



Dynamics of HDI Index: Temporal Dependence Based on D-vine Copulas Model for Three-Way Data

Marta Nai Ruscone¹ · Daniel Fernández^{2,3}

Accepted: 30 March 2021 / Published online: 22 May 2021
© The Author(s) 2021

Abstract

The HDI (Human Development Index) is a widely used index based on the average of measures of health, education, and income. It assesses the progress of countries worldwide. The publicly available data set associated with the HDI can be seen as a table with 3 dimensions (three-way table): countries, indexes regarding progress, and years (from 2010 to 2018). Thus, modeling the serial dependence structure of this type of intricate three-way tables is a challenge. D-vine copulas are a special class of multivariate copulas that are particularly suited for modeling serial dependence. This work aims to assess the evolution of the dependence relationship between the indexes of the HDI data set over time through D-vine copulas, which has not been fully used before in the area, as far as we are concerned. We tested our approach to European and African countries and compare their results.

Keywords Copulas · Longitudinal data · Human development index · Dependence

1 Introduction

The human development approach is based on expanding the richness of human life, rather than simply the richness of the economy in which human beings live. It is an approach that is focused on people and their opportunities and choices (see <http://hdr.undp.org/en/human-dev>). There are well-known indexes to measure human development in terms of economic factors such as the Gross Domestic Product (GPD) per capita, the index of innovation and ICT index. However, it is more difficult to capture the human development in terms of non-economic factors (McGillivray 2005). There have been several efforts to describe human welfare based on non-economic facts. The probably more used and accepted is the Human Development Index (HDI). It was introduced in the first Human Development Report,

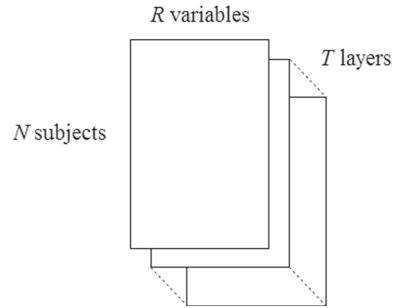
✉ Marta Nai Ruscone
marta.nairuscone@unige.it

¹ Department of Mathematics – DIMA, University of Genova, Via Dodecaneso 35, 16146 Genova, Italy

² Serra Hünter fellow – Department of Statistics and Operations Research, Polytechnic University of Catalonia-BarcelonaTech, 08028 Barcelona, Spain

³ Institute of Mathematics of UPC - BarcelonaTech (IMTech), 08028 Barcelona, Spain

Fig. 1 Three-way data. In the HDI data set, N represents countries, R represents indexes, and T represents years, from 2010 to 2018



presented by the United Nations Development Program in 1990 (Undp 1990). The use of the HDI and its association with other index has been largely studied. For instance, Alijanzadeh, Asefzadeh, and Moosaniaye (2016) studied the association between the HDI and the infant mortality rate, Almasi-Hashiani et al. (2016) analyzed the association between the HDI and mortality rates in other ages and Yakunina and Bychkov (2015) computed the correlation among the HDI and other indexes such as the index of innovation, ICT index, Gross national income and life expectancy at birth (Alijanzadeh et al. 2016).

The HDI is based on the average over three main pillars: health, education, and income. Health is measured by life expectancy. Education is based on mean years of schooling for adults aged 25 and expected years of schooling for children of school age. Finally, income is based on the GDP per capita (see please here for technical notes of the computation of the HDI: http://hdr.undp.org/sites/default/files/hdr2020_technical_notes.pdf). There is an assumption of minimum values for each of the variables. Thus, 0 years is assumed for education, 20 years for life expectancy, and 163 US dollars for income. The HDI provides a measure ranging between 0 and 1. Thus, the higher is the score the more is human development. For instance, Germany and New Zealand scored the highest on the HDI in 2018, with 0.946 and 0.923, respectively. On the other hand, Niger and Eritrea had the lowest HDI with 0.247 and 0.269, respectively.

The HDI data set (we will refer to the data set in this way in this article from now on) can be seen as a table with 3 dimensions (three-way table, see Fig. 1). A first dimension (rows) is described by 188 countries. The second dimension (columns) would be the 4 indexes: the HDI, life expectancy, income and education. The third dimension (layers) would be the years ranging from 2010 to 2018. The first two dimensions, i.e., countries and indexes are labeled as the domain indicators. Additionally, we cannot assume normality in the variables regarding the indexes. Moreover, the variable time (years) is involved in the data and, therefore, we deal with non-normal multivariate longitudinal data, which is quite common in this type of framework.

As we described above, the HDI index is formulated as a linear combination of the other three indexes. Therefore, one of the main challenges of working with this type of intricate 3-way data sets is dealing with modeling a very complex dependence structure. Modeling dependence structures has become more and more popular in all areas of applications over the last two decades. Particularly, copulas have gained large popularity since they allow to model marginal distributions and the dependence structure separately (Sklar 1959). Consequently, copulas were also applied for modeling longitudinal data. This approach was first used by Lambert and Vandenhende (2002) who developed a model for longitudinal data, where the dependence is described by the copula, although their work only applied the Gaussian copula. Shen and Weissfeld (2006) modelled serial dependence for continuous

longitudinal data with a non-ignorable non-monotone missing-data process using a Gaussian copula. Other examples are Lindsey (1999), who used the Gaussian copula among other multivariate models with correlation matrices for non-linear repeated measurements. Further, Sun et al. (2008) argued that elliptical copulas are better suited than Archimedean copulas (Joe 1997; Nelsen 2007) for modeling serial dependence in the context of longitudinal data.

D-vine copulas are a special class of multivariate copulas that are particularly suited for modeling serial dependence (Bedford and Cooke 2002; Aas et al. 2009). Smith et al. (2010) used them to model longitudinal data in a Bayesian approach. Multivariate time series are considered in Smith (2015) and Nai Ruscone and Osmetti (2016). In Joe (2014) (Chapter 7.5) discrete longitudinal count data are modeled using D-vines. Shi and Yang (2018) used a mixed D-vine to model semi-continuous longitudinal claims. The main advantage of this approach is to be able to overcome the problem of the multivariate copulas (Nelsen 2007), which suffer from rather inflexible structures in high dimensions. Therefore, it is a reasonable approach to model our data.

The main aim of this article is to assess the evolution of the dependence relationship between the domain indicators (i.e., countries, indexes) over time (years) through a copula method. Thus, we present a model based on D-vine for longitudinal data, which allows us to represent a unified and general strategy for jointly modeling the variable and the temporal dependence. This approach allows us to get insights into the data via a new flexible way of modeling the dependence between the components of the non-normal multivariate longitudinal data.

The paper is organized as follows: Sect. 2 briefly introduces the D-vine copula model approach for a three-way data set. The dependence relationships between domain indicators over time in the HDI data set are described in Sect. 3. Finally, Sect. 4 contains the conclusions and an outlook on future research.

2 D-vine-based dependence model: a brief review

As mentioned in the previous section, we will apply a D-vine copula model approach to the HDI data set. In this section, we briefly introduce this kind of model. For a more thorough introduction, we refer to Stöber and Czado (2012) and Joe (2014). Sects. 2.1 and 2.2 introduce the general setting of copula and Vine copula, respectively. The particular case of characterizing the dependence relationship for a three-way data set is described in Sect. 2.3.

2.1 Copulas

Copulas provide a great deal of flexibility in modeling multivariate distributions, allowing the researcher to specify the models for the marginal distributions separately from the dependence structure (copula) that links them to form a joint distribution. In addition to flexibility, this helps to expose and understand the various fallacies associated with correlation. The mathematical description of the copula is as follows: let F be a d -dimensional distribution function of the random vector $\mathbf{X} = (X_1, \dots, X_d)$ with univariate marginals F_1, \dots, F_d . A d -variate copula is a multivariate cumulative distribution function (cdf) $C : [0, 1]^d \rightarrow [0, 1]$ with $d \in \mathbb{N}$ (with \mathbb{N} set of natural numbers) and a specific set of uniformly distributed marginals on the interval $[0, 1]$, i.e. $U(0, 1)$. The Sklar's Theorem (Sklar

Table 1 Four copulas functions and ranges of the dependent parameters. Note that x_1 and x_2 are uniform from $[0, 1]$, $\Phi_G(x_1, x_2)$ is the standard bivariate normal distribution with correlation parameter λ , Φ is the distribution function of the standard normal distribution

Copula	$C_\lambda(x_1, x_2)$	Range of λ
Gaussian	$\Phi_\lambda(\Phi_1^{-1}(x_1), \Phi_2^{-1}(x_2))$	$(-1, 1)$
Clayton	$(x_1^{-\lambda} + x_2^{-\lambda} - 1)^{-1/\lambda}$	$[0, \infty)$
Gumbel	$\exp\left\{-\left[(-\ln(x_1))^\lambda + (-\ln(x_2))^\lambda\right]^{1/\lambda}\right\}$	$[1, \infty)$
Frank	$-\frac{1}{\lambda} \log\left(1 + \frac{(e^{-\lambda x_1} - 1)(e^{-\lambda x_2} - 1)}{e^{-\lambda} - 1}\right)$	$[-\infty, \infty) \setminus \{0\}$

1959) provides a link between an arbitrary joint distribution and its marginal distributions and dependence structure. This result has been very important for applications since it allows the marginals and the dependence structure to be modeled separately.

The Sklar’s theorem (Sklar 1959) states that every multivariate distribution F with marginals F_1, \dots, F_d can be written as:

$$F(x_1, \dots, x_d) = C(F_1(x_1), \dots, F_d(x_d)), \quad \text{for all } x \in \mathbb{R}^d. \tag{1}$$

for some appropriate d -dimensional copulas C . For an absolutely continuous F with strictly increasing continuous marginals F_1, \dots, F_d the joint density function f is given by:

$$f(x_1, \dots, x_d) = \left[\prod_{k=1}^d f_k(x_k) \right] c(F_1(x_1), \dots, F_d(x_d)), \quad \text{for all } x \in \mathbb{R}^d, \tag{2}$$

where $c(\cdot)$ denotes the copula density. More details on copulas theory can be found in Nelsen (2007) and Joe (1997) contains a thorough overview of copulas.

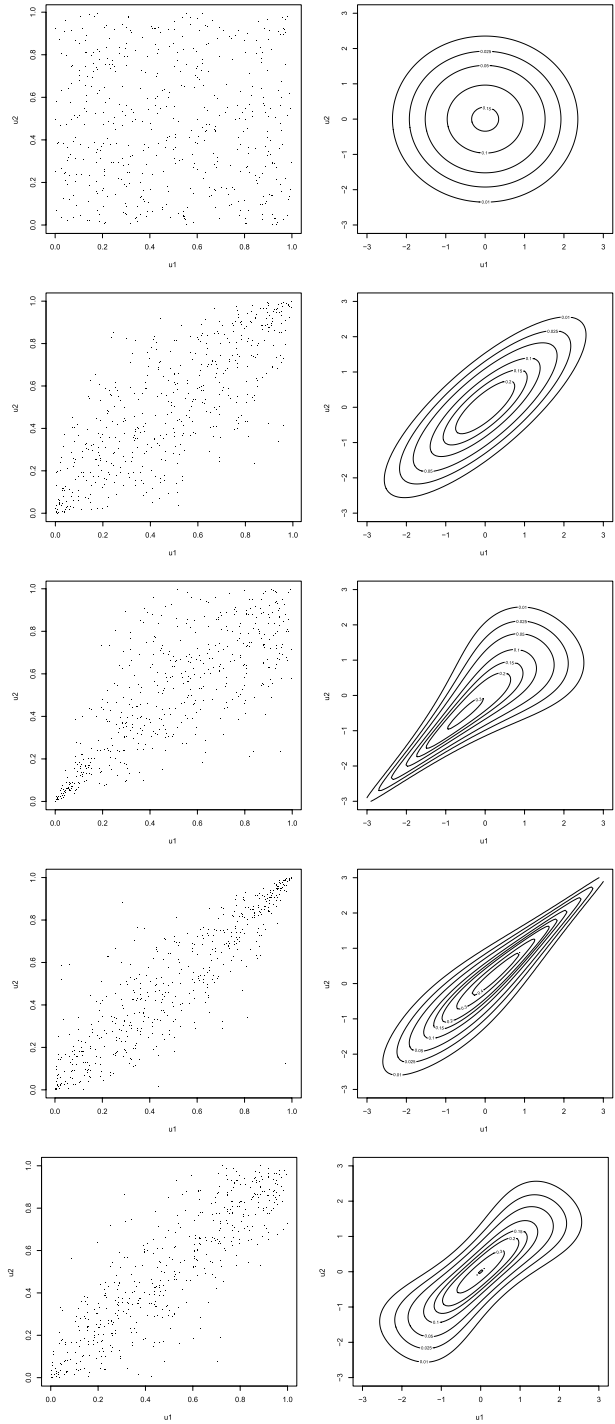
While there is exhaustive literature on bivariate copulas families for the bidimensional case (see Table 1 and Fig. 2), their extension to the multivariate case is not straightforward (Joe 1997). Standard multivariate copulas, such as elliptical or Archimedean (Nelsen 2007), lack of flexibility and accuracy in modeling the dependence structure in a higher dimension. These multivariate extensions imply additional restrictions on the parameters that limit their flexibility. To overcome these problems, vine copulas are proposed by (Joe 1997) and developed in more detail by several authors (see please e.g. Bedford and Cooke 2001a, b; Kurowicka and Cooke 2006; Aas et al. 2009; Berg and Aas 2009).

2.2 Vine copulas

Vine copulas (Bedford and Cooke 2002) is a rich and flexible class of multivariate copulas based on the idea of Joe (1997) to decompose the dependence into a cascade of dependencies between (unconditional/conditional) pairs and permits to construct flexible high-dimensional copulas by using only bivariate copulas as a building blocks (see please Aas et al. (2009)).

Bedford and colleagues (Bedford and Cooke 2002) introduced a graphical model, called regular vine (R-vine), that organizes all valid decompositions and Kurowicka and Cooke (2006) described it in more detail. It involves the specification of a sequence of trees where each edge corresponds to bivariate copulas, the so-called pair-copulas. Then these pair-copulas constitute the building blocks of the joint regular vine distribution.

Fig. 2 Plot and contour plots of bivariate copula models with normal standard margins and dependence parameter λ . From the first row: Independence, Gaussian ($\lambda = 0.8$), Clayton ($\lambda = 3$), Gumbel ($\lambda = 4$) and Frank ($\lambda = 10$) copula



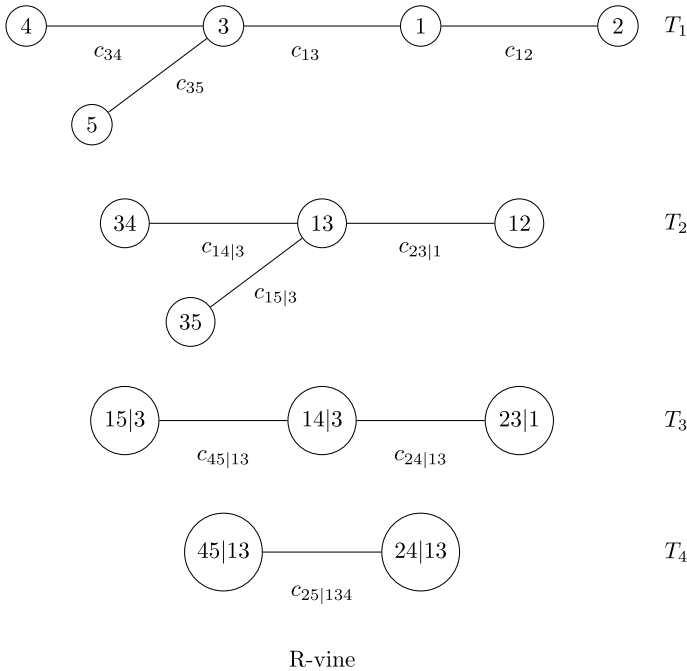


Fig. 3 Five dimensional R-vine tree specification. The figure shows examples of a five-dimensional R-vine copulas with five random variables, four trees and ten edges. The nodes in the first tree correspond to the five random variables that are being modeled and each edge corresponds to a bivariate unconditional or conditional pair-copulas

According to Definition (4.4) of Kurowicka and Cooke (2006), a R-vine V on d variables consists of connected trees T_1, \dots, T_{d-1} with nodes N_i and edges E_i for $i = 1, \dots, d - 1$, which satisfy the following:

- T_1 has nodes $N_1 = \{1, \dots, d\}$ and edges E_1
- For $i = 2, \dots, d - 1$ the tree T_i has nodes $N_i = E_i - 1$.
- (proximity condition) If two edges in tree T_i are to be joined by an edge in tree T_{i+1} they must share a common node.

To develop a statistical model on R-vine trees with the node set $\mathbb{N} := \{N_1, \dots, N_{d-1}\}$ and the edge set $\mathbb{E} := \{E_1, \dots, E_{d-1}\}$, one associates each edge $e = j(e), k(e)|D(e)$ in E_i with a bivariate copula density $c_{j(e),k(e)|D(e)}$. The nodes $j(e)$ and $k(e)$ are called the conditioned nodes, while $D(e)$ is the conditioning set.

More generally, a d -dimensional R-vine is a set of $d - 1$ trees such that the first tree comprises d nodes, identifying $d - 1$ pairs of variables and also $d - 1$ corresponding edges. Therefore each subsequent tree is derived such that all the edges of the tree T_i turn into nodes of the tree $i + 1$; furthermore, two edges in T_i , which become nodes in T_{i+1} are joined by an edge in T_{i+1} only if these edges share a common node in T_i . In Fig. 3 a graphical representation of a regular vine tree sequence of five variable is given. The node in the first tree represent the random variables X_1, X_2, \dots, X_5 . The edges are identified with a bivariate copulas (called pair-copulas), where edge $(j(e), k(e))$ described the dependence between

$X_{j(e)}$ and $X_{k(e)}$. In the second tree, the nodes are just the edges of the first tree. The edges are annotated by $(j(e), k(e)|D(e))$ ¹ and describe the dependence between $X_{j(e)}$ and $X_{k(e)}$ conditional on $X_{D(e)}$. In subsequent trees, the number of conditioning variables increases.

In Theorem (4.2) of Kurowicka and Cooke (2006) it is proven that the joint density of \mathbf{X} is uniquely determined and given by

$$f(x_1, \dots, x_d) = \left[\prod_{k=1}^d f_k(x_k) \right] \left[\prod_{i=1}^{d-1} \prod_{e \in E_i} c_{j(e),k(e)|D(e)}(F(x_{j(e)}|\mathbf{x}_{D(e)}), F(x_{k(e)}|\mathbf{x}_{D(e)})) \right], \tag{3}$$

where $x_{D(e)}$ denotes the sub-vector of $x = (x_1, \dots, x_d)'$ determined by the indices in $D(e)$. The rightmost part of equation (3), which involves the product of d marginal densities f_k and $d(d - 1)/2$ bivariate copulas densities evaluated at the conditional distribution functions $F(\cdot|\cdot)$, is called an R-vine copulas.

One well-known specifications of such representations of R-vine were identified by (Bedford and Cooke 2002): the drawable vine (D-vine, in short). Particularly, an R-vine is called a D-vine if each node in T_1 has a degree at most 2, where the degree of a node denotes the number of connections or edges the node has to other nodes (path structure).

For distinct indices i, j, i_1, \dots, i_k with $i < j$ and $i_1 < \dots < i_k$ we use the abbreviation

$$c_{i,j|i_1, \dots, i_k} := c_{i,j|i_1, \dots, i_k}(F(x_i|x_{i_1}, \dots, x_{i_k}), F(x_j|x_{i_1}, \dots, x_{i_k})). \tag{4}$$

Using this notation the D-vine density is given by

$$f(x_1, \dots, x_d) = \left[\prod_{k=1}^d f_k(x_k) \right] \left[\prod_{j=1}^{d-1} \prod_{i=1}^{d-j} c_{i,i+j|i+1, \dots, i+j-1} \right] \tag{5}$$

See Aas et al. (2009) for more on D-vine copulas.

Fig. (4) depicts the hierarchical nature of the R-vine copulas by plotting the tree structure of D-vine model.

Fitting an R-vine copulas specification to a given dataset requires the following separate tasks:

- (a) selection of the R-vine (structure), i. e., selecting which unconditioned and conditioned pair to use,
- (b) choice of a bivariate copulas family for each pair selected in (a),
- (c) estimation of the corresponding parameters for each copula.

Since all the three step are needed for an R-vine copulas specification, one way of finding the “best” model is to accomplish steps (b) and (c) for all the possible R-vine constructions.

To select one possible R-vine for a given dataset it is necessary to decide for which pairs of variables we want to specify copulas. We proceed sequentially, starting by defining the first tree $T_1 = (N_1, E_1)$ for the R-vine, continuing with the second tree, and so on. The trees are selected in such a way that the chosen pairs model the strongest pairwise dependencies present. The copulas families specified in the first tree of the R-vine have the greatest influence on the model fit. Later, we will refer to this method as the sequential method. This sequential

¹ As in (Aas et al. 2009) it is assumed here that this conditional distribution is independent of the conditioning variable $\mathbf{X}_{D(e)}$

approach is reasonable because the copulas families specified in the first tree of the R-vine have the greatest influence on the model fit (Joe et al. 2010). Similarly, D-vines are also constructed by choosing a specific order of the variables (time ordered). Then in the first tree, the dependence of the first and second variable, of the second and third, of the third and fourth, and so on, is modeled using pair-copulas, i.e., if we assume the order 1, 2, ..., d , we model the pairs (1, 2), (2, 3), (3, 4), etc. In the second tree, conditional dependence of the first and third given the second variable (the pair (1, 3|2)), the second and fourth given the third (the pair (2, 4|3)), and so on, is modeled. In the same way, pairwise dependencies of two variables a and b are modeled in subsequent trees conditioned on those variables which lie between the variables a and b in the first tree, e.g., the pair (1, 5|2, 3, 4). That is each D-vine tree has a path structure. Given the tree model structure we used the Dißmann's algorithm Dißmann et al. (2013) to fit the pair-copulas families and parameters. The bivariate copulas for each pair is selected based on the Akaike Information Criterion (AIC) (Akaike 1998) and the copulas selection is performed sequentially (Dißmann et al. 2013) from the lower to the higher trees. Dißmann's algorithm starts with the estimation of the first tree and estimates the unconditional pair-copulas (and their parameters) via maximum-likelihood estimation. Then the observations are transformed into pseudo-observations needed for the estimation of the second tree using the estimated pair-copulas of tree 1. Continuing this way the vine is built up tree-by-tree.

2.3 Application: D-vine for the HDI data set

The empirical evidence shows that normality is not a rule in practice, thus alternatives to the multivariate normal model must be found. With the aim of relaxing the normality assumption, the procedure based on D-vine copulas theory will be applied here to infer the dependence relationships between the HDI indexes over the years (2010–2018).

This relationship follows dynamics that are not homogeneous over the times and among the HDI indicators as they are characterized by a structurally different underlying phenomena. The D-vine approach allows the relaxation of the assumption of multi-normality that is typical of the usual model for multivariate longitudinal data and it can easily accommodate complex dependence structures such as asymmetric dependence or strong joint tail behavior Joe et al. (2010).

Let's suppose the existence of dependence among the HDI indexes invariant over time. Therefore, the change over time of the distribution of the indexes is due only to a change of the marginal conditional (to the past) distributions of each index and not to the change of the dependence structure across the indexes. The proposed model considers two different levels of analysis. In the first level, a multivariate copulas describes the relations of the indexes observed at a specific time. In the second level of analysis, each longitudinal series, which corresponds to a given index over time, is modeled separately using a pair copulas decomposition to relate the distributions of the variables describing the observation given in different times. The model that we consider starts from a given C_λ a multivariate copulas with parameter λ of the multivariate index variable

$$(\mathbf{Y}^{(1)}, \mathbf{Y}^{(2)}, \dots, \mathbf{Y}^{(R)})$$

such that the joint cumulative distribution function (cdf) is

$$F(\mathbf{y}^{(1)}, \mathbf{y}^{(2)}, \dots, \mathbf{y}^{(R)}) = C_\lambda(F_1(\mathbf{y}^{(1)}), F_2(\mathbf{y}^{(2)}), \dots, F_R(\mathbf{y}^{(R)})) \quad (6)$$

and the joint density function (df) is

$$f(\mathbf{y}^{(1)}, \mathbf{y}^{(2)}, \dots, \mathbf{y}^{(R)}) = c_\lambda(F_1(\mathbf{y}^{(1)}), F_2(\mathbf{y}^{(2)}), \dots, F_R(\mathbf{y}^{(R)})) \prod_{r=1}^R f_r(\mathbf{y}_r) \tag{7}$$

Since longitudinal data are considered each index is observed over times. Therefore, for each index $\mathbf{Y}^{(r)}$ we observe n independent observations on a dependent time series vector $\mathbf{y}^{(r)} = (y_1^{(r)}, y_2^{(r)}, \dots, y_T^{(r)})'$. We model each continuous process which generates the longitudinal data using a pair copulas decomposition (as in (Smith and Khaled 2012), (Nai Ruscone and Osmetti 2016)). In so doing we decompose the distribution of the process at point in time, conditional to the past, into the product of a sequence of bivariate copulas density and marginal density. The advantage is that the marginal distribution of the process at each point can be modeled arbitrarily, while the dependence over time is captured by a sequence of bivariate copulas.

Let (y_1, y_2, \dots, y_T) (a simplification of the notation dropping the index r) be the univariate series, the joint density function can be decomposed in a product of the conditional (to the past) distributions:

$$f(y_1, y_2, \dots, y_T) = \prod_{i=1}^T f(y_i | y_{i-1}, \dots, y_1) f(y_1) \tag{8}$$

By using a pair copula decomposition we have:

$$f(y_i | y_{i-1}, \dots, y_1) = \prod_{j=1}^{i-1} c_{i,j}(F(y_i | y_{i-1}, \dots, y_{j+1}), F(y_j | y_{i-1}, \dots, y_{j+1}); \theta_{i,j}) f(y_i)$$

where $F(y_i)$ and $f(y_i)$ are the cdf and the df of the marginal Y_i and $c_{i,j} = c_{i,j|t-1, t-2, \dots, j+1}$ are the pair copulas with parameters $\theta_{i,j}$. Therefore, the joint distribution of the process becomes a D-vine copulas model of order T , which is a product of T marginal densities and $T(T - 1)/2$ pair copulas densities:

$$f(y_1, y_2, \dots, y_T) = \prod_{t=2}^T \left[\prod_{j=1}^{t-1} c_{t,j}(u_{t|j+1}, u_{j|t-1}; \theta_{t,j}) f(y_t) \right] f(y_1) \tag{9}$$

where $u_{t|j+1} = F(y_t | y_{t-1}, \dots, y_{j+1})$ and $u_{j|t-1} = F(y_j | y_{t-1}, \dots, y_{j+1})$.

By substituting (9) and the correspondent cumulative distribution function in (7) we obtain the model for multivariate longitudinal data:

$$f(Y^{(1)}, Y^{(2)}, \dots, Y^{(R)}) = \prod_{r=1}^R \left(\prod_{t=2}^T \left[\prod_{j=1}^{t-1} c_{t,j}(u_{t|j+1}^{(r)}, u_{j|t-1}^{(r)}; \theta_{t,j}^{(r)}) f(y_t^{(r)}) \right] f(y_1^{(r)}) \right) \cdot c_\lambda \left(\prod_{t=2}^T \left[\prod_{j=1}^{t-1} C_{t,j}^{(1)}(u_{t|j+1}^{(1)}, u_{j|t-1}^{(1)}; \theta_{t,j}^{(1)}) \right] \dots, \prod_{t=2}^T \left[\prod_{j=1}^{t-1} C_{t,j}^{(R)}(u_{t|j+1}^{(R)}, u_{j|t-1}^{(R)}; \theta_{t,j}^{(R)}) \right] \right) \tag{10}$$

In (10) λ describes the dependence between the indexes and $\theta_{i,j}^{(r)}$ describes the dependence between the r -th index at time t and the one at time j .

We define the distribution of the R index variables at time t conditional to their past. Let now $y_t^{*(r)} = y_t^{(r)} | y_{t-1}^{(r)}, \dots, y_1^{(r)}$, we consider the factorization:

$$f(y_t^{*(1)}, y_t^{*(2)}, \dots, y_t^{*(R)}) = \prod_{r=2}^R \left[f_{\lambda_r}(y_t^{*(r)} | y_t^{*(r-1)}, \dots, y_t^{*(1)}) \right] f(y_t^{*(1)})$$

Then by applying the PCC (respect to the time) to each conditional distribution we define the joint distribution as the extension of the D-vine model in R dimensions.

$$f(\mathbf{y}^{(1)}, \mathbf{y}^{(2)}, \dots, \mathbf{y}^{(R)}) = \prod_{r=1}^R \left\{ \prod_{t=2}^T \left[\prod_{j=1}^{t-1} c_{tj}(\mathbf{u}_{t|j+1}, \mathbf{u}_{j|t-1}; \theta_{t,s}) f_{\lambda_t}(y_t^{*(r)} | y_t^{*(r-1)}, \dots, y_t^{*(1)}) \right] \cdot f_{\lambda_1}(y_1^{*(r)} | y_1^{*(r-1)}, \dots, y_1^{*(1)}) \right\} \tag{11}$$

where $u_{t|j+1} = F(y_t^{*(r)} | y_t^{*(r-1)}, \dots, y_t^{*(1)})$. The df f_{λ_t} can be also defined by a PCC (between the indexes). In (11) λ_t describes the dependence between the indexes at time t and $\theta_{ij}^{(r)}$ describes the dependence between the r-th index at time t and the one at time j.

Note that in both models we suppose that the response at time t is independent from the past of the other variables.

The code of the algorithm is based on functions in the R packages *CDVine* (Schepsmeier et al. 2015) and *VineCopula* (Schepsmeier et al. 2015) and the MLE (Maximum Likelihood Estimation) used here therefore corresponds to the maximum pseudo likelihood method (MPL) (Genest et al. 1995).

3 Dependence relationships among indicators in the HDI data set

As we described in Sect. 1, the HDI data set used to apply the D-vine copulas methods described in 2.2 is composed of 4 indexes in a study period of 9 years (2010-2018): Education, Income, Life Expectancy, and the HDI itself (see Appendix Figs. 5, 6, 7, 8, 9, 10, 11, 12, 13). The region distribution over the 5 continents is as follows: 53 countries from Africa, 35 countries from America, 48 from Asia, 41 from Europa, and 11 from Oceania. Thus, we will find estimates for the set of pair-copula families in Eq. (11) and the relative set of parameters. Additionally, in the case of the HDI data set, we also want to jointly assess the evolution of the dependence relationship between the four domain indexes over time. Therefore, this copula approach will allow us to model the dependence structure between the indexes and the temporal structure from 2010 to 2018 at the same time.

For fitting the model in Eq.(11) with flexible lower/upper tail dependence, we considered the following copulas families, which are classified according to their different strengths of tail behavior in the estimation process of the copulas:

- Gaussian/Normal (tail-symmetric, with no tail dependence at all)
- Student-t (tail-symmetric, with upper and lower tail dependence)
- Clayton (tail-asymmetric, with only lower tail dependence)
- Gumbel (tail-asymmetric, with only upper tail dependence)
- Frank (tail-symmetric, with no tail dependence)

for more details see please Nelsen (2007). The suitable copulas approaches in the case of positive dependence, are the Gaussian, Student-t, Gumbel and Frank copulas. Otherwise, Clayton copula is the appropriate when we observed negative dependence see Tab. 1 and Fig. 2.

Given those different copulas candidates, the next step is to fit the model shown in Eq.(11) and select the best candidate for the HDI data set. In the first step, each longitudinal series of data corresponding to a given index is modelled separately using a D-vine copulas to relate the marginal distributions of the indexes at each time of observation. In other words, we assume that we can explain the effect of the past indexes on the r th index by the history of the r th index and by the current values of the other indexes. In the second step, at each observation time, the conditional (on the past) distributions of each index are related using another D-vine copulas describing the relationship between the corresponding variables.

As we want to test whether there are differences between countries according to location, we stratified the HDI data set into 2 data sets: countries with the lowest HDI (African countries), countries with the highest HDI (European countries). Thus, we compare the dependence between these two data sets, but also comparing both data sets with the whole group of countries.

Following the decomposition shown in the Eq.(11), the 4 D-vine copulas (each one with 8 trees) is fitted.

Tables 5–2 listed the $\theta_{i,j}^{(r)}$ estimates, which describe the dependence between the r th index at time t and the one at time j . Additionally, Figs. 14, 15, 16, 17 depict the graphical representation of the D-vine fitting results by means of contour plot. The results between Tables and Figures are equivalent. For each Figure, the contour of the D-vine for the whole countries (top), only the European countries (middle) and only the African countries (bottom) are depicted. These contour plots graphically show all the conditional/ unconditional copulas estimates in every tree D-vine (t the bottom of every picture the first tree, and on the top of every copulas the label as in Fig. 4.) As we can observe the dependence structure in all these D-vine cases (World, Europe, and Africa) is slightly different. Table 5 shows in the first two trees all symmetric dependence in all the three cases (World, Europe, and Africa) but in the Table 2 and the Table 3 we can observe upper tail dependence (Gumbel copula) in the first two trees in the case of Europe and Africa. This means that indexes emphasize dependence among the higher level of two consecutive years. For instance in Africa, c_{89} (Table 2) we got among HDI of 2017 and 2018 a Gumbel (18.36) copula, therefore there is dependence in the case of the higher values among the indexes HDI in those two years. One of the main results we observed in African countries is that there is a significant upper tail dependence in the case of both the life expectancy index and the HDI. This implies that there has been a positive tendency and improvement in living conditions in Africa over the last years. In the case of the European countries, we observed a quite significant right tail dependence in the HDI, which in this case might mean that the level of HDI is keeping slightly a positive tendency over the studied years. This makes sense as Europe has less to improve than in Africa. From the copulas families selected, we see evidence of different types of asymmetric dependence. This demonstrates that the choice of D-vine was appropriate, since it guarantees enough flexibility to model the dependence structure. According to Joe et al. (2010) the copulas families in the first tree $c_{12}, c_{23}, c_{34}, c_{45}, c_{56}, c_{67}, c_{78}, c_{89}$ of the model have the greatest influence on the model fitting.

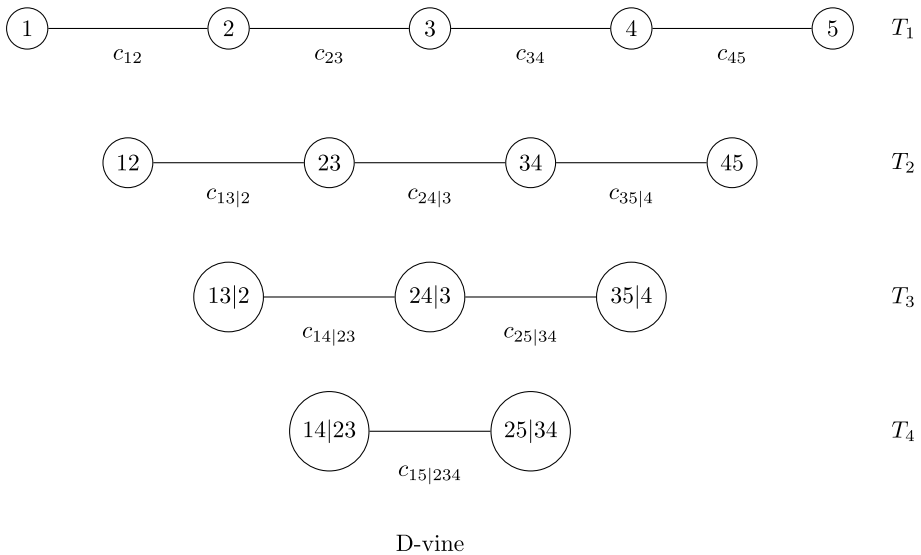


Fig. 4 Five dimensional D-vine copulas. The figure shows examples of a five-dimensional D-vine copulas with five random variables, four trees and ten edges. The nodes in the first tree correspond to the five random variables that are being modeled and each edge corresponds to a bivariate conditional or unconditional pair-copulas

4 Discussion

The HDI is one of the more used measures to rank countries into three tiers of human development: lifespan, education level and gross national income per capita. Its associated data set includes a temporal structure that motivates a dependency modeling, which has not been fully used before in the area, as far as we are concerned. Our research work employed a D-vine copulas approach to jointly modeling the dependence among the four HDI indexes (HDI itself, life expectancy, income, and education) at a given time and across time. This modeling approach allows us to describe a complex pattern of dependence structure in the HDI data set using a flexible high-dimensional model based on a D-vine copulas. Thus, it overcomes the problem of modeling simultaneous dependence between non-normal responses over time. Particularly, this allows modeling non-linear and asymmetric cross-sectional and serial dependence, using a D-vine copulas. The use of the methods described in this article may be advantageous for practitioners in the field because they can jointly model cross-section dependence and serial dependence in a really flexible way. As the joint estimation of the D-vine copulas might be computationally slow and time consuming in high dimensional spaces, future research direction would include exploring possibilities of speeding up the estimation process. One option would be to implement a fast sequential alternative based on a simplified D-vine. Another possibility would be to use parallel computing approach via the R package *parallel*, which is included in the R base.

Another future research direction we would like to intent is to develop a model-based clustering approach, which will take into account both cross-sectional and serial dependence for a comprehensive class of continuous three-way data sets. This modeling

strategy will determine clustering structures of countries according to low-, medium-, high-, and very high-HDI levels based on both cross-sectional and serial dependence and not only taking into account cross-sectional dependence.

The R code to fit the models is available upon request from the first author.

Appendix

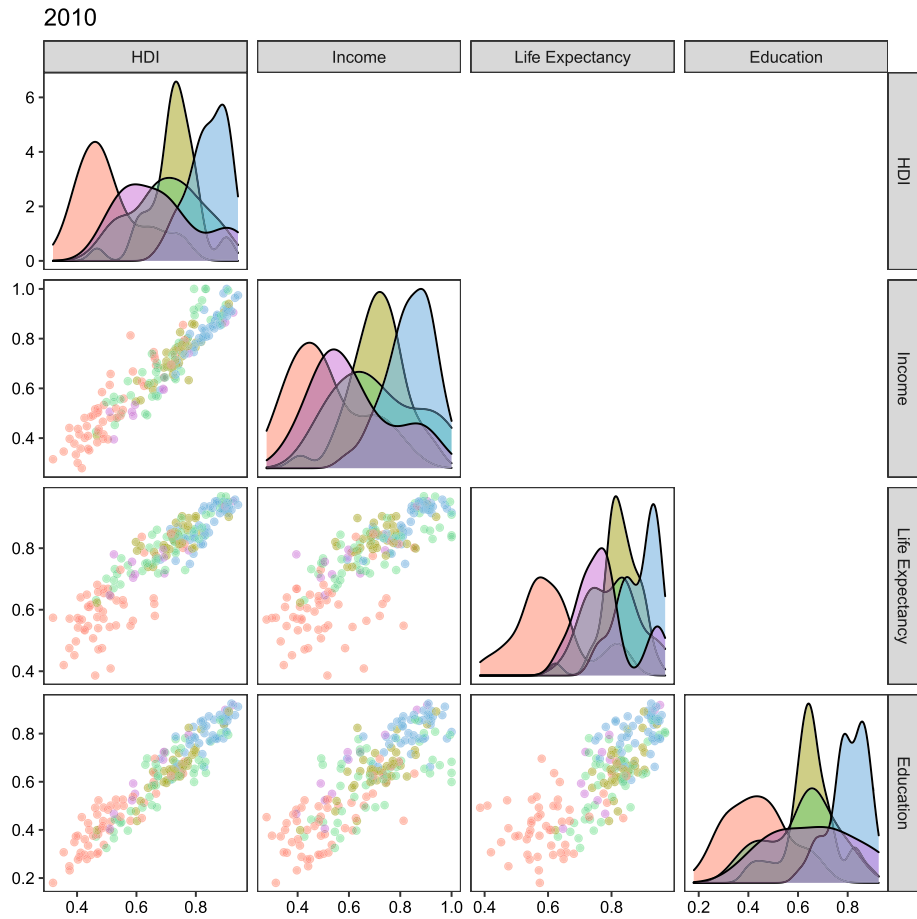


Fig. 5 Pairwise scatter plot of the indexes, year 2010. Density distribution by continent on the diagonal. Red: Africa; yellow: America; green: Asia; Blue: Europe; pink: Oceania

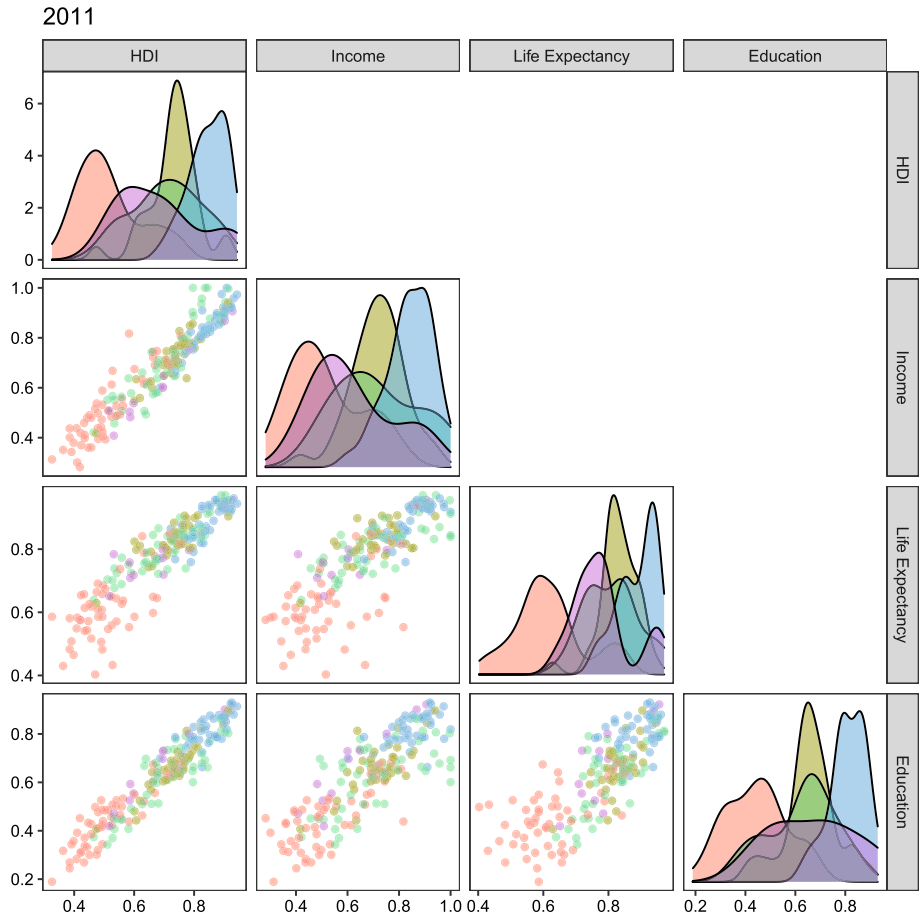


Fig. 6 Pairwise scatter plot of the indexes, year 2011. Density distribution by continent on the diagonal. Red: Africa; yellow: America; green: Asia; Blue: Europe; pink: Oceania

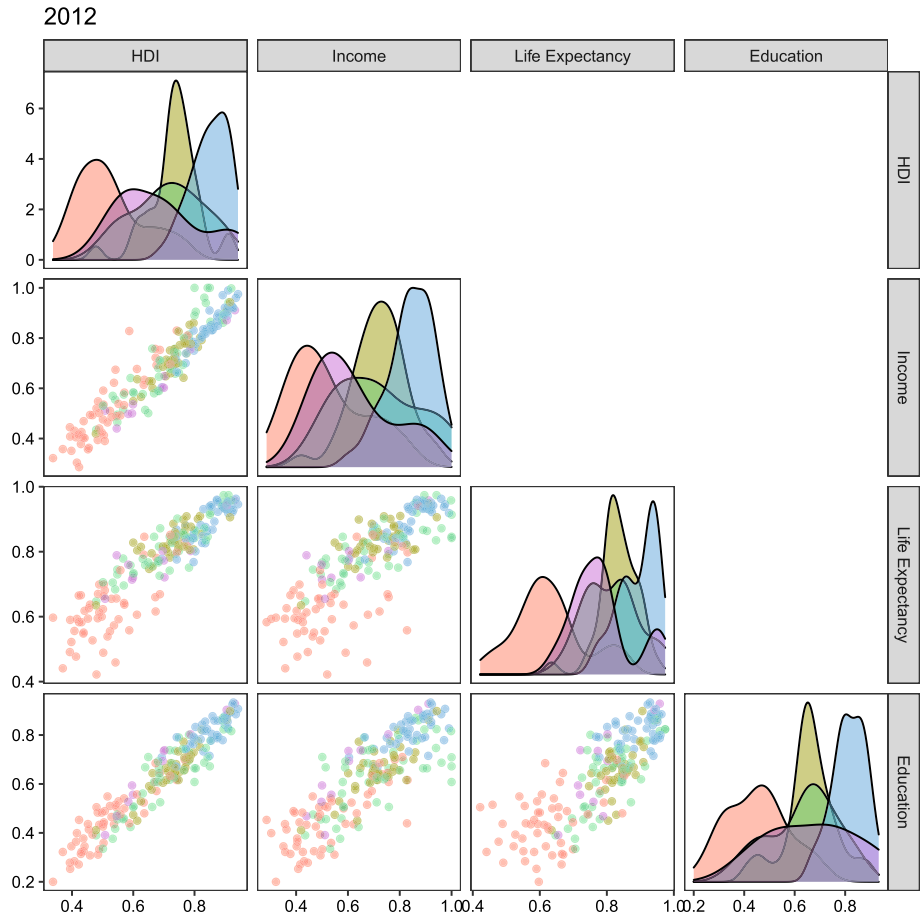


Fig. 7 Pairwise scatter plot of the indexes, year 2012. Density distribution by continent on the diagonal. Red: Africa; yellow: America; green: Asia; Blue: Europe; pink: Oceania

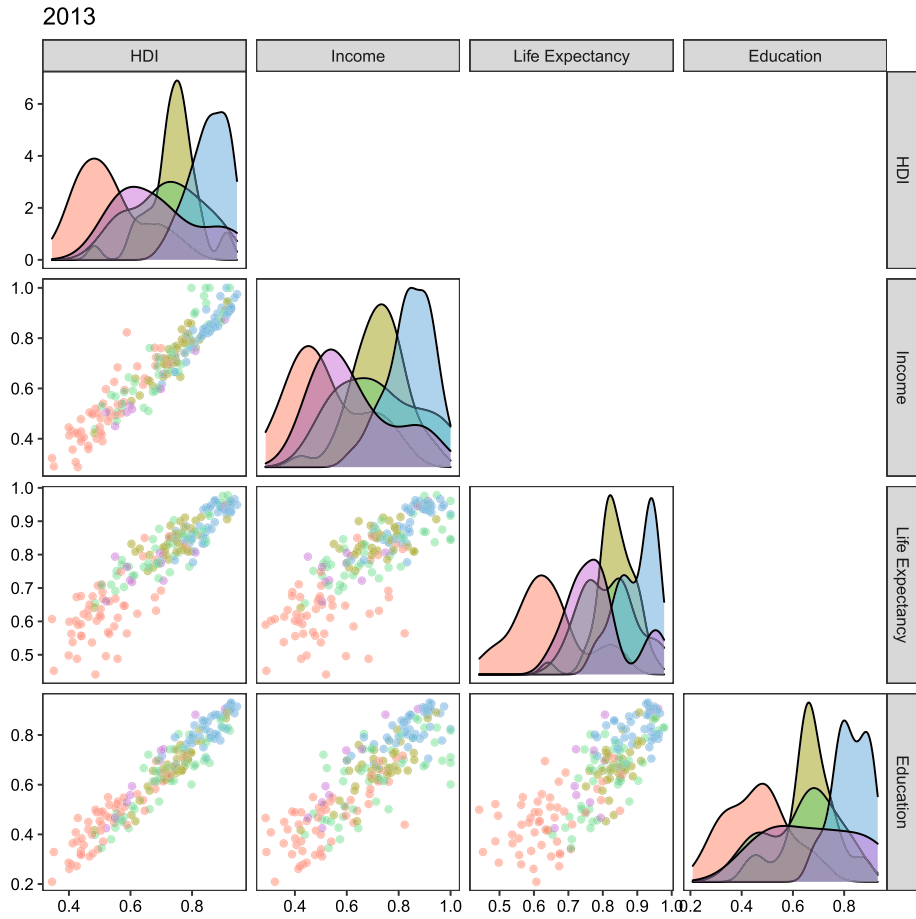


Fig. 8 Pairwise scatter plot of the indexes, year 2013. Density distribution by continent on the diagonal. Red: Africa; yellow: America; green: Asia; Blue: Europe; pink: Oceania

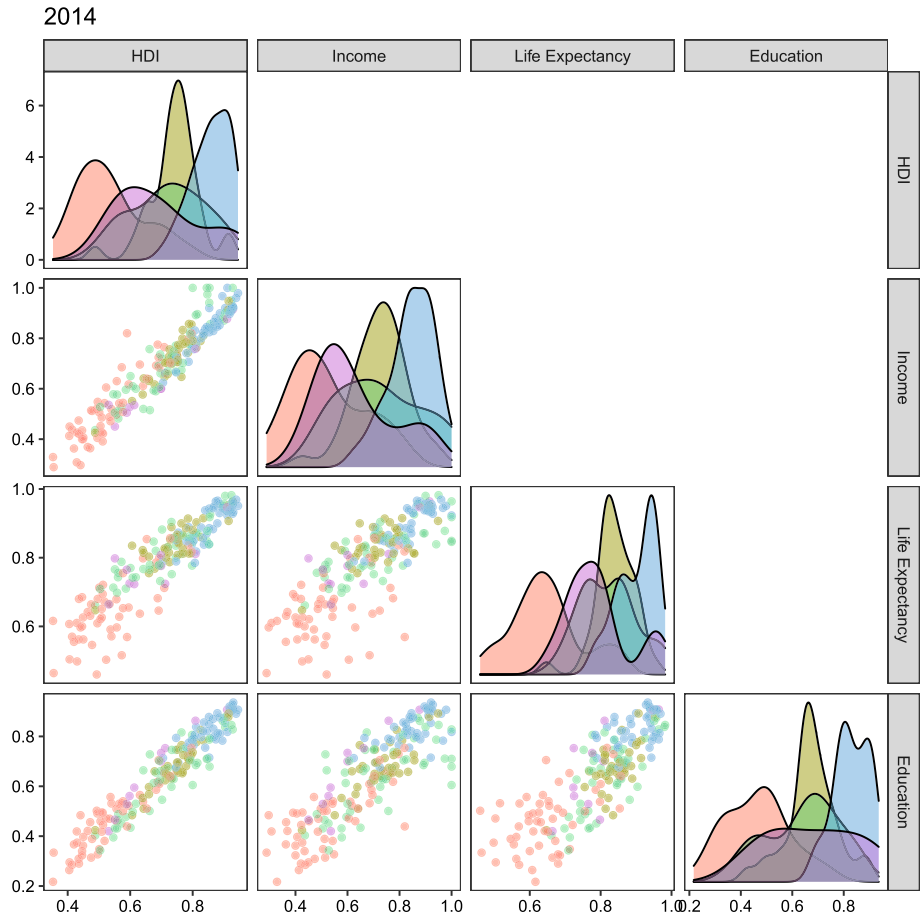


Fig. 9 Pairwise scatter plot of the indexes, year 2014. Density distribution by continent on the diagonal. Red: Africa; yellow: America; green: Asia; Blue: Europe; pink: Oceania

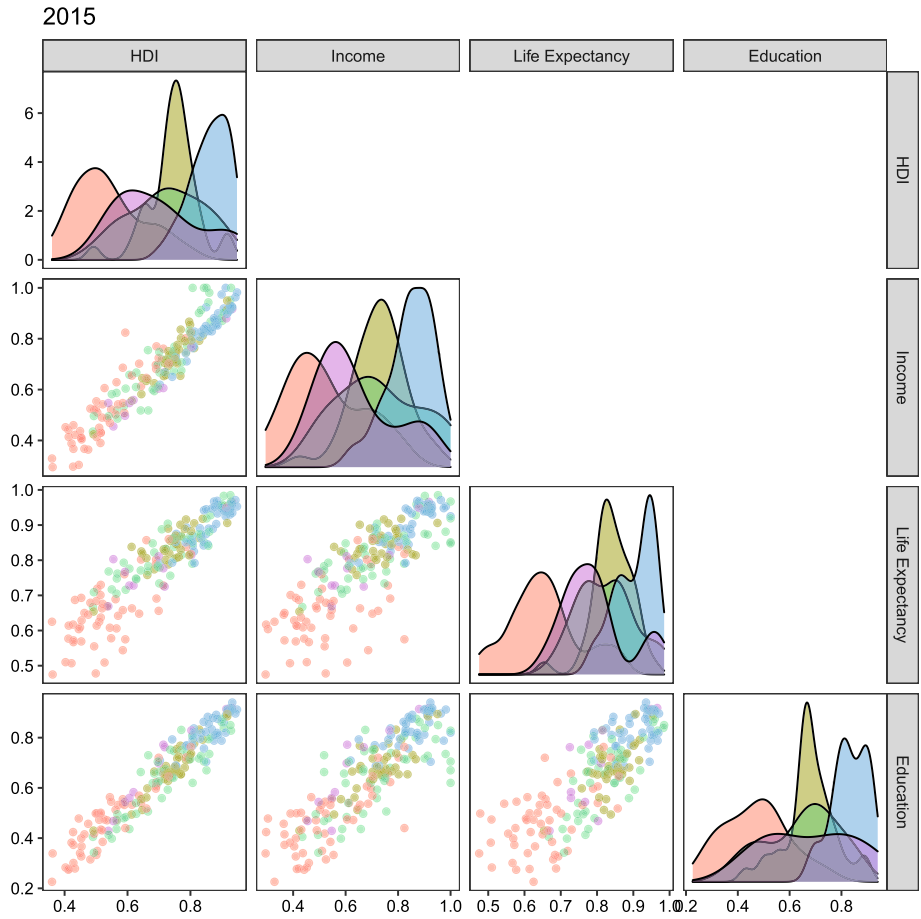


Fig. 10 Pairwise scatter plot of the indexes, year 2015. Density distribution by continent on the diagonal. Red: Africa; yellow: America; green: Asia; Blue: Europe; pink: Oceania

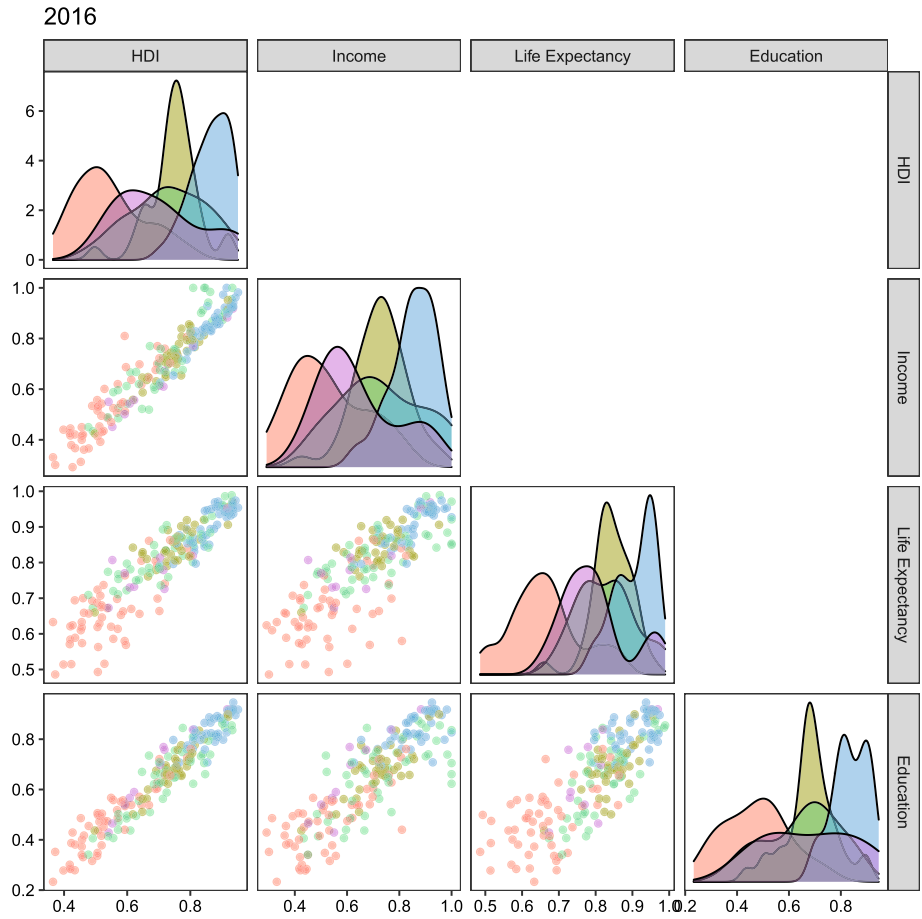


Fig. 11 Pairwise scatter plot of the indexes, year 2016. Density distribution by continent on the diagonal. Red: Africa; yellow: America; green: Asia; Blue: Europe; pink: Oceania

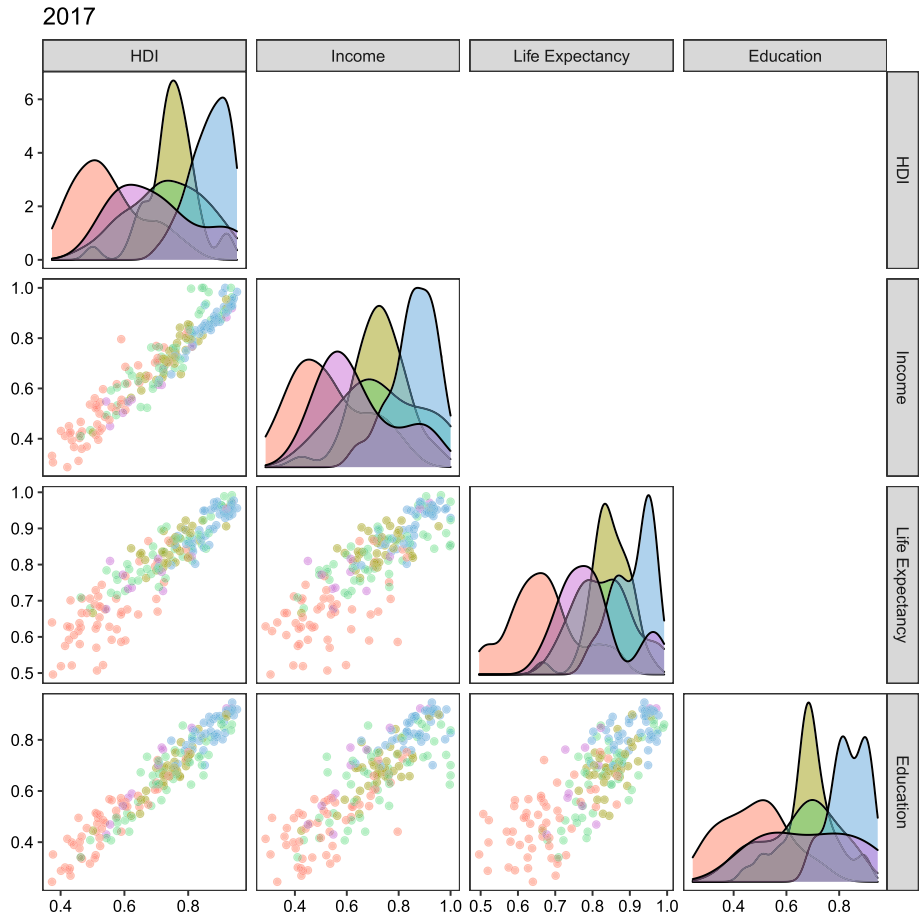


Fig. 12 Pairwise scatter plot of the indexes, year 2017. Density distribution by continent on the diagonal. Red: Africa; yellow: America; green: Asia; Blue: Europe; pink: Oceania

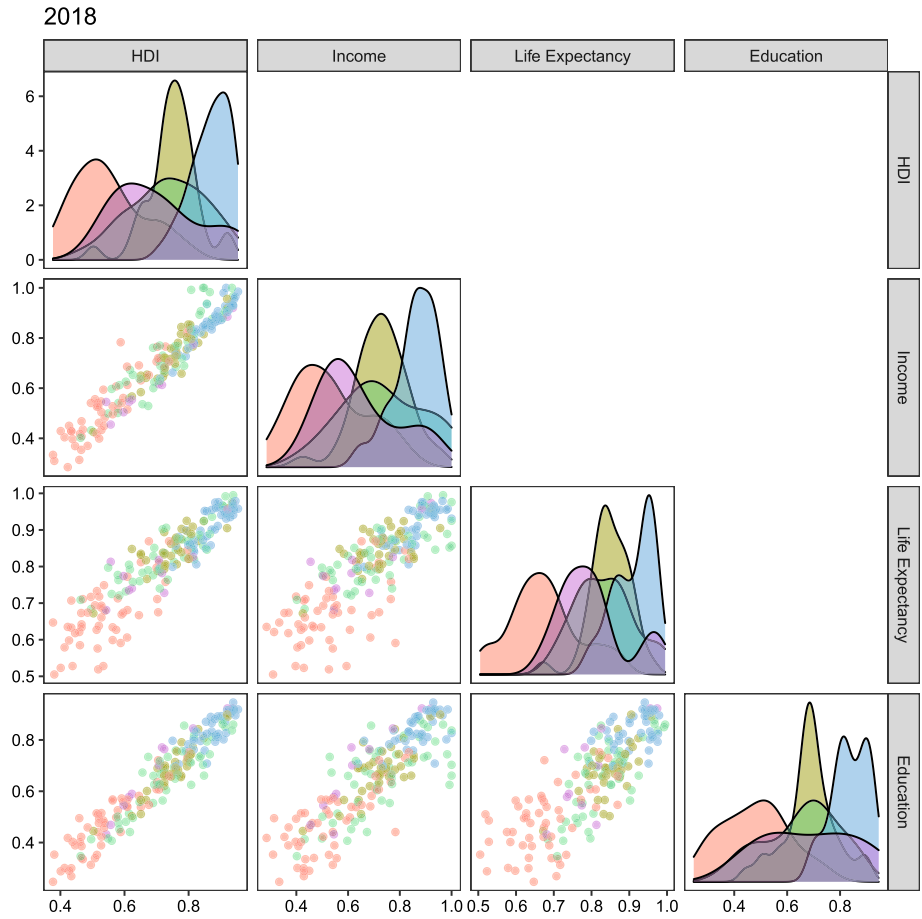


Fig. 13 Pairwise scatter plot of the indexes, year 2018. Density distribution by continent on the diagonal. Red: Africa; yellow: America; green: Asia; Blue: Europe; pink: Oceania

Table 2 D-vine parameter estimates for the HDI in all the 8 trees (T_1 the first tree, T_8 the last tree). From the left to the right: the whole countries, only European countries, and only African countries. The selected copulas family (*Cop*: Gauss, t-Student (labeled as t), Clayton, Frank, and Gumbel copula) and the copulas parameters (*par.1* and *par.2*) can be one or two according to the type of copulas. *N/A* identifies absence of one parameter)

	All Countries				Europe			Africa		
	<i>Cop</i>	<i>fam</i>	<i>par.1</i>	<i>par.2</i>	<i>fam</i>	<i>par.1</i>	<i>par.2</i>	<i>fam.</i>	<i>par.1</i>	<i>par.2</i>
T_1	c_{12}	t	0.995	3.345	Gumbel	5.504	N/A	Clayton	12.406	N/A
	c_{23}	t	0.993	2.390	Gumbel	8.33	N/A	t	0.988	2
	c_{34}	t	0.994	2.000	Gauss	0.982	N/A	t	0.99	2.913
	c_{45}	t	0.996	2.000	Gumbel	10.315	N/A	Gumbel	17.438	N/A
	c_{56}	t	0.996	2.867	Gauss	0.982	N/A	Gumbel	16.041	N/A
	c_{67}	t	0.996	2.355	Frank	27.975	N/A	t	0.994	2
	c_{78}	t	0.996	2.209	Gumbel	11.541	N/A	t	0.997	2
	c_{89}	t	0.998	2.375	t	0.997	2	Gumbel	18.836	N/A
	T_2	$c_{13 2}$	t	-0.098	4.414	Gumbel	1.027	N/A	Gumbel	1.193
$c_{24 3}$		t	-0.052	3.217	Gumbel	1.109	N/A	t	-0.091	3.255
$c_{35 4}$		t	-0.072	4.933	Frank	1.2	N/A	Gauss	-0.121	N/A
$c_{46 5}$		Gauss	-0.143	N/A	t	-0.117	3.563	Gauss	-0.019	N/A
$c_{57 6}$		Gauss	-0.094	N/A	Gauss	-0.346	N/A	Frank	1.503	N/A
$c_{68 7}$		Gauss	-0.226	N/A	Gumbel	1.139	N/A	Gauss	0.116	N/A
$c_{79 8}$		t	-0.264	3.277	Clayton	0.259	N/A	Gauss	-0.186	N/A
T_3	$c_{14 23}$	t	-0.282	3.268	Gauss	0.214	N/A	Gauss	-0.25	N/A
	$c_{25 34}$	Gauss	-0.133	N/A	t	-0.07	2.726	Gumbel	1.102	N/A
	$c_{36 45}$	t	-0.055	6.524	Gauss	0.36	N/A	t	-0.127	2.923
	$c_{47 56}$	t	-0.119	8.646	Frank	-2.936	N/A	t	0.017	6.149
	$c_{58 67}$	t	-0.085	4.915	t	-0.418	2	t	0.084	2.438
	$c_{69 78}$	Gauss	-0.147	N/A	Frank	-1.613	N/A	Frank	0.76	N/A
T_4	$c_{15 234}$	Frank	-0.449	N/A	Frank	-0.275	N/A	Gauss	0.185	N/A
	$c_{26 345}$	t	-0.195	5.260	Frank	-0.583	N/A	Frank	-0.855	N/A
	$c_{37 456}$	Frank	-0.738	N/A	Frank	0.755	N/A	Frank	-1.235	N/A
	$c_{48 567}$	Gumbel	1.041	N/A	Clayton	0.253	N/A	Gauss	-0.032	N/A
	$c_{59 678}$	Gauss	-0.233	N/A	Frank	-0.786	N/A	Frank	-0.229	N/A
T_5	$c_{16 2345}$	Clayton	0.040	N/A	Frank	-0.691	N/A	Clayton	0.381	N/A
	$c_{27 3456}$	Gauss	0.049	N/A	Gauss	0.116	N/A	Gauss	-0.039	N/A
	$c_{38 4567}$	Gumbel	1.058	N/A	Gauss	0.042	N/A	Frank	-0.405	N/A
	$c_{49 5678}$	t	0.019	6.723	t	-0.364	2.273	Frank	-1.323	N/A
T_6	$c_{17 23456}$	Frank	-0.204	N/A	Gauss	-0.135	N/A	Gumbel	1.034	N/A
	$c_{28 34567}$	t	0.085	5.975	Gauss	-0.017	N/A	Frank	-0.286	N/A
	$c_{39 45678}$	t	0.040	6.429	Frank	0.184	N/A	Gumbel	1.182	N/A
T_7	$c_{18 234567}$	Clayton	0.063	N/A	Frank	-1.286	N/A	Frank	-1.113	N/A
	$c_{29 345678}$	Frank	-0.389	N/A	Gauss	-0.278	N/A	Gauss	-0.135	N/A
T_8	$c_{19 2345678}$	Gauss	0.090	N/A	t	0.388	N/A	Gauss	-0.192	N/A

Table 3 D-vine parameter estimates for the Life Expectancy in all the 8 trees (T_1 the first tree, T_8 the last tree). From the left to the right: the whole countries, only European countries, and only African countries. The selected copulas family (*Cop*: Gauss, t-Student (labeled as t), Clayton, Frank, and Gumbel copula) and the copulas parameters (*par.1* and *par.2*) can be one or two according to the type of copulas. *N/A* identifies absence of one parameter)

	All Countries				Europe			Africa		
	<i>Cop</i>	<i>fam</i>	<i>par.1</i>	<i>par.2</i>	<i>fam</i>	<i>par.1</i>	<i>par.2</i>	<i>fam</i>	<i>par.1</i>	<i>par.2</i>
T_1	c_{12}	t	0.998	2.899	Gumbel	3.248	N/A	Frank	74.256	N/A
	c_{23}	t	0.998	4.409	Gauss	0.955	N/A	t	0.994	2.125
	c_{34}	t	0.999	3.673	t	0.952	2	Gumbel	15.303	N/A
	c_{45}	Gumbel	39.708	N/A	Gumbel	8.216	N/A	Gumbel	27.351	N/A
	c_{56}	Gumbel	40.606	N/A	Gumbel	7.498	N/A	Gumbel	19.475	N/A
	c_{67}	Gumbel	42.232	N/A	Gumbel	7.269	N/A	Gumbel	26.213	N/A
	c_{78}	t	0.999	5.39	Gauss	0.991	N/A	Gumbel	26.708	N/A
	c_{89}	Gumbel	62.675	N/A	t	0.998	2.352	Gumbel	15.185	N/A
	T_2	$c_{13 2}$	t	-0.279	3.684	Clayton	0.217	N/A	Frank	-1.715
$c_{24 3}$		t	-0.168	5.463	t	0.295	2.285	t	-0.431	2.737
$c_{35 4}$		t	-0.061	8.327	Clayton	0.547	N/A	Frank	-1.266	N/A
$c_{46 5}$		t	-0.159	6.042	t	-0.173	2	Frank	-1.871	N/A
$c_{57 6}$		Frank	-1.565	N/A	t	-0.206	3.436	Gauss	-0.295	N/A
$c_{68 7}$		Frank	-1.435	N/A	t	-0.006	4.236	Gauss	-0.203	N/A
$c_{79 8}$		t	-0.309	6.47	Frank	1.344	N/A	t	-0.252	4.126
T_3	$c_{14 23}$	t	-0.171	4.362	Gauss	0.149	N/A	Gauss	-0.174	N/A
	$c_{25 34}$	Frank	-2.148	N/A	t	-0.098	2.578	Gauss	-0.089	N/A
	$c_{36 45}$	t	-0.091	7.043	Gumbel	1.232	N/A	Frank	0.993	N/A
	$c_{47 56}$	Frank	-0.853	N/A	Clayton	0.132	N/A	t	0.003	4.689
	$c_{58 67}$	Gauss	-0.081	N/A	Gauss	-0.09	N/A	Frank	-3.95	N/A
	$c_{69 78}$	Frank	-1.136	N/A	Frank	-0.244	N/A	Gumbel	1.112	N/A
T_4	$c_{15 234}$	t	-0.163	10.192	Clayton	0.356	N/A	Frank	-0.673	N/A
	$c_{26 345}$	Frank	-0.487	N/A	Frank	-0.517	N/A	Frank	0.247	N/A
	$c_{37 456}$	Frank	-0.643	N/A	Gauss	-0.115	N/A	Frank	-1.595	N/A
	$c_{48 567}$	t	-0.028	11.186	t	0.124	3.024	t	-0.198	2.241
T_5	$c_{59 678}$	Frank	-1.02	N/A	Clayton	0.136	N/A	t	-0.2	3.696
	$c_{16 2345}$	Gauss	-0.094	N/A	Frank	-1.458	N/A	Gumbel	1.049	N/A
	$c_{27 3456}$	Gumbel	1.044	N/A	Clayton	0.265	N/A	Gauss	-0.116	N/A
	$c_{38 4567}$	Gauss	-0.078	N/A	Clayton	0.104	N/A	Frank	-0.731	N/A
	$c_{49 5678}$	Frank	-0.12	N/A	t	-0.087	3.923	Clayton	0.081	N/A
T_6	$c_{17 23456}$	Frank	0.17	N/A	Frank	-0.663	N/A	Gumbel	1.119	N/A
	$c_{28 34567}$	t	0.239	5.986	Gauss	-0.169	N/A	Gauss	0.041	N/A
	$c_{39 45678}$	Frank	-0.616	N/A	Clayton	0.482	N/A	Frank	1.411	N/A
T_7	$c_{18 234567}$	Gauss	-0.014	N/A	Frank	-0.06	N/A	Gauss	0.231	N/A
	$c_{29 345678}$	Clayton	0.098	N/A	Gauss	-0.113	N/A	Gauss	-0.158	N/A
T_8	$c_{19 2345678}$	Gauss	-0.043	N/A	Clayton	0.446	N/A	Gauss	-0.184	N/A

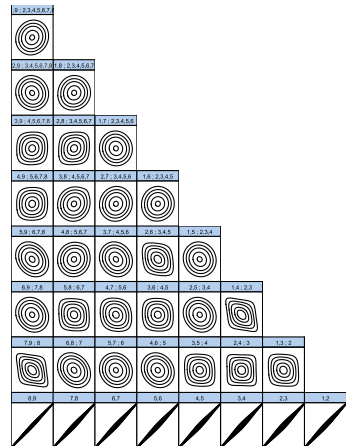
Table 4 D-vine parameter estimates for the Income in all the 8 trees (T_1 the first tree, T_8 the last tree). From the left to the right: the whole countries, only European countries, and only African countries. The selected copulas family (*Cop*: Gauss, t-Student (labeled as t), Clayton, Frank, and Gumbel copula) and the copulas parameters (*par.1* and *par.2*) can be one or two according to the type of copulas. *N/A* identifies absence of one parameter)

	All Countries				Europe			Africa		
	<i>Cop</i>	<i>fam</i>	<i>par.1</i>	<i>par.2</i>	<i>fam</i>	<i>par.1</i>	<i>par.2</i>	<i>fam.</i>	<i>par.1</i>	<i>par.2</i>
T_1	c_{12}	t	0.992	2.000	Gauss	0.941	N/A	t	0.978	2
	c_{23}	t	0.988	2	t	0.977	5.498	t	0.984	2
	c_{34}	t	0.988	2	t	0.97	5.582	t	0.979	2
	c_{45}	t	0.995	2	t	0.991	2	t	0.991	2
	c_{56}	t	0.993	2	t	0.988	2	t	0.99	2.147
	c_{67}	t	0.995	4.529	Gauss	0.992	N/A	Gauss	0.988	N/A
	c_{78}	t	0.995	3.118	Gumbel	16.878	N/A	t	0.99	2
	c_{89}	t	0.998	2	t	1	2	t	0.998	2
	T_2	$c_{13 2}$	t	-0.058	3.524	Gauss	0.161	N/A	t	0.254
$c_{24 3}$		Gumbel	1.098	N/A	Gauss	0.459	N/A	t	0.188	4.395
$c_{35 4}$		t	0.126	3,967	Frank	0.347	N/A	t	0.148	4.59
$c_{46 5}$		Frank	-1.434	N/A	Frank	0.435	N/A	Gauss	-0.204	N/A
$c_{57 6}$		Gauss	-0.156	N/A	Gauss	-0.235	N/A	Gauss	-0.176	N/A
$c_{68 7}$		t	-0.183	4.977	Gumbel	1.111	N/A	Gauss	-0.312	N/A
$c_{79 8}$		t	-0.22	5.915	Gauss	0.081	N/A	Gauss	-0.38	N/A
T_3		$c_{14 23}$	t	-0.092	8.904	Frank	0.896	N/A	Clayton	0.205
	$c_{25 34}$	Frank	-1.06	N/A	Gauss	-0.275	N/A	Frank	-1.321	N/A
	$c_{36 45}$	t	0.024	5.505	Frank	1.612	N/A	Clayton	0.254	N/A
	$c_{47 56}$	Clayton	0.123	N/A	Clayton	0.129	N/A	Frank	1.727	N/A
	$c_{58 67}$	Frank	-0.54	N/A	Gauss	-0.28	N/A	t	-0.077	4.124
	$c_{69 78}$	t	-0.136	6.542	Frank	0.458	N/A	Frank	0.852	N/A
	T_4	$c_{15 234}$	t	-0.14	5.405	Gumbel	1.135	N/A	Gauss	-0.144
$c_{26 345}$		t	-0.036	6.62	Frank	0.436	N/A	Gumbel	1.062	N/A
$c_{37 456}$		Clayton	0.048	N/A	Gauss	-0.05	N/A	Frank	1.54	N/A
$c_{48 567}$		Frank	-0.284	N/A	t	0	2.683	Gumbel	1.088	N/A
$c_{59 678}$		Gauss	-0.09	N/A	Frank	1.711	N/A	Frank	-1.234	N/A
T_5		$c_{16 2345}$	Gauss	-0.04	N/A	Frank	-1.037	N/A	Gauss	-0.169
	$c_{27 3456}$	Clayton	0.079	N/A	Gauss	-0.01	N/A	Gumbel	1.085	N/A
	$c_{38 4567}$	Gumbel	1.051	N/A	Frank	-0.239	N/A	Gauss	0.066	N/A
	$c_{49 5678}$	Gumbel	1.03	N/A	Frank	-0.035	N/A	Frank	1.266	N/A
	T_6	$c_{17 23456}$	Gauss	-0.151	N/A	Frank	-1.531	N/A	Frank	-1.3
$c_{28 34567}$		t	0.215	5.361	Frank	-1.941	N/A	Frank	0.38	N/A
$c_{39 45678}$		Frank	0.204	N/A	Gumbel	1.047	N/A	Gauss	-0.056	N/A
T_7		$c_{18 234567}$	Frank	-0.347	N/A	Frank	0.685	N/A	Gauss	0.071
	$c_{29 345678}$	Gauss	0.134	N/A	Gumbel	1.012	N/A	Gumbel	1.138	N/A
	T_8	$c_{19 2345678}$	Frank	-0.686	N/A	t	-0.066	7.495	Gauss	-0.025

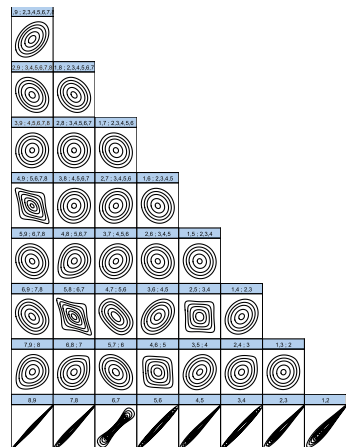
Table 5 D-vine parameter estimates for the Education in all the 8 trees (T_1 the first tree, T_8 the last tree). From the left to the right: the whole countries, only European countries, and only African countries. The selected copulas family (*Cop*: Gauss, t-Student (labeled as t), Clayton, Frank, and Gumbel copula) and the copulas parameters (*par.1* and *par.2*) can be one or two according to the type of copulas. *N/A* identifies absence of one parameter)

	All Countries				Europe			Africa		
	<i>Cop</i>	<i>fam</i>	<i>par.1</i>	<i>par.2</i>	<i>fam</i>	<i>par.1</i>	<i>par.2</i>	<i>fam.</i>	<i>par.1</i>	<i>par.2</i>
T_1	c_{12}	t	0.997	2.338	t	0.987	2	Frank	48.894	N/A
	c_{23}	t	0.993	3.254	Frank	22.929	N/A	t	0.987	3.657
	c_{34}	t	0.988	3.498	Gauss	0.962	N/A	t	0.983	4.346
	c_{45}	t	0.984	4.887	Gauss	0.946	N/A	t	0.98	4.539
	c_{56}	t	0.984	5.703	Gauss	0.944	N/A	t	0.976	5.602
	c_{67}	t	0.983	5.699	Gauss	0.933	N/A	t	0.975	4.788
	c_{78}	t	0.984	5.267	Gauss	0.944	N/A	t	0.971	4.557
	c_{89}	t	0.984	4.921	Gauss	0.943	N/A	t	0.971	5.321
	T_2	$c_{13 2}$	t	0.815	3,154	Gauss	0.73	N/A	Frank	9.483
$c_{24 3}$		t	0.664	3.199	Gauss	0.435	N/A	Frank	6.249	N/A
$c_{35 4}$		t	0.581	3.279	Clayton	0.543	N/A	t	0.599	4.585
$c_{46 5}$		t	0.524	5.224	Clayton	0.826	N/A	Gumbel	1.363	N/A
$c_{57 6}$		t	0.492	5.319	Clayton	0.677	N/A	Gauss	0.52	N/A
$c_{68 7}$		t	0.47	5.468	Clayton	0.727	N/A	Gauss	0.479	N/A
$c_{79 8}$		t	0.464	5.789	Clayton	0.657	N/A	Gauss	0.462	N/A
T_3		$c_{14 23}$	Frank	5.762	N/A	Frank	5.948	N/A	t	0.739
	$c_{25 34}$	Frank	4.92	N/A	t	0.668	2	Gumbel	1.515	N/A
	$c_{36 45}$	Frank	4.991	N/A	Frank	4.976	N/A	Gumbel	1.458	N/A
	$c_{47 56}$	Frank	4.134	N/A	t	0,613	2.591	Gauss	0.502	N/A
	$c_{58 67}$	Frank	3.876	N/A	t	0,591	2.926	Gauss	0.505	N/A
	$c_{69 78}$	Frank	3.946	N/A	Gauss	0.658	N/A	Gauss	0.501	N/A
T_4	$c_{15 234}$	t	0.858	3.251	Gauss	0.846	N/A	Frank	7.072	N/A
	$c_{26 345}$	t	0.699	3.79	Gumbel	1.994	N/A	Frank	6.179	N/A
	$c_{37 456}$	t	0.628	4.807	Gauss	0.668	N/A	t	0.636	4.359
	$c_{48 567}$	t	0.555	4.859	Gauss	0,66	N/A	Clayton	0.936	N/A
	$c_{59 678}$	t	0.536	4.825	Gauss	0,65	N/A	Clayton	0.887	N/A
T_5	$c_{16 2345}$	t	0.686	4.949	Frank	4.768	N/A	Clayton	2.558	N/A
	$c_{27 3456}$	t	0.576	5.233	Clayton	0.739	N/A	Frank	2.721	N/A
	$c_{38 4567}$	t	0.517	5.45	Frank	2.231	N/A	Gauss	0.194	N/A
	$c_{49 5678}$	t	0,489	5.878	Frank	1.946	N/A	Gauss	0.207	N/A
T_6	$c_{17 23456}$	t	0.787	3.872	t	0.392	2.89	Gumbel	1.411	N/A
	$c_{28 34567}$	t	0.691	5.252	Gumbel	1.359	N/A	Gumbel	1.275	N/A
	$c_{39 45678}$	t	0.699	6.377	Frank	3.973	N/A	Gumbel	1.313	N/A
T_7	$c_{18 234567}$	Frank	9.055	N/A	Gumbel	2.836	N/A	Frank	9.403	N/A
	$c_{29 345678}$	Frank	8.353	N/A	Frank	6.152	N/A	Frank	9.39	N/A
T_8	$c_{19 2345678}$	t	0.935	2	Frank	7.524	N/A	t	0.959	2

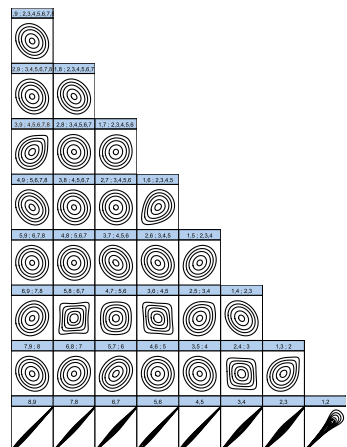
Fig. 14 Contour plots of D-vine Trees (see Table 2) for *HDI*. From the top to the bottom: the whole countries, only European countries and only African countries



All Countries

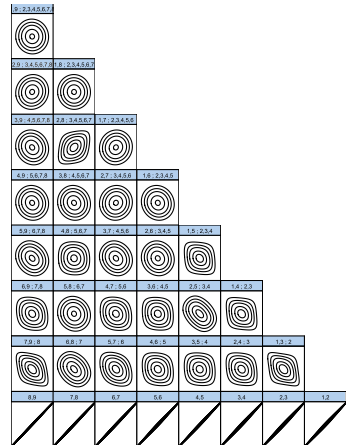


European Countries

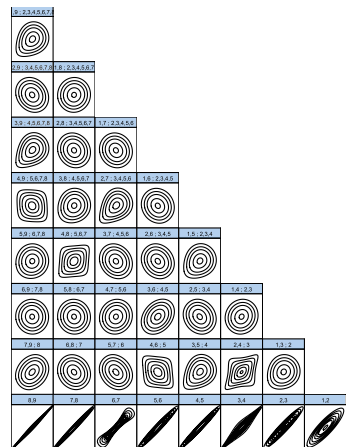


African Countries

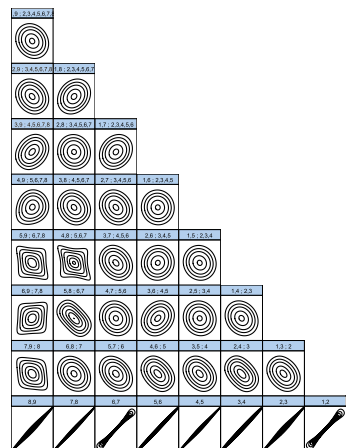
Fig. 15 Contour plots of D-vine Trees (see Table 3) for *Life expectancy*. From the top to the bottom: the whole countries, only European countries and only African countries



All Countries

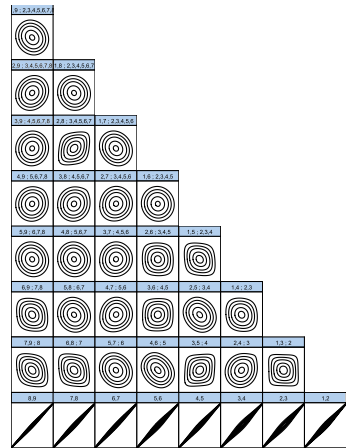


European Countries

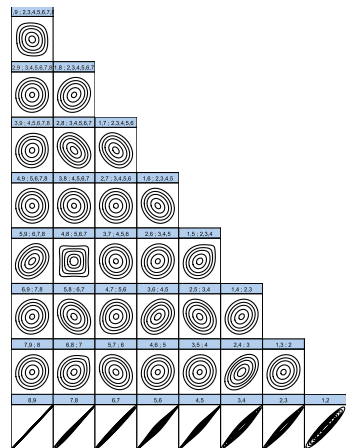


African Countries

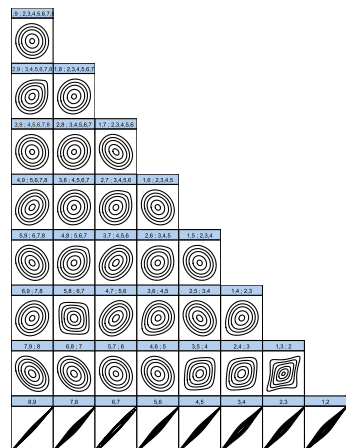
Fig. 16 Contour plots of D-vine Trees (see Table 4) for *Income index*. From the top to the bottom: the whole countries, only European countries and only African countries



All Countries

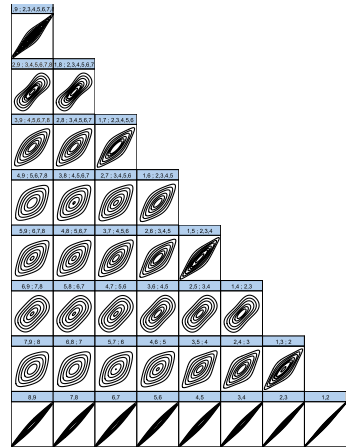


European Countries

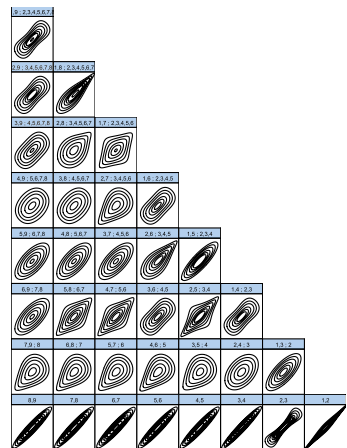


African Countries

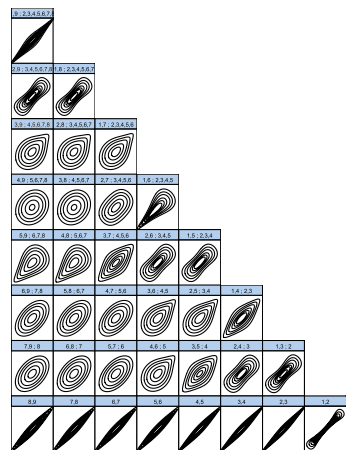
Fig. 17 Contour plots of D-vine Trees (see Table 4) for *Education*. From the top to the bottom: the whole countries, only European countries and only African countries



All Countries



European Countries



African Countries

Funding This work has been supported by Marsden grant E2987-3648 administrated by the Royal Society of New Zealand) and by grant 2017 SGR 622 (GRBIO) administrated by the Departament d'Economia i Coneixement de la Generalitat de Catalunya (Spain). Open access funding provided by Università degli Studi di Genova within the CRUI-CARE Agreement.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Aas, K., Czado, C., Frigessi, A., & Bakken, H. (2009). Pair-copula constructions of multiple dependence. *Insurance: Mathematics and economics*, 44(2), 182–198.
- Akaike, H. (1998). Information theory and an extension of the maximum likelihood principle. In E. Parzen, K. Tanabe, & G. Kitagawa (Eds.), *Selected papers of Hirotugu Akaike* (pp. 199–213). New York: Springer.
- Alijanzadeh, M., Asefzadeh, S., & Moosaniaye, Z. S. A. (2016). Correlation between human development index and infant mortality rate worldwide. *Biotechnology and Health Sciences*, 3(1), 1–5.
- Almasi-Hashiani, A., Sepidarkish, M., Vesali, S., & Omani Samani, R. (2016). The correlation of human development index on fertility and mortality rate: A global ecological study. *International Journal of Pediatrics*, 4(12), 4071–4080.
- Bedford, T., & Cooke, R. M. (2001a). Probability density decomposition for conditionally dependent random variables modeled by vines. *Annals of Mathematics and Artificial intelligence*, 32(1–4), 245–268.
- Bedford, T., & Cooke, R. M. (2001b). Probability density decomposition for conditionally dependent random variables modeled by vines. *Annals of Mathematics and Artificial intelligence*, 32(1), 245–268.
- Bedford, T., & Cooke, R. M. (2002). Vines: A new graphical model for dependent random variables. *Annals of Statistics*, 30(4), 1031–1068.
- Berg, D., & Aas, K. (2009). Models for construction of multivariate dependence: A comparison study. *The European Journal of Finance*, 15(7), 639–659.
- Dißmann, J., Brechmann, E. C., Czado, C., & Kurowicka, D. (2013). Selecting and estimating regular vine copulae and application to financial returns. *Computational Statistics & Data Analysis*, 59, 52–69.
- Genest, C., Ghoudi, K., & Rivest, L.-P. (1995). A semiparametric estimation procedure of dependence parameters in multivariate families of distributions. *Biometrika*, 82(3), 543–552.
- Joe, H. (1997). *Multivariate models and multivariate dependence concepts*. Boca Raton: CRC Press.
- Joe, H. (2014). *Dependence modeling with copulas*. Boca Raton: CRC Press.
- Joe, H., Li, H., & Nikoloulopoulos, A. K. (2010). Tail dependence functions and vine copulas. *Journal of Multivariate Analysis*, 101(1), 252–270.
- Kurowicka, D., & Cooke, R. M. (2006). *Uncertainty analysis with high dimensional dependence modelling*. Hoboken: Wiley.
- Lambert, P., & Vandenhende, F. (2002). A copula-based model for multivariate non-normal longitudinal data: analysis of a dose titration safety study on a new antidepressant. *Statistics in medicine*, 21(21), 3197–3217.
- Lindsey, J. K. (1999). *Models for repeated measurements, Second Edition*. Oxford University Press.
- McGillivray, M. (2005). Measuring non-economic well-being achievement. *Review of Income and Wealth*, 51(2), 337–364.
- Nai Ruscone, M., & Osmetti, S. A. (2016). Modelling the dependence in multivariate longitudinal data by pair copula decomposition. In M. B. Ferraro, P. Giordani, B. Vantaggi, M. Gagolewski, M. A. Gil, P. Grzegorzewski, & O. Hryniewicz (Eds.), *International Conference on Soft Methods in Probability and Statistics* (pp. 373–380). Cham: Springer.
- Nelsen, R. B. (2007). *An introduction to copulas*. New York: Springer.
- Schepsmeier, U., Stoeber, J., Brechmann, E. C., Graeler, B., Nagler, T., Erhardt, T. (2017). *VineCopula: Statistical Inference of Vine Copulas. R package version 2.1.2*. Available on CRAN.

- Shen, C., & Weissfeld, L. (2006). A copula model for repeated measurements with non-ignorable non-monotone missing outcome. *Statistics in Medicine*, 25(14), 2427–2440.
- Shi, P., & Yang, L. (2018). Pair copula constructions for insurance experience rating. *Journal of the American Statistical Association*, 113(521), 122–133.
- Sklar, M. (1959). Fonctions de repartition an dimensions et leurs marges. *Publ. Inst. Statist. Univ. Paris*, 8, 229–231.
- Smith, M., Min, A., Almeida, C., & Czado, C. (2010). Modeling longitudinal data using a pair-copula decomposition of serial dependence. *Journal of the American Statistical Association*, 105(492), 1467–1479.
- Smith, M. S. (2015). Copula modelling of dependence in multivariate time series. *International Journal of Forecasting*, 31(3), 815–833.
- Smith, M. S., & Khaled, M. A. (2012). Estimation of copula models with discrete margins via bayesian data augmentation. *Journal of the American Statistical Association*, 107(497), 290–303.
- Stöber, J., & Czado, C. (2012). *Sampling pair copula constructions with applications to mathematical finance. Simulating Copulas: Stochastic Models, Sampling Algorithms, and Applications*. World Scientific Publishing Co, Singapore.
- Sun, J., Frees, E. W., & Rosenberg, M. A. (2008). Heavy-tailed longitudinal data modeling using copulas. *Insurance: Mathematics and Economics*, 42(2), 817–830.
- Undp (1990). *Human development report 1990: Concept and measurement of human development*. New York. Available at <http://www.hdr.undp.org/en/reports/global/hdr1990>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.