the region where dependency between extreme values is observed and is used to calculate the loss risk (loss probability). Underestimating loss probability implies lower premium rates, reducing the financial gains by the insurer.
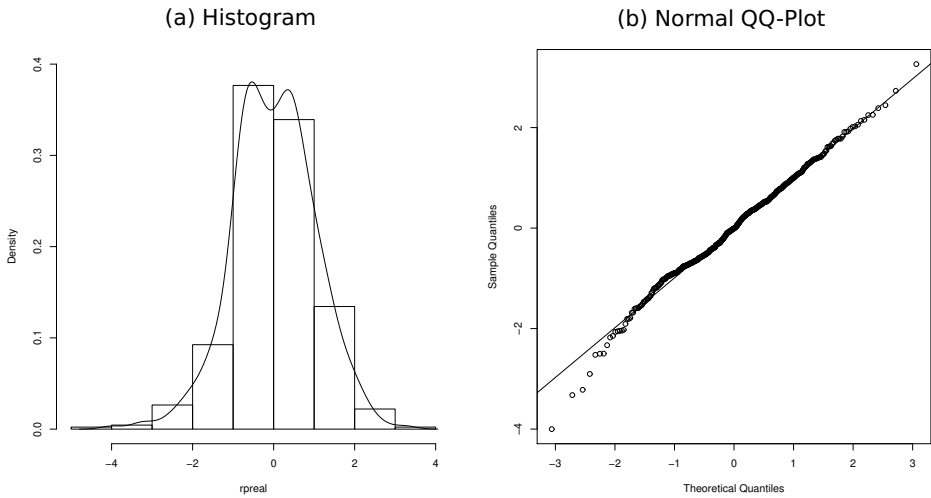
(a) Histogram

(b) Normal QQ-Plot

Figure 3. Original Price in Reais, Adjusted Price (R\$), Histogram and Normal QQ-Plot for the adjustment price series in Reais.

The Ljung-Box test was applied for the adjusted price series (p-value = 0.03963) and it is squared (p-value = 0.00864). There is absence of serial autocorrelation and presence of conditional heteroscedasticity at a significance level of 0.01. Therefore, the modeling of heteroscedasticity was performed using models from the GARCH(m,n) family.

Table 4 presents the values of the AIC and BIC criteria and the maximum log likelihood for the GARCH(m, n) family models with t-Student errors for the adjusted price series. The model chosen was GARCH (1,1) with t-Student errors by the criteria AIC and BIC.

The Ljung-Box test was applied to the residuals of the chosen model, with p-value 0.626727. Considering a significance level 0.01, the test indicates no autocorrelation. On the other hand, the Ljung-Box test result for the squared residuals of the chosen model presented p-value 0.04475, which indicates absence of conditional heteroscedasticity, considering a 0.01 significance level.

Table 4. AIC and BIC criteria and log-likelihood maximum values for models adjusted to the price series in Reais

	Criteria	**GARCH(1,1)**	GARCH(2,1)	GARCH(2,2)
	AIC	2.85	2.86	2.86
Price (R\$)	BIC	2.90	2.91	2.92
	LLV	642.24	642.42	641.74

Therefore, with the simulated yield series of the OLLN distribution for each selected municipality and the residual of the GARCH model (1,1) for the adjusted price, the two-dimensional copula model was adjusted.

Table 5 presents the AIC and BIC values for the adjusted Copulas for the selected municipalities. For both municipalities, the copula that best fits the data structure is the Frank Copula with parameter $\delta = -0.3358$ and standard error (0.289).

Table 5. AIC and BIC Criteria Values for Copula Selection

	Criteria	Frank	AMH	Joe	Gaussiana	t	Gumbel
Cascavel	AIC	0.55	0.66	2.00	1.11	3.11	2.00
	BIC	4.67	4.78	6.12	5.23	11.35	6.12
Toledo	AIC	0.55	0.66	2.00	1.11	3.11	2.00
	BIC	4.67	4.78	6.12	5.23	11.35	6.12

Therefore, the premium rate for Cascavel and Toledo is the same, as these municipalities have the same dependency structure between productivity and price. The discussion of revenue insurance premium rates using the selected Copula is presented in subsection 6.1.

5.3 *Three-dimensional Analysis for Revenue Contract*

In this section, the dependency structure between yield, CME price, and the EXG rate is considered, that is, the variability of the exchange rate in the modeling is taken into account.

Table 6 presents the descriptive statistics for the adjusted price series (US$) and the EXG rate

Table 6. Descriptive Statistics for Adjusted Price Series (US$) and EXG rate (Reais)

Series	Mean	Median	Std Dev	Asymmetry	Kurtosis	Max.Value	Min.Value	CV (%)
Price	21.25	21.15	1.36	0.34	2.04	24.65	19.19	6.40
EXG rate	3.47	3.39	0.31	0.46	1.91	4.19	2.99	9.05

Note that the adjusted price has an average 21.254 (US$ per 60kg bag), standard deviation(Std Dev) 1.361, and coefficient of variation 6.40%. On the other hand, the average EXG rate is 3.47 Reais and standard deviation 0.314. In addition, the coefficient of EXG rate variation is higher than that of the price, respectively, with 9.05% and 6.40%. Indicating that the exchange rate risk is greater than the price risk, which is expected, as the exchange rate is influenced by many other variables than just commodity price.

Figures 4 and 5 show the adjusted price (differential price) for CME and the EXG rate, their histograms, and the normal Q×Q-plot graph, respectively. Both price series fluctuate around zero, appearing to be stationary; however, they have a time-dependent variability (volatility). In addition, the histogram for both series has a higher central part than a normal one, and values are quite far from the central positions of the data. These facts are characteristic of financial returns and are described by the Kurtosis measurement (5.168 and 6.689 for CME and EXG rate, respectively). In addition, the

(a) CME Adjusted Prices (US$)

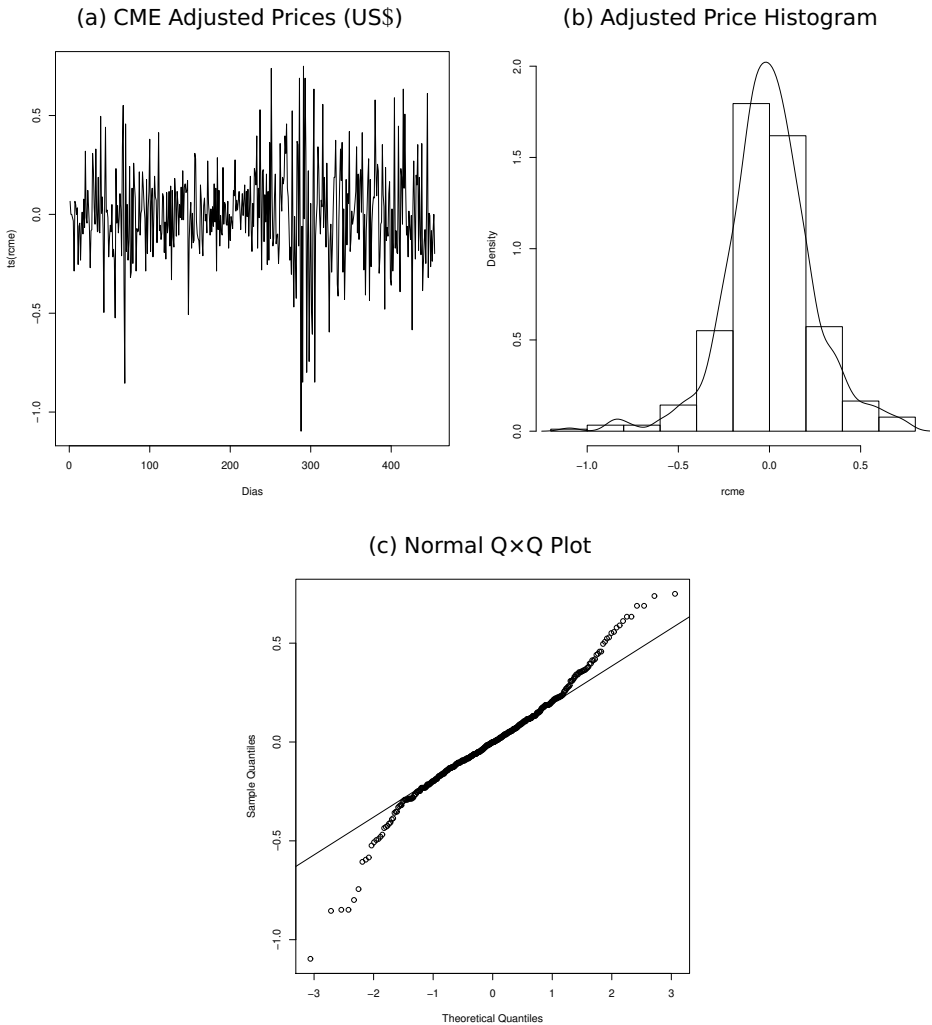### (b) Adjusted Price Histogram

(c) Normal Q×Q Plot

Figure 4. Price, Histogram, and Normal QQ-Plot for the CME Daily Adjustment Price Series.

normal QQ-plot graph shows that the adjusted series have tails heavier than normal thus it is suggested that they are leptocurtic.

The CME price and EXG rate adjusted price series are white noise, that is, they do not present temporal autocorrelation, according to the Ljung-Box test at 0.05 significance level, whose results are presented in Table 7. Therefore, both non-autocorrelated price series are adopted.

The existence of conditional heteroscedasticity was confirmed by the Ljung-Box tests at 0.05 significance level for both squared series (Table 7).

For conditional heteroscedasticity modeling, we used the GARCH models (m, n) with t-Student errors. Table 8 shows the AIC, BIC, and log-likelihood maximum (LLM) values

(a) Adjusted EXG rate
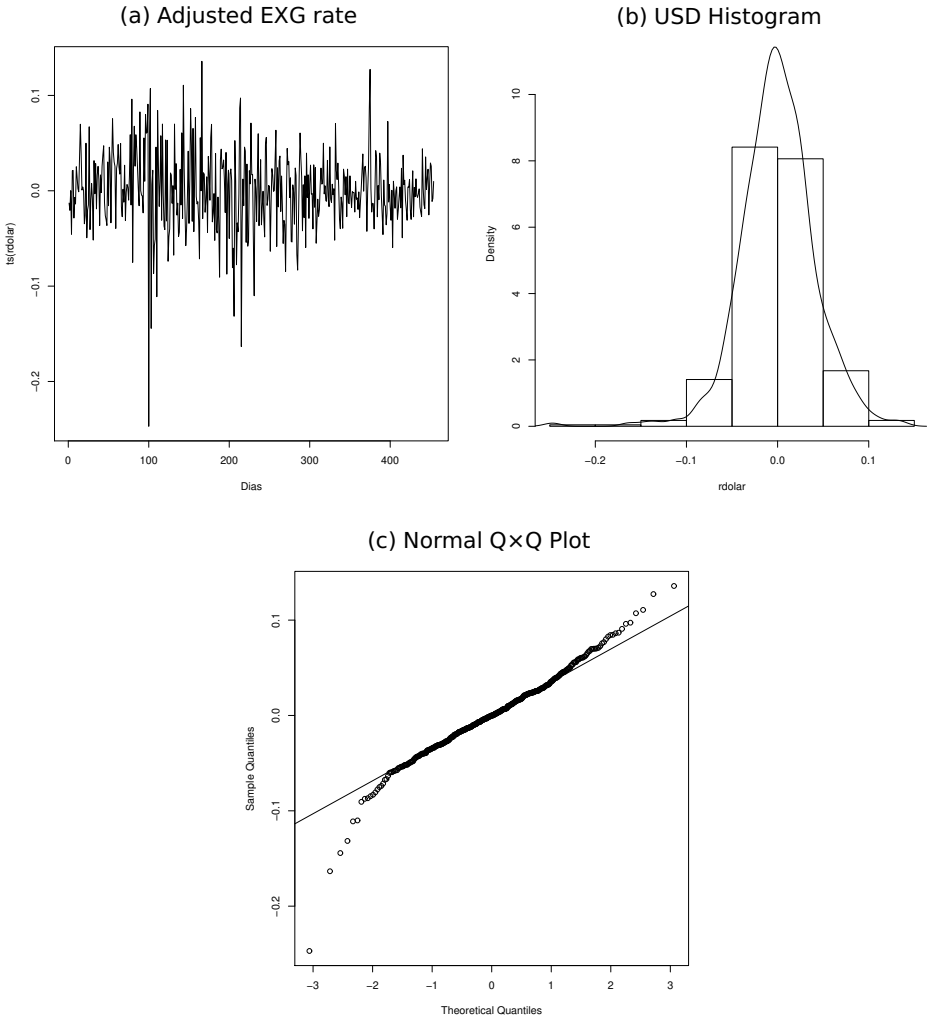
(b) USD Histogram

(c) Normal Q×Q Plot

Figure 5. Price, Histogram, and Normal QQ-Plot for the EXG rate series.

Table 7. Ljung-Box test results for squared CME prices (rcme) and EXG rate (r dolar) adjusted price series.

Series	rcme	rdolar	rcme×rcme	rdolar×rdolar
p-valor	0.54	0.37	0.05	1.31×10^{-4}

for the CME and dollar price-adjusted models. By the BIC criterion, the lowest value for both series is for GARCH (1,1) with t-Student errors.

The Ljung-Box tests for GARCH model residuals (1,1) with t-Student errors for CME price and USD quote presented p-value=0.5505 and p-value=0.7138, respectively. The Ljung-Box test results for the squared residuals of the model presented p-value=0.4886 and p-value=0.2777 for CME and EXG rate, respectively. These results indicate that the

model assumptions are verified, that is, absence of temporal autocorrelation and conditional heteroscedasticity in the residuals, respectively. Therefore, the GARCH models (1,1) with t-Student errors were selected for adjusted CME prices and EXG rates.

Table 8. Values of the AIC and BIC criteria and maximum log-likelihood(LLM) for GARCH models (m, n) with t-Student errors to the CME prices and EXG rate series

	Criteria	**GARCH(1,1)**	GARCH(2,1)	GARCH(2,2)
CME price	AIC	−0.18	−0.18	−0.18
	BIC	−0.13	−0.13	−0.12
	LLM	−45.42	−46.99	−47.95
EXG rate	AIC	−3.74	−3.73	−3.74
	BIC	−3.69	−3.68	−3.68
	LLM	−853.24	−853.13	−855.90

Figure 6 presents Chi-Plot for the pairs of variables (Toledo yield, CME price), (Toledo Yield, EXG rate), and (CME price, EXG rate). The graphs suggest a negative dependence between the pair of variables (CME price, EXG rate) and independence between the pairs (Yield, CME price) and (Yield,EXG rate). Thus, the CME price variable may be negatively correlated with the EXG rate quote, which may reduce the impact of price risk on insurance pricing.

Figure 7 presents the result of the empirical copula independence test of random variables proposed by Genest and Rémillard (2004). The heights of the bars represent the values of the test statistics by subset of variables and the black ball represents the critical value of the test. In addition, the first bar represents the test result for the set of variables 1,2, where 1 represents the variable yield, 2 the CME price, and 3 the EXG rate series.

At 0.05 of significance level, the subset variables (1,2) and (1,3) can be considered independent. On the other hand, the subset of variables (2,3) is considered dependent and variables 1,2 and 3 are considered mutually independent. Therefore, the yield, price, and EXG rate variables were randomly generated and are considered independent.

Table 9 presents the values of AIC and BIC criteria and the maximum log-likelihood (MLL) for the different copulas for the yield variables, adjusted CME, and EXG rate price in Toledo and Cascavel. According to all selection criteria, the t-Student Copula best represents the data dependency structure.

Table 10 presents the selected models with the selected parameter vector of the trio of variables. Note that for both municipalities, the selected copula was t-Student with parameters $(\rho_1, \rho_2, \mathrm{df}) = (0.5458; 0.0896; 1.84)$. The inference was based on the pseudo-maximum likelihood with symmetric positive matrix structure characterized by toeplitz (dispstr = toeplitz).

Therefore, the revenue insurance premium rate for both municipalities is the same, as the copula parameters are the same. This was expected, since we are working with the same price and EXG rate series, changing only the simulated yield (with very similar parameters between both municipalities).
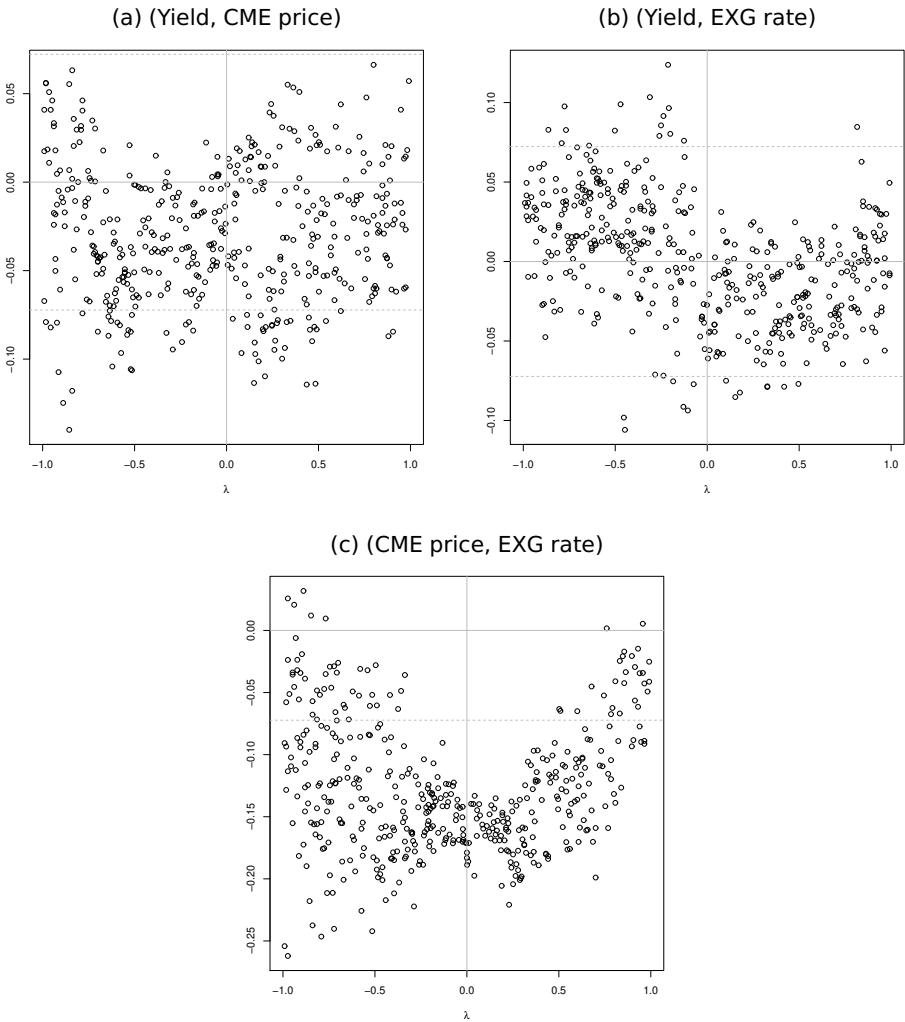
(a) (Yield, CME price)

(b) (Yield, EXG rate)

(c) (CME price, EXG rate)

Figure 6. Chi-Plot for variable pairs (Yield, CME price)

Table 9. AIC, BIC, and maximum log-likelihood (MLL) criteria for Copula Selection.

	Criteria	Clayton	Frank	Gumbel	Joe	Gaussian	t
	AIC	−116.40	−115.66	−101.40	−67.12	−106.27	−450.55
Toledo	BIC	−112.28	−111.55	−97.28	−63.00	−102.15	−438.20
	MLL	59.20	58.83	51.70	34.56	54.14	228.28
	AIC	−116.40	−115.66	−101.40	−67.12	−106.27	−450.55
Cascavel	BIC	−112.28	−111.55	−97.28	−63.00	−102.15	−438.20
	MLL	59.20	58.83	51.70	34.56	54.14	228.28

The fact that the price and EXG rate series are the same in both municipalities makes yield important in differentiating the rates of municipalities. The risk rate cannot be
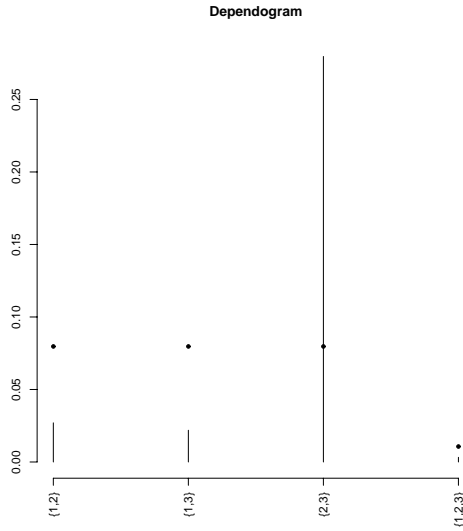
Figure 7. Dependogram.

viewed solely as a direct proportionality of risk. For example, the standard deviation of simulated yield in Toledo is approximately 39% higher than that of Cascavel; neverthe-less, it cannot be said that the Toledo rate is 39% higher than that of Cascavel.

Table 10. Selected models for revenue insurance pricing with dependency parameters.

Data	Copula	Parameters
(Toledo,CME price,EXG rate)	t-Student	$(\rho_1,\rho_2,\mathrm{df}) = (0.5458; 0.0896; 1.84)$
(Cascavel,CME price,EXG rate)	t-Student	$(\rho_1,\rho_2,\mathrm{df}) = (0.5458; 0.0896; 1.84)$

6. Discussion

6.1 *Two-dimensional Revenue Agreement*

Table 11 presents the pure rate and the commercial rate (with 20% loading) for the bivariate case for Frank Copula ($\delta = -0.3358$), with N = number of observations.

For N = 454, we used the simulated yield series of the OLLN distribution with N = 454 observations and the adjusted price series in Reais, where (Y_e) was calculated by the arithmetic average of the last 15 days to from Aug 8, 2016 (date of insurance contracting).

On the other hand, for N = 20,000, we considered 20,000 simulated observations of adjusted Copula of adjusted series. In addition, for the expected price Y_e, we used the average of the Copula simulated price vector with 20,000 observations.

Note that the simulated pure rate (for N = 20,000) of the coverage level (CL) in-surance premium of 60% and 70% corresponds to 4.682% and 7.115%, respectively.

Considering the commercial EXG rate for the coverage level of 60% and 70% is 9.365% and 14.230% respectively. Note that the higher the level of coverage, the higher the insurance premium rate.

Table 11. Pure and Commercial Rate in % for the two-dimensional revenue insurance premium for the municipalities of Toledo and Cascavel using the Frank copula ($\delta = -0.3358$)

N		60%	65%	70%	75%	80%
454	Pure rate	4.50	5.39	6.45	7.65	8.95
	Commercial Rate	9.00	10.77	12.89	15.31	17.90
20 000	Pure rate	4.68	5.84	7.12	8.48	9.97
	Commercial Rate	9.37	11.68	14.23	16.97	19.93

Table 12 presents the average commercial revenue insurance premium rate using the selected copula with N = 454 observations and insurer A rates for coverage levels(CL) $60 - 69\%$ and $70 - 79\%$. Note that the rate calculated in this study is well above those applied by the insurance market. For example, insurer A rate is 63.32% and 40.28% of the rate calculated by the Copula for CL of 60-69% and 70-79%, respectively.

Therefore, there is evidence of an underestimation by insurer A, which may lead to a large loss for the insurer, as it may be considering a much lower risk.

Table 12. Average commercial rate for revenue insurance using copula and insurer A rate for Toledo and in parenthesis for Cascavel

Model	$60 - 69\%$	$70 - 79\%$
Copula	9.89	14.10
insurer A	6.26	5.68 (5.52)

6.2 *Three-dimensional revenue agreement*

From a three-dimensional approach, the selected Copula for both municipalities is the t-Student with parameters $(\rho_1, \rho_2, \mathrm{df}) = (0.5458; 0.0896; 1.84)$. Table 13 presents the pure premium revenue insurance rate and the commercial rate in % with 20% charge rate and coverage level 60%, 65%, 70%, 75% and 80% for N = 454 (adjusted prices), and N = 20,000 simulated observations of the selected Copula.

Note that for CL = 65% and CL = 80%, the pure premium rate is 1.360% and 3.230%, respectively.

Comparing Tables 11 and 13, for all levels of coverage, the rate calculated by the bivariate Copula is overestimated when compared to the three-dimensional Copula.

Moreover, when considering exchange as a random variable, the rate calculated by the three-dimensional approach (CL = 80%) is almost threefold lower than that calculated under the two-dimensional approach, in which EXG rate variability is not taken into account.

Therefore, when the EXG rate variable is added to the modeling, there is a decrease in the revenue insurance premium rate. This decrease may be related to the incorporation of the negative correlation between price (US$) and the EXG rate into the model, which may be leading to risk neutralization between these variables. In other words, the negative dependence that exists between prices and the EXG rate may be nullifying the risk of these variables in modeling.

Table 13. Pure and Commercial Charging Rates (20%) for revenue Insurance using t-Student copula for Toledo (Yield, CME Price, EXG rate)

		N	60%	65%	70%	75%	80%
N=454	Pure rate		0.955	1.410	1.937	2.589	3.378
	Commercial rate		1.911	2.820	3.874	5.178	6.755
N=20000	Pure rate		0.947	1.360	1.869	2.486	3.230
	Commercial rate		1.894	2.719	3.738	4.973	6.459

Table 14 presents the commercial rate for coverage level $60-69\%$ and $70-79\%$ calculated by the three-dimensional copula and insurer A. For both coverage level ranges, insurer A is overestimating the insurance rate when compared to the copula methodology.

Rate overestimation can hamper securing insurance throughout the Brazilian territory as well as attracting producers with greater risk of compensation, thus increasing the problem of adverse selection.

Table 14. Commercial Average Rates for Revenue Insurance using t-Student Copula and Insurer A.

CL	Copula	Insurer A
$60-69\%$	2.36	6.26
$70-79\%$	4.52	5.68

7. Conclusion

Only two insurance companies offer agricultural revenue insurance in the Brazilian insurance market. Insurance company A takes into account the series of municipal annual yields and the physical price (CEPEA/ESALQ indicator) in pricing its municipality. On the other hand, insurer B differs by using the futures series negotiated at CME, already converted into Reais.

Therefore, this study proposes alternative methods for the calculation of the premium insurance premium rate. The Copula method was analyzed under two approaches: two-dimensional and three-dimensional. The two-dimensional approach takes

into account yield and price risk in Reais. On the other hand, the three-dimensional approach takes into account yield, CME price, and exchange rate variables.Moreover, understanding the relationships among these three variables is the main contribution of the study.

This study used the annual soybean yield series for the municipalities of Toledo and Cascavel, the soybean prices traded at CBOT/CME due March 2017, and the USD exchange rate during the same period. Yield was simulated from the OLLN distribution, price and USD series were adjusted by GARCH (1,1) models with t-Student errors for volatility modeling.

From the two-dimensional approach, the best-fit dependency structure was the Frank Copula with the negative dependency parameter ($\delta = -0.3358$). The results suggest that the insurance market underestimates the insurance premium rate when compared to the copula methodology. Underestimating the rate leads the insurer to risk loss, which may lead to financial losses. In addition, insurers take into account a lower risk than should be taken into account when pricing the product.

From the three-dimensional approach, the best-fitting copula was t-Student with parameters $(\rho_1,\rho_2,\mathrm{df}) = (0.5458; 0.0896; 1.84)$. The rate calculated under this approach is lower than the one presented under the two-dimensional approach, because in this case, the random USD rate variable in the modeling was taken into account. Thus, there is the incorporation of an exogenous variable into the modeling, which has a great influence on soybean prices in the Brazilian context. This decrease may be related to the incorporation of negative dependency between price (US$) and the USD rate into the modeling, which may lead to risk neutralization between these variables.

In addition, the results suggest that rates applied by insurers may be overpriced when compared to the three-dimensional copulation model. This overpricing can attract high-risk producers, increasing the problem of adverse selection and making it difficult to massify the crop insurance among farmers.

Bibliography

Akaike, Hirotugu (1974): "A new look at the statistical model identification," *Automatic Control, IEEE Transactions on*, 19 (6), 716–723. [10]

Azzalini, Adelchi (1985): "A class of distributions which includes the normal ones," *Scandinavian journal of statistics*, 171–178. [5]

Azzalini, Adelchi and Antonella Capitanio (2003): "Distributions generated by perturbation of symmetry with emphasis on a multivariate skew t-distribution," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 65 (2), 367–389. [5]

Barchart (2017): "Advanced Commodities Service," Available in: http://acs.barchart.com (accessed in 05/02/2017). [3]

Bollerslev, Tim (1986): "Generalized autoregressive conditional heteroskedasticity," *Journal of econometrics*, 31 (3), 307–327. [8]

Botts, Ralph R and James N Boles (1958): "Use of normal-curve theory in crop insurance ratemaking," *Journal of Farm Economics*, 40 (3), 733–740. [5]

Brisolara, Cláudio Silveira (2013): "Proposicões para o desenvolvimento do seguro de receita agrícola no Brasil: do modelo teórico ao cálculo das taxas de prêmio," Ph.D. thesis, Universidade de São Paulo. [3]

Chen, Xiaohong and Yanqin Fan (2006): "Estimation of copula-based semiparametric time series models," *Journal of Econometrics*, 130 (2), 307–335. [10]

Cherubini, Umberto, Elisa Luciano, and Walter Vecchiato (2004): *Copula methods in finance*, John Wiley & Sons. [9]

Duarte, Gislaine V, Altemir Braga, Daniel L Miquelluti, and Vitor A Ozaki (2017): "Modeling of soybean yield using symmetric, asymmetric and bimodal distributions: implications for crop insurance," *Journal of Applied Statistics*, 1–18. [1, 2, 5, 12]

Duarte, Gislaine Vieira and Vitor Augusto Ozaki (2019): "Pricing Crop Revenue Insurance Using Parametric Copulas," *Revista Brasileira de Economia*, 73 (3), 325–343. [2]

Engle, Robert F (1982): "Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation," *Econometrica: Journal of the Econometric Society*, 987–1007. [7]

Fisher, NI and P Switzer (2001): "Graphical assessment of dependence: Is a picture worth 100 tests?" *The American Statistician*, 55 (3), 233–239. [10]

Gallagher, Paul (1987): "US soybean yields: estimation and forecasting with nonsymmetric disturbances," *American Journal of Agricultural Economics*, 69 (4), 796–803. [4, 5]

Genest, Christian and Anne-Catherine Favre (2007): "Everything you always wanted to know about copula modeling but were afraid to ask," *Journal of hydrologic engineering*, 12 (4), 347–368. [10]

Genest, Christian, Kilani Ghoudi, and L-P Rivest (1995): "A semiparametric estimation procedure of dependence parameters in multivariate families of distributions," *Biometrika*, 82 (3), 543–552. [10]

Genest, Christian and Bruno Rémillard (2004): "Test of independence and randomness based on the empirical copula process," *Test*, 13 (2), 335–369. [11, 18]

Ipardes (2017): "Instituto Paranaense de Desenvolvimento Econômico e Social. Base de Dados do Estado (BDEweb)," Accessed in: 20/dezembro/2017. [3]

Ipea (2017): "Câmbio - Séries Históricas- IDEADATA," Avaliable in: http://www.ipeadata.gov.br/Default.aspx (Accessed in: 05/02/2017). [4]

Joe, Harry (2014): *Dependence modeling with copulas*, CRC Press. [9, 10]

Just, Richard E and Quinn Weninger (1999): "Are crop yields normally distributed?" *American Journal of Agricultural Economics*, 81 (2), 287–304. [5]

Kojadinovic, Ivan and Jun Yan (2011): "Tests of serial independence for continuous multivariate time series based on a Möbius decomposition of the independence empirical copula process," *Annals of the Institute of Statistical Mathematics*, 63 (2), 347–373. [11]

MAPA (2017): "Atlas do Seguro Rural. Indicador das taxas." Accessed in 21 of february-2017. [4]

Miqueleto, G. J. (2011): "Contribuições para o desenvolvimento do seguro agrícola de renda para o Brasil: evidências teóricas e empíricas." Ph.D. thesis, Escola Superior de Agricultura "Luiz de Queiroz " -Universidade de São Paulo. [3, 11]

Morettin, Pedro Alberto (2008): *Econometria financeira: um curso em séries temporais financeiras*, Edgard Blücher. [7, 9]

Nelson, Daniel B (1991): "Conditional heteroskedasticity in asset returns: A new approach," *Econometrica: Journal of the Econometric Society*, 347–370. [9]

Ozaki, Vitor A and Ralph S Silva (2009): "Bayesian ratemaking procedure of crop insurance contracts with skewed distribution," *Journal of Applied Statistics*, 36 (4), 443–452. [5]

Schwarz, Gideon et al. (1978): "Estimating the dimension of a model," *The annals of statistics*, 6 (2), 461–464. [10]

Sklar, M (1959): *Fonctions de répartition à n dimensions et leurs marges*, Université Paris 8. [9]

Tejeda, Hernan A and Barry K Goodwin (2008): "Modeling Crop prices through a Burr distribution and Analysis of Correlation between Crop Prices and Yields using a Copula method." Tech. Rep. [4, 5]

Tursunalieva, Ainura and Param Silvapulle (2007): "Assessing and modeling the changes in dependence between exchange rates," *Working Paper of Monash University*. [10]