Vine-copula Based Models for Farmland Portfolio Management

Xiaoguang Feng
Graduate Student
Department of Economics
Iowa State University
xgfeng@iastate.edu

Dermot J. Hayes
Pioneer Chair of Agribusiness
Professor of Economics
Professor of Finance
Iowa State University
dhayes@iastate.edu

Selected Paper prepared for presentation at the 2016 Agricultural & Applied Economics Association Annual Meeting, Boston, Massachusetts, July 31-August 2

Introduction

U.S. farmland has achieved total returns of 10%-13% over the past decade with volatility of only 4%-5% (NCREIF Farmland Index). In addition, farmland returns have had low or negative correlation with traditional asset classes. These characteristics make farmland an attractive asset class for investors. Farmland, as a real asset, can also provide a hedge against inflation because farmland returns exhibit positive correlation with inflation. Over the past decade, annual U.S. farmland total return exceeds U.S. inflation rate by 3.55% (NCREIF Farmland Index and Consumer Price Index - Urban). With growing global demand for agricultural commodities and limited land to expand capacity, some investors expect that farmland will continue to generate superior returns for the foreseeable future.

Efficient risk management and portfolio management are critical to create optimal risk/return profile for all investments. An essential issue in portfolio risk management is how marginal time series and the correlation structure of a large number of asset returns are treated. Most previous studies on farmland portfolio analysis were performed under the Capital Asset Pricing Model (CAPM) framework (Barry, 1980; Hennings, Sherrick, and Barry, 2005; Noland, Norvell, Paulson, and Schnitkey, 2011). The linear correlation assumption implied by the CAPM, however, is not adequate to capture complex correlation structure such as tail dependence and asymmetry that potentially exist among farmland asset returns. In addition, the normality assumption of the CAPM for asset returns has proven to be inappropriate in agriculture (Just and Weninger, 1999). Copula modeling is a suitable alternative. Margins and dependence can be separated by the copula function. The choice of marginal distribution is arbitrary and various copula types exhibiting flexible and complex correlation structures are available. Chen, Wilson, Larsen, and Dahl (2014) used the

Gaussian copula to model joint distribution of agricultural asset returns to account for non-normal margins. However, the Gaussian copula can only capture symmetric correlation structure and allows no tail-dependence. Besides, the Gaussian copula, restrictions exist for most other multivariate copulas (Student's t copula, Archimedean copulas, etc.). This inflexibility issue can be overcome by the pair-copula modeling proposed by Joe (1996). In particular, the regular vine (R-vines) representation of pair-wise copulas specifies arbitrary bivariate copulas as building blocks and hence can model more complicated correlation structure.

This study applies vine copulas to model farmland asset returns. We focus on annual state-level cropland returns for 24 major U.S. agricultural producing states. This data set covers the period spanning from 1967 to 2014. Following Brechmann and Czado (2013), ARMA models with appropriate error distribution are fitted to each return. R-vine copulas are then used to model the correlation structure of standardized residuals obtained from the marginal time series models. Given the high dimensionality of the vine copula modeling, a sequential maximum likelihood method is applied to specify the R-vine structures and estimate the parameters. The vine model mitigates the curse of dimensionality and facilitates interpretation of the correlation structure. This model loosens the restrictive normality and linearity assumptions under the classical CAPM framework, and allows for complex and flexible correlation structure such as tail-dependence. We compare this model to relevant benchmark models using the Gaussian and t copulas. The results show that the vine-copula based model provides a better a fit as indicated by modeling-fitting criteria. We show that, farmland portfolio management can benefit in terms of forecasting tail risk (Value-at-Risk) and constructing optimal portfolio more accurately for both passive and active portfolio management. Our results show that the model provides an approach to precisely assessing and allocating risk of the farmland portfolio under

the modern risk management framework.

Empirical Framework

The copula was first introduced by Sklar (1959). Sklar's theorem states that if F is an arbitrary k-dimensional joint continuous distribution function, then the associated copula is unique and defined as a continuous function $C : [0,1]^k \to [0,1]$ that satisfies the equation

$$(1) \quad F(x_1, \ldots, x_k) = C\left[F_1(x_1), \ldots, F_k(x_k)\right], \qquad x_1, \ldots, x_k \in \mathbb{R},$$

where $F_1(x_1), \ldots, F_k(x_k)$ are the respective marginal distributions.

In this way, the joint distribution of $x_1, \ldots, x_k$ can be described by the marginal distributions F_i and the correlation structure captured by the copula C. Note that the copula function is flexible in the sense that the variables x_i can be modeled with any kind of marginal distributions. In turn, if the marginal distributions are continuous, a unique copula exists corresponding to the joint distribution. That is,

$$(2) \quad C(u_1, \ldots, u_k) = F\left[F_1^{-1}(u_1), \ldots, F_k^{-1}(u_k)\right], \qquad u_1, \ldots, u_k \in [0,1],$$

where $F_1^{-1}(\cdot), \ldots, F_k^{-1}(\cdot)$ are the corresponding quantile functions. Therefore, the copula can be defined as an arbitrary multivariate distribution on $[0,1]^k$ with all marginal distributions being uniform.

Let c denote the density function of the copula C, which can be described as

$$(3) \quad c(u_1, \ldots, u_k) = \frac{\partial^k C(u_1, \ldots, u_k)}{\partial u_1 \cdots \partial u_k},$$

The corresponding joint density function of $x_1, \ldots, x_k$ can then be written as

$$(4) \quad f(x_1, \ldots, x_k) = c\left[F_1(x_1), \ldots, F_k(x_k)\right] \prod_{i=1}^{k} f_i(x_i).$$

where $f_1(x_1), \ldots, f_k(x_k)$ are marginal density functions.

Basic copula families are generally composed of parametric and nonparametric

copulas. Empirical studies typically use parametric copulas because of their superiority in simulation. There are a large number of different parametric copula families. The most frequently used are elliptical copulas and Archimedean copulas. Despite the effectiveness of basic copulas for modeling low dimentional such as pair-wise correlation, they have strict restrictions in terms of the correlation structure. For example, elliptical copulas imply symmetric correlation structure in the tails. Archimedean copulas, while allowing for asymmetric tail dependence, imply symmetry of the permutation of variables and represent multivariate correlation structure with only one single parameter.

Vine copulas, introduced by Aas et al. (2009), overcome the restrictions imposed by basic copulas and exploit the usefulness of basic copulas in bivariate case as well. For a set of k random variables with density function $f(x_1, x_2, \ldots, x_k)$, it holds that

$$f(x_1, \ldots, x_k) = f_k(x_k) f(x_{k-1} \mid x_k) f(x_{k-2} \mid x_{k-1}, x_k) \ldots f(x_1 \mid x_2, \ldots, x_k). \tag{5}$$

Joe (1996) shows that each of the components in equation (5) can be decomposed into the product of a pair-wise copula and a conditional marginal density:

$$f(x \mid v) = c_{x, v_k \mid v_{-k}}(F(x \mid v_{-k}), F(v_k \mid v_{-k})) f(x \mid v_{-k}). \tag{6}$$

Following this composition, the joint density $f(x_1, x_2, \ldots, x_k)$ can be represented in terms of only pair-wise copulas. In the case of three random variables, for example, the density can be written as

$$(7) \quad f(x_1, x_2, x_3) = f_1(x_1) f_2(x_2) f_3(x_3) c_{1,2}(F_1(x_1), F_2(x_2)) \\ c_{1,3}(F_1(x_1), F_3(x_3)) c_{2,3|1}(F(x_2 \mid x_1), F(x_3 \mid x_1)).$$

The pair-wise copulas $C_{1,2}$,$C_{1,3}$,and $C_{2,3|1}$ are chosen independently so that a wide range of correlation structures can be modeled. The constrution can be generalized in essentially the same way for correlation structures with higher dimensions. Vines are used to represent this pair-wise copula construction graphically. Kurowicka and Cooke (2006) show that a regular vine (R-vine) on k random variables consists of a sequence of linked trees $T_1, \ldots, T_{k-1}$. The copula density function is uniquely determined by

$$(8) \quad c(F_1(x_1), \ldots, F_k(x_k)) = \prod_{i=1}^{k-1} \prod_{e \in E_i} c_{j(e),k(e)|D(e)}(F(x_{j(e)} \mid x_{D(e)}), F(x_{k(e)} \mid x_{D(e)})),$$

where each edge $e = j(e), k(e) \mid D(e)$ in E_i is associated with a bivariate copula density $c_{j(e),k(e)|D(e)}$ and $x_{D(e)}$ represents the subvector of $(x_1, x_2, \ldots, x_k)$ indicated by the indices contained in $D(e)$. Figure 1 shows an example of the trees in the case of three variables.

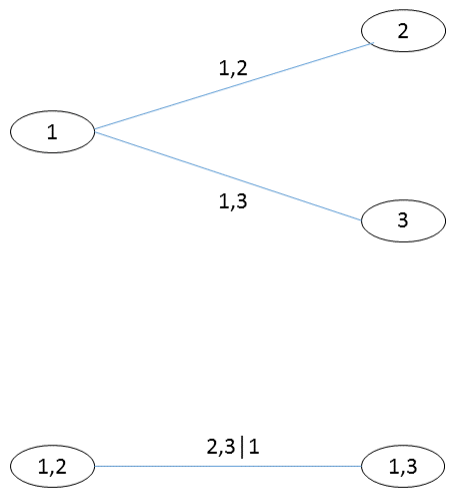

Figure 1. Three-dimensional R-vine trees

Many different combinations of pair-wise copulas are possible for a vine copula specification. In the case of k random variables, there are $\frac{k!}{2} \cdot \frac{(k-2)!}{2!}$ different R-vines. Following Aas et al. (2009), a heuristic method is used to specify the R-vine trees. This approach, while captures the strongest dependencies in the lowest level trees, avoids numerical errors in higher level tress as well (Joe et al., 2010). For the lowest level tree in a R-vine, we select a tree on all nodes that leads to the maximum for the sum of pairwise dependencies. Kendall's τ is used as a measure of association between the dependency and the copula parameter since it is indifferent to nonlinear transformation. Therefore, the lowest level tree is selected by solving the following optimization problem

$$(9) \quad max \sum_{edges\ e=\{i,j\}\ in\ the\ tree} \mid \tau_{ij} \mid,$$

where τ_{ij} is the Kendall's τ associated with the edge $e = \{i, j\}$ in the tree. Given the specified tree, the pair-wise copulas are selected from a range of copula families using the Akaike information criterion (AIC). This procedure is iterated sequentially for higher level trees until the whole vine structure has been determined. Commonly used Gaussian and t copulas are selected as benchmarks for model-fitting comparison.

Application

Our data set consists of state-level average cash rents and land values in 24 US states spanning from 1967 to 2014. All the data are taken from the USDA databases. Cash rents are used as an approximation for the net income of land assets. Annual land asset return is calculated as the sum of income return and capital appreciation for each of the states. This creates a 24-dimentional time series data set of land asset returns.

We investigate the correlation structure of the land asset returns using R-vine copulas. A commonly used two-step procedure is adopted to estimate the parameters of the copula model. The method is called inference for margins (IFM) (Joe and Xu, 1996). Individual land asset returns are first modeled by univariate time series models. ARMA(1,1), AR(1), MA(1), and white noise models with Student's t error distribution are first used to account for potentially heavy tails. The standardized residuals are tested using Kolmogorov-Smirnov goodness-of-fit test and the model with the highest p-value is selected if the p-value is greater than 5%. If the p-value is less than 5%, we stepwisely increase the terms in the ARMA model until the p-value

from the respective Kolmogorov-Smirnov test on the standardized residuals is greater than 5%. Also, the degree of freedom of the Student's t error distribution is greater than ten, a normal distribution is used if the corresponding p-value is greater than 5%.

The standardized residuals are then modeled by the R-vine copula using the maximum likelihood method. The vine copula model is compared with the Gaussian copula and Student's t copula in terms of model fit. The Gaussian copula and Student's t copula are also estimated using the maximum likehood method. Table 1 shows the results for the three alternative models. It is obvious that the vine copula model provides a superior fit than the other two benchmark models, indicating the correlation structure among land asset returns is more complicated than what the standard models imply.

Table 1. Goodness-of-fit Statistics of Alternative Copula Models for Farmland Asset Returns

	Log-lik.	# of Par.	AIC	BIC
R-vine	1204.61	443	-1523.22	-694.28
Gaussian	647.88	276	-743.76	-227.31
Student's t	658.40	277	-764.80	-248.34

Optimal Portfolio Construction

With the estimated R-vine copula model and marginal time series of the farmland asset returns, the optimal portfolio is constructed by the following procedure:

(1) A sample of the standardized residuals is simulated from the R-vine copula model.

(2) Forward looking asset returns are projected using the simulated residuals and the estimated marginal time series.

(3) Portfolio return is the weighted average of individual asset return in the portfolio.

The portfolio return is maximized with respect to a risk measure (standard deviation, value-at-risk, etc.)

Table 2 reports the risk/return characteristics of the minimum variance portfolio according to the R-vine copula model as well as tobtained from the mean-variance optimization. The results show that the both the expected return and standard deviation are higher using mean-variance optimization based on historical data than those obtained from copula models with a forward-looking point of view. This might indicate the farmland market is going to experience downward trends in the near future. Investors therefore should be alerted on the declining rate of return from this alternative asset type. The R-vine copula model identifies less risky portfolio with a higher expected return compared to the Gaussian or Student's t copula model. This shows the superiority of R-vine copulas in modeling potentially complicated correlation structure among farmland asset returns.

Table 2. Risk/return Profile of the Minimum Variance Portfolio Based on Alternative Models

	Expected Return	Standard Deviation
R-vine	8.09%	4.75%
Gaussian	7.57%	4.92%
Student's t	7.64%	5.05%
Mean-variance	10.20%	6.60%

Conclusions

The vine-copula based model used in this study can serve as an initiative for more elaborate models for farmland portfolio management. One direction for future research would be to explore dynamic vine-copula structures to take into account the dynamics of correlations among farmland asset returns for forward-looking portfo-

lio management. Another direction could be the consideration of estimation risk to account for the uncertainty of correlation parameters in the vine-copula model.

References

Aas, K., Czado, C., Frigessi, A. and Bakken, H., 2009. Pair-copula constructions of multiple dependence. Insurance: Mathematics and economics, 44(2), pp.182-198.

Barry, P.J., 1980. Capital asset pricing and farm real estate. American Journal of Agricultural Economics, 62(3), pp.549-553.

Brechmann, E.C. and Czado, C., 2013. Risk management with high-dimensional vine copulas: An analysis of the Euro Stoxx 50. Statistics & Risk Modeling, 30(4), pp.307-342.

Chen, S., Wilson, W.W., Larsen, R. and Dahl, B., 2014. Investing in Agriculture as an Asset Class. Agribusiness.

Hennings, E., Sherrick, B.J. and Barry, P.J., 2005. Portfolio diversification using farmland investments. Urbana, 51, p.61801.

Joe, H., 1996. Families of m-variate distributions with given margins and m (m-1)/2 bivariate dependence parameters. Lecture Notes-Monograph Series, pp.120-141.

Joe, H., Li, H. and Nikoloulopoulos, A.K., 2010. Tail dependence functions and vine copulas. Journal of Multivariate Analysis, 101(1), pp.252-270.

Joe, H., and J.J. Xu. 1996. *The Estimation Method of Inference Functions for Margins for Multivariate Models.* vol. 166. Technical Report.

Just, R.E. and Weninger, Q., 1999. Are crop yields normally distributed?. American Journal of Agricultural Economics, 81(2), pp.287-304.

Kurowicka, D. and Cooke, R.M., 2006. Uncertainty analysis with high dimensional dependence modelling. John Wiley & Sons.

Noland, K., Norvell, J., Paulson, N.D. and Schnitkey, G.D., 2011. The role of farmland in an investment portfolio: analysis of Illinois endowment farms. Journal of the ASFMRA.

Sklar, A. 1959. Distribution Functions in n Dimensions and Their Margins. *Statistics Publications, University of Paris* 8: 229-231.