Stochastic Environmental Research and Risk Assessment (2018) 32:2787–2807
https://doi.org/10.1007/s00477-018-1599-9

ORIGINAL PAPER

Vine copula models for predicting water flow discharge at King George Island, Antarctica

Mario Gómez[1] · M. Concepción Ausín[1,2] · M. Carmen Domínguez[3]

Published online: 27 August 2018

Abstract

In order to understand the behavior of the glaciers, their mass balance should be studied. The loss of water produced by melting, known as glacier discharge, is one of the components of this mass balance. In this paper, a vine copula structure is proposed to model the multivariate and nonlinear dependence among the glacier discharge and other related meteorological variables such as temperature, humidity, solar radiation and precipitation. The multivariate distribution of these variables is expressed as a mixture of four components according to the presence or not of positive discharge and/or positive precipitation. Then, each of the four subgroups is modelled with a vine copula. The conditional probability of zero discharge for given meteorological conditions is obtained from the proposed joint distribution. Moreover, the structure of the vine copula allows us to derive the conditional distribution of the glacier discharge for the given meteorological conditions. Three different prediction methods for the values of the discharge are used and compared. The proposed methodology is applied to a large database collected since 2002 by the GLACKMA association from a measurement station located in the King George Island in the Antarctica. Seasonal effects are included by using different parameters for each season. We have found that the proposed vine copula model outperforms a previous work where we only used the temperature to predict the glacier discharge using a time-varying bivariate copula.

Keywords Glacier discharge · Vine copula · Prediction · Meteorological · Finite mixtures

1 Introduction

The study of the mass balance in glaciers is crucial for the correct quantification of water resources (Hamlet and Lettenmaier 1999; Marsh 1999). Mass balance is the difference between accumulation (mainly in form of fallen snow) and ablation (produced by sublimation, calving and melting). Glacier discharge is defined as the rate of flow of meltwater through a vertical section perpendicular to the direction of the flow (Cogley et al. 2011). It is produced by surface runoff or by flowing inside the glacier and exiting through the front or the base.

The Antarctic Peninsula as being affected by recent rapid regional climate warming, which refers to those areas where the regional changes have been more profound than the worldwide mean, as noted by the Intergovernmental Panel on Climate Change (IPCC) (Turner et al. 2005; Vaughan et al. 2003). Ablation periods of the glaciers in this area have been increasing over time (Domínguez and Eraso 2007). As a consequence, glaciers have been retreating and thinning, and surface melting has tended to increase (Rückamp et al. 2011; Barrand et al. 2013).

The study of the relationship between glacier behavior and climate is a fundamental issue in glaciology. These relationships can be analyzed with the energy balance equations which evaluate the most important energy fluxes between the atmosphere and the glacier surface. These

✉ Mario Gómez
mario.gomez@uc3m.es

M. Concepción Ausín
causin@est-econ.uc3m.es

M. Carmen Domínguez
karmenka@usal.es

1 Department of Statistics, Universidad Carlos III de Madrid, Getafe, Spain

2 UC3M-BS Institute of Financial Big Data (IFiBiD), Getafe, Spain

3 Department of Applied Mathematics, University of Salamanca, Salamanca, Spain

equations are computed from physically based calculations (see e.g. Braun 2001; Sicart et al. 2008) and involve complex equations and measurements. Alternatively, temperature index models use only the air temperature to empirically model the relationship between melt rate and (positive) air temperature (degree-day model; Hock 2003). A complete review of these methods can be found in Hock (2003). There are some studies that incorporate more variables into this model, such as the direct solar radiation (Hock 1999) or the albedo and the shortwave radiation (Pellicciotti et al. 2005). The main problem in this type of model is that they implicitly assume a linear relationship between temperature and discharge.

In this paper we propose the use of multivariate copulas to model the non-linear relationships between the discharge, temperature and several other meteorological variables. Copulas are statistical tools within statistics that allow us to model these relationships independently of the marginal distributions choice (Genest and Favre 2007). In climate science and hydrology, many research authors model the relationship between pairs of variables using bidimensional copulas (see e.g. De Michele and Salvadori 2003; Carnicero et al. 2013; Sarhadi et al. 2016). However, the number of publications studying the relationships among more than two variables is much smaller. Standard multivariate copulas are available, such as the multivariate Gaussian or the t-Student copulas. However, they have shown to be rather inflexible as, for example, they assume that the dependence is symmetric in both tails, which is not realistic in this context. Alternatively, it is possible to model multivariate distributions using vine copulas. These are flexible dependence models that are built from recursive conditioning of the underlying joint density functions resulting in a product of conditional and unconditional bivariate copula densities. Vine copulas have been successfully used in a small number of papers in hydrology. For example, Gyasi-Agyei and Melching (2012) model the internal dependence structure between net storm event depth, maximum wet periods depth, and the total wet periods duration. Gyasi-Agyei (2013) models the dependence between total depth, total duration of wet periods, and the maximum proportional depth of a wet period in a rainfall disaggregation model. Xiong et al. (2015) study the dependence between annual maximum daily discharge, annual maximum 3-day flood volume and annual maximum 15-day flood volume to understand the change-point detection of multivariate hydrological series. Note that these papers deal with three hydrological variables, while in our work we study the relation between five variables which, in addition, may take discrete values.

This work has two main goals. First, we wish to predict the conditional probability of having no glacier discharge given the observations of temperature, humidity, radiation and precipitation. Also, we want to predict the values of the discharge given the specific observations of the meteorological variables, with the conditional distribution of the discharge obtained through the vine copula. This paper extends our previous bivariate copula model (Gomez et al. 2017), based on a time-varying relationship between discharge and temperature, by the inclusion of three new meteorological variables which, as in the case of the precipitation, may take zero values. We consider vine copulas to define the dependence structure between all variable of our new model. We also propose a new way of dealing with zero values, in the glacier discharge, compared to previous work in Gomez et al. (2017). These were considered as missing observations, while it is now assumed that the glacier discharge may be equal to zero with positive probability. Another difference is the way in which seasonality has been taken into account. Due to the increase in the number of variables, we have divided each hydrologic year into four periods with different behavior in the discharge regime instead of assuming a time-varying seasonality.

The remainder of the paper is organized as follows. Firstly, copulas, vine copulas and the estimation method of inference function for margins are briefly introduced in Sect. 2. Next, the study area and the considered database are described in Sect. 3. This is followed by the proposal of a multivariate model based on vine copulas and a description of the estimation method in Sect. 4. The proposed methodology is applied, in Sect. 5, to the GLACKMA database. Finally, Sect. 6 includes an extensive discussion about findings, limitations and possible future extensions.

2 Background

Copulas are multivariate distributions defined on the unit hypercube with uniform marginal distributions. Copulas allow to define the dependence structure between random variables independently of their marginal behavior. See Nelsen (2006) for an extensive review. Sklar's theorem (Sklar 1959) proves that for any m-dimensional distribution, F, there exists a m-dimensional copula, C, such that for all $x_1, \ldots, x_m$,

$$F(x_1, \ldots, x_m) = C(F_1(x_1), \ldots, F_m(x_m)), \tag{1}$$

where $F_1, \ldots, F_m$ are the marginal distribution functions corresponding to F. If the margins are continuous, then the copula is unique and the joint density function is,

$$f(x_1, \ldots, x_m) = c(F_1(x_1), \ldots, F_m(x_m)) \prod_{j=1}^{m} f_j(x_j), \tag{2}$$

where c is the m-dimensional copula density and $f_1,\ldots,f_m$ are the marginal density functions.

2.1 Vine copulas

Vine copulas or pair-copula construction is a dependence model to model multivariate data using a product of unconditional and conditional bivariate copulas, called pair-copulas (Aas et al. 2009), and marginal densities. For high dimensions, many different vine structures can be used for the same multivariate density function. Bedford and Cooke (2001) introduced a graphical structure, called regular vine structure, to organize the pair-copula construction as a sequence of nested trees with undirected edges, which will be later illustrated. According to the structure of the trees, we can distinguish two families of regular vines: c-vines and d-vines. D-vines and C-vines have a practical interpretation; D-vine: temporal/serial ordering of variables and C-vine: ordering of the variables by importance. In this paper, we will only consider c-vines for simplicity and also because it seems reasonable to consider the temperature as the main variable in the first tree, as will be later explained. A c-vine can be constructed as follows. Consider a multivariate density as a product of conditional densities,

$$f(x_1,\ldots,x_m) = f_1(x_1)\prod_{j=2}^{m} f_{j|1:j-1}\left(x_j \mid x_1,\ldots,x_{j-1}\right) \quad (3)$$

where, using (2),

$$f_{j|1:j-1}\left(x_j \mid x_{1:j-1}\right) = c_{j-1,j|1:j-2}\left(F_{j-1|1:j-2}(x_{j-1}|x_{1:j-2}),\ F_{j|1:j-2}(x_j|x_{1:j-2})\right) f_{j|1:j-2}(x_j|x_{1:j-2}), \quad (4)$$

where $F(\cdot|\cdot)$ and $f(\cdot|\cdot)$ denote conditional cdfs and densities, respectively, and $x_{1:j-1} = \{x_1,\ldots,x_{j-1}\}$. Then, applying (4) recursively, the joint density (3) can be expressed as:

$$f(x_1,\ldots,x_m) = \prod_{j=1}^{m} f_j(x_j)\prod_{j=2}^{m}\prod_{k=1}^{j-1} c_{j-k,j|1:j-k-1}\left(F_{j-k|1:j-k-1}(x_{j-k}|x_{1:j-k-1}), F_{j|1:j-k-1}(x_j|x_{1:j-k-1})\right), \quad (5)$$

where the conditional distribution functions in (5) can be obtained recursively using the following result obtained by Joe (1996),

$$F(x_{j-k}|x_{1:j-k-1}) = \frac{\partial C_{j-k,j-k-1|1:j-k-2}\left(F_{j-k|1:j-k-2}(x_{j-k}|x_{1:j-k-2}), F_{j-k-1|1:j-k-2}(x_{j-k-1}|x_{1:j-k-2})\right)}{\partial F_{j-k-1|1:j-k-2}(x_{j-k-1}|x_{1:j-k-2})}. \quad (6)$$

This derivative, which is usually called h function, has been derived explicitly for many Archimedean and elliptical copulas, see Aas et al. (2009).

2.2 Inference function for margins method

Consider a copula-based parametric model,

$$F(x_1,\ldots,x_m;\nu_1,\ldots,\nu_m,\theta_c) = C(F_1(x_1;\nu_1),\ldots,F_m(x_m;\nu_m);\theta_c), \quad (7)$$

where $(\nu_1,\ldots,\nu_m)$ denote the marginal parameters and θ_c are the copula parameters. Maximum likelihood estimation of both marginal and copula parameters may become very complicated, especially when the dimension m is high. Note that given a data set, $\{(x_{i1},\ldots,x_{im})\}$, for $i=1,\ldots,n$, the log-likelihood can be written as, see (2),

$$L(\nu_1,\ldots,\nu_m,\theta_c) = \sum_{i=1}^{n}\log c(F_1(x_{i1};\nu_1),\ldots,F_m(x_{im};\nu_m);\theta_c) + \sum_{i=1}^{n}\sum_{j=1}^{m}\log f_j(x_{ij};\nu_j),$$

Alternatively, Joe (1997) proposes the so-called Inference Function for Margins (IFM) method. The IFM method is a two-step estimation approach where, in the first step, the marginal parameters are separately estimated as if the variables were independent,

$$\hat{\nu}_j = \arg\max_{\nu_j}\sum_{i=1}^{n}\log f_j(x_{ij};\nu_j), \quad \text{for, } j=1,\ldots,m.$$

Then, in a second step, the copula parameters are estimated by,

$$\hat{\theta}_c = \arg\max_{\theta_c}\sum_{i=1}^{n}\log c\left(\hat{F}_1(x_{i1};\hat{\nu}_1),\ldots,\hat{F}_m(x_{im};\hat{\nu}_m);\theta_c\right),$$

where $\hat{\nu}_j$, for $j=1,\ldots,m$, are the estimated marginal parameters from the first step and $\hat{F}_i$ indicates that the parametric forms for marginal distributions are estimates themselves and not only the corresponding parameter estimates.

Note that the probability integral transformed values based on the marginal estimates, $\hat{F}_j(x_i;\hat{\nu}_j)$, are approximately standard uniformly distributed if the parametric fits are sufficiently good, otherwise the copula parameter estimates will be biased.

As usually considered in vine copulas, we impose the so-called simplifying assumption in (4) which assumes that each pair-copula of conditional distributions are independent of he variables on which they are conditioned. Although, this assumption has been criticized (Acar et al. 2012 and Spanhel and Kurz 2015). Killiches et al. (2017) propose the use of this assumption especially when the number of parameters is large and this assumption simply is common practice in vine copula modeling. Moreover, Haff et al. (2010) came to the conclusion that vine copulas built with this assumption are "a rather good solution, even when the simplifying assumption is far from being fulfilled by the actual model".

3 Study area and data

King George is the largest of the South Shetland Islands, which is an archipelago located off the coast of the Antarctic Peninsula in the Southern Ocean. See Braun (2001) for a complete description of the island. The study area is located in the southwestern part of King George island, where GLACKMA has placed one of their eight Pilot Experimental Catchment Areas, called CPE-KG-62°S whose coordinates are Lat:S 62°11'035, Long:W 58°54'414, so as to study the discharge of the Collins Glacier, see www.glackma.es. Specific glacier discharge per unit area, measured in $m^3 \cdot s^{-1} \cdot km^{-2}$, is estimated as an exponential function of the level of the river which receives the melted water from the catchment area. Also, we have selected a collection of meteorological variables such as the air temperature ($^\circ C$), the relative humidity (%), the solar radiation (W/m^2) and the precipitation (mm). These meteorological data have been provided by the Bellinghausen Russian base (via GLACKMA), sited near of the catchment area. See Domínguez and Eraso (2007) and Gomez et al. (2017) for a description of the catchment area and further details about the variables.

The available data are from 10/01/2002 to 09/30/2012 and are composed by daily measurements of the mean daily temperature, humidity, radiation and discharge and the daily cumulative precipitation. A preliminary study of these data shows that discharge and precipitation have a large number of zero-values. In particular, the value of the discharge was zero in 62% of the observed days and the value of the precipitation was zero in 31% of the observed days. This fact will have a definite impact in the design of the vine copula model. Figure 1 shows the scatter plot of each couple of variables and the histogram of each individual variable for those days when the discharge was larger than zero. Apparently, there are relationships between the variables also these relations seem not to be linear. We therefore suggest the use of copulas to model these non-linear relationships. The Kendall rank correlation coefficients of the variables help us to choose the order of the variables in the vine structure.

4 Methodology

In this section, we introduce a method to predict values of the specific glacier discharge per unit area given the observed values of the meteorological variables. First, we propose a vine copula model to describe the multivariate joint density function of the five variables where two of them have a large number of zero values. Then, we obtain the conditional probability of having no discharge. Finally, we derive the conditional distribution of the discharge given the other meteorological variables.

4.1 Multivariate copula model

Let T, H, R, P and D be random variables, where the temperature is $T \in \mathbb{R}$; the humidity is $H \in (0,1)$; the radiation is $R \in \mathbb{R}^+$; the precipitation is $P \in \mathbb{R}^+ \cup \{0\}$ and the discharge is $D \in \mathbb{R}^+ \cup \{0\}$. As commented in the previous section, in practice, it is observed that both, the precipitation and the discharge, show a large number of zero values. This fact has a quite important impact in the construction of our proposed model. We define the joint density function as a mixture of four different multivariate densities, depending on the joint probability of observing zero or positive values for the discharge and precipitation. Thus, the joint density function of the multivariate variable (T, H, R, P, D) is decomposed as,

$$f(t,h,r,p,d) = \begin{cases} f^{00}_{thr}(t,h,r), & \text{with } p^{00} = \Pr(D=0, P=0), \quad (8a) \\ f^{10}_{thrd}(t,h,r,d), & \text{with } p^{10} = \Pr(D>0, P=0), \quad (8b) \\ f^{01}_{thrp}(t,h,r,p), & \text{with } p^{01} = \Pr(D=0, P>0), \quad (8c) \\ f^{11}_{thrpd}(t,h,r,p,d), & \text{with } p^{11} = \Pr(D>0, P>0), \quad (8d) \end{cases}$$

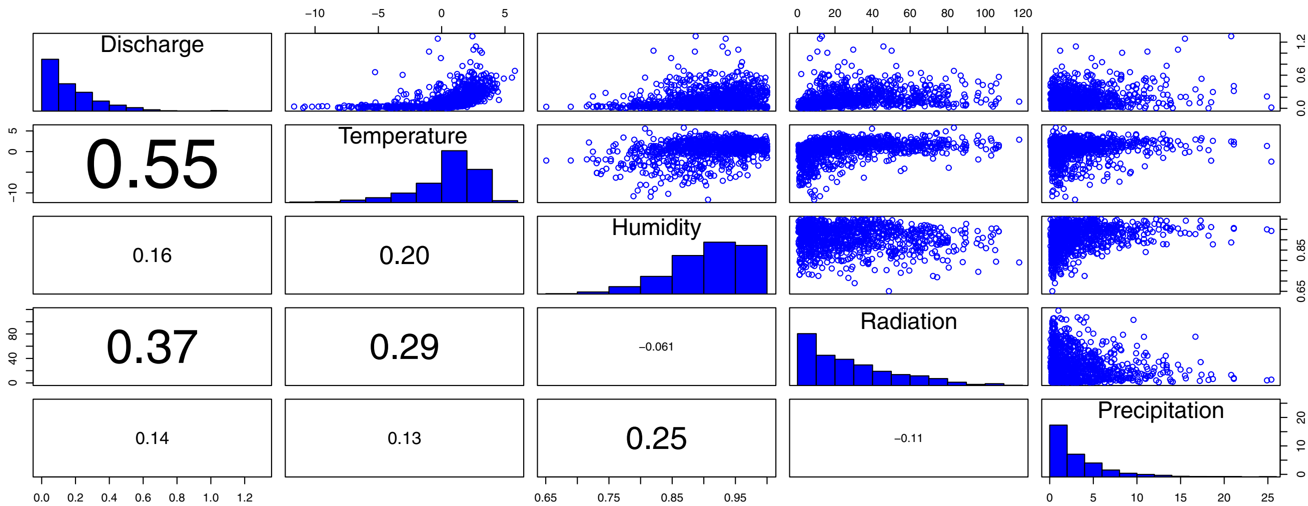

Fig. 1 Scatter plot, histograms and Kendall's τ of the five variables when the values of the discharge and precipitation are greater than zero. Sizes of the τ values are proportional to their absolute values

where f^{00}_{thr}, f^{10}_{thrd}, f^{01}_{thrp} and f^{11}_{thrpd} denote each of the four different multivariate density functions, where the first and second superscripts refer, respectively, to the presence (1) or not (0) of positive discharge and precipitation. In practice, given a data set for the five variables (T, H, R, P, D), we will divide the data into four subsamples according to having or not discharge and precipitation. The joint probabilities of having or not discharge and precipitation, p^{jk}, for $j = 0, 1$ and $k = 0, 1$, will be obtained empirically. Then, the four subsamples will be modeled independently allowing for different marginal parameters and c-vine copula models constructed from different pair-copula families, as explained below.

Our proposed model builds upon the ideas by Erhardt and Czado (2012) and Brechmann et al. (2014) who propose a zero-inflated copula-based model where marginal variables may take zero or positive values. However, our approach is different since they use, depending on each subgroup, a multivariate margin of the joint density. By doing so, in some way they assume the same probabilistic law for all subgroups and therefore, all observations (independently of their membership to a specific subgroup) contribute to the estimation of joint copula parameters. Alternatively, we allow for different multivariate distribution models for each subgroup.

We can define each of the four joint density functions, (8a) to (8d), in terms of c-vine copulas. For example, the multivariate density for the group with highest dimension, (8d), can be expressed as, see (3),

$$f^{11}_{thrpd}(t,h,r,p,d) = f^{11}_{thrpd}(d \mid t,h,r,p) \cdot f^{11}_{thrp}(p \mid t,h,r) \cdot f^{11}_{thr}(r \mid t,h) \cdot f^{11}_{th}(h \mid t) \cdot f^{11}_{t}(t), \tag{9}$$

where $f^{11}_{thrp}(.)$ is the four dimensional margin of f^{11}_{thrpd}, $f^{11}_{thrp}(.|.)$ are the univariate conditional densities and so on. We use this decomposition because conditioning is in the order of the importance of variables and corresponds to the main nodes in the C-vine tree levels.

By the Sklar theorem (2), we know that,

$$f^{11}_{th}(h \mid t) = c^{11}_{th}\left(F^{11}_{t}(t), F^{11}_{h}(h)\right) \cdot f^{11}_{h}(h), \tag{10}$$

and similarly, see (4),

$$\begin{aligned} f^{11}_{thr}(r \mid t,h) &= c^{11}_{hr|t}\left(F^{11}_{th}(h \mid t), F^{11}_{tr}(r \mid t)\right) \cdot f^{11}_{tr}(r \mid t) \\ &\overset{(10)}{=} c^{11}_{hr|t}\left(F^{11}_{th}(h \mid t), F^{11}_{tr}(r \mid t)\right) \\ &\quad \cdot c^{11}_{tr}\left(F^{11}_{t}(t), F^{11}_{r}(r)\right) \cdot f^{11}_{r}(r), \end{aligned} \tag{11}$$

and, see (4),

$$\begin{aligned} f^{11}_{thrp}(p \mid t,h,r) = {} & c^{11}_{rp|th}\left(F^{11}_{thr}(r \mid t,h), F^{11}_{thp}(p \mid t,h)\right) \\ & \cdot f^{11}_{thp}(p \mid t,h) \overset{(11)}{=} c^{11}_{rp|th}\left(F^{11}_{thr}(r \mid t,h), F^{11}_{thp}(p \mid t,h)\right) \\ & \cdot c^{11}_{hp|t}\left(F^{11}_{th}(h \mid t), F^{11}_{tp}(p \mid t)\right) \\ & \cdot c^{11}_{tp}\left(F^{11}_{t}(t), F^{11}_{p}(p)\right) \cdot f^{11}_{p}(p), \end{aligned} \tag{12}$$

and, see (4),

$$\begin{aligned} f^{11}_{thrpd}(d \mid t,h,r,p) = {} & c^{11}_{pd|thr}\left(F^{11}_{thrp}(p \mid t,h,r), F^{11}_{thrd}(d \mid t,h,r)\right) \\ & \cdot f^{11}_{thrd}(d \mid t,h,r) \overset{(12)}{=} c^{11}_{pd|thr}\left(F^{11}_{thrp}(p \mid t,h,r),\right. \\ & \left. F^{11}_{thrd}(d \mid t,h,r)\right) \cdot c^{11}_{rd|th}\left(F^{11}_{thr}(r \mid t,h),\right. \\ & \left. F^{11}_{thd}(d \mid t,h)\right) \cdot c^{11}_{hd|t}\left(F^{11}_{th}(h \mid t), F^{11}_{td}(d \mid t)\right) \\ & \cdot c^{11}_{td}\left(F^{11}_{t}(t), F^{11}_{d}(d)\right) \cdot f^{11}_{d}(d). \end{aligned} \tag{13}$$

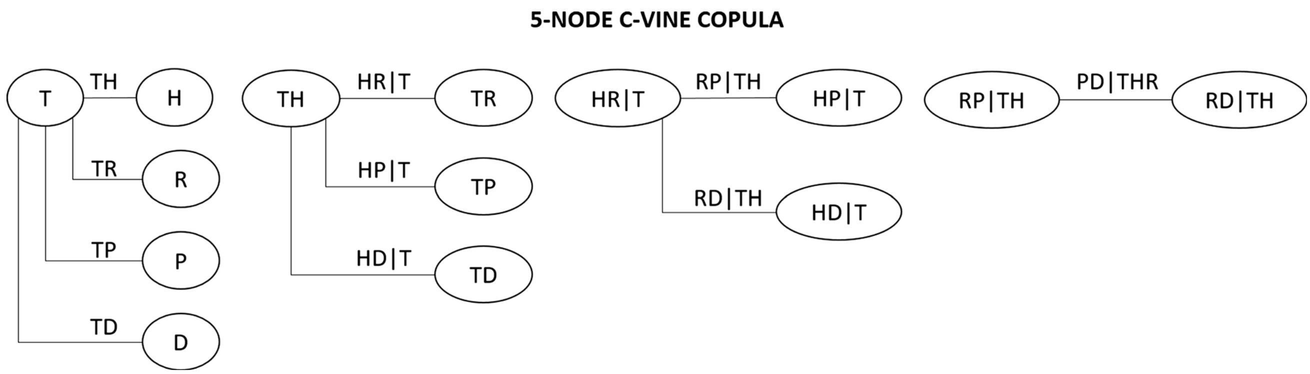

Fig. 2 Structure of a c-vine copula with 5 variables, 4 trees and 10 edges, where T is the main variable for the first tree. Each edge may be associated with a pair-copula

Then, we can apply (10), (11), (12) and (13) in (9) to obtain the multivariate density function as the product of bivariate conditional and unconditional copula densities and marginal density functions,

$$\begin{aligned} f^{11}_{thrpd}(t,h,r,p,d) = & f^{11}_t(t)\cdot f^{11}_h(h)\cdot f^{11}_r(r)\cdot f^{11}_p(p)\cdot f^{11}_d(d)\\ & \cdot c^{11}_{th}(u^{11}_t,u^{11}_h)\cdot c^{11}_{tr}(u^{11}_t,u^{11}_r)\cdot c^{11}_{tp}(u^{11}_t,u^{11}_p)\\ & \cdot c^{11}_{td}(u^{11}_t,u^{11}_d)\cdot c^{11}_{hr|t}(u^{11}_{h|t},u^{11}_{r|t})\cdot c^{11}_{hp|t}(u^{11}_{h|t},u^{11}_{p|t})\\ & \cdot c^{11}_{hd|t}(u^{11}_{h|t},u^{11}_{d|t})\cdot c^{11}_{rp|th}(u^{11}_{r|th},u^{11}_{p|th})\\ & \cdot c^{11}_{rd|th}(u^{11}_{r|th},u^{11}_{d|th})\cdot c^{11}_{pd|thr}(u^{11}_{p|thr},u^{11}_{d|thr}), \end{aligned} \tag{14}$$

where, using (6), we have that,

$$u^{11}_x = F^{11}_x(x), \quad \text{for } x = t,h,r,p,d,$$

$$u^{11}_{x|t} = F^{11}_{tx}(x \mid t) = \frac{\partial C^{11}_{tx}(u^{11}_t,u^{11}_x)}{\partial u^{11}_t}, \quad \text{for } x = h,r,p,d,$$

$$u^{11}_{x|th} = F^{11}_{thx}(x \mid t,h) = \frac{\partial C^{11}_{hx|t}(u^{11}_{h|t},u^{11}_{x|t})}{\partial u^{11}_{h|t}}, \quad \text{for } x = r,p,d,$$

$$u^{11}_{x|thr} = F^{11}_{thrx}(x \mid t,h,r) = \frac{\partial C^{11}_{rx|th}(u^{11}_{r|th},u^{11}_{x|th})}{\partial u^{11}_{r|th}}, \quad \text{for } x = p,d.$$

Figure 2 shows the graphical structure of the 5-dimensional c-vine of the density (14). Each line in (14) corresponds to one of the four nested trees in the figure, except for the first line which corresponds to the marginal densities. Each edge connects two nodes whose relationship can be modeled with a bivariate copula.

Similar expressions can be obtained for (8a), (8b) and (8c), which are shown in Appendix 1.

4.2 Marginal distributions

Now, we wish to define the marginal distribution function for each of the five variables, (T,H,R,P,D), in each of the four groups, (j,k), for $j = 0,1$ and $k = 0,1$, according to the presence or not of discharge and precipitation. Thus, each marginal distribution will be defined as a mixture of four components. For example, the density function of the temperature will be expressed as:

$$f_t(t) = \begin{cases} f^{00}_t(t), \text{with } p^{00},\\ f^{10}_t(t), \text{with } p^{10},\\ f^{01}_t(t), \text{with } p^{01},\\ f^{11}_t(t), \text{with } p^{11}. \end{cases}$$

Note that, as commented before, in practice we will divide the whole data set in four subsamples for the marginal distribution according to having or not discharge and precipitation. Therefore, we will have four subsamples from the marginal distribution of each variable. Thus, we will allow a different marginal distribution with different parameters for each of the four groups in each variable.

Further, we assume a parametric model based on finite mixture models for the density in each of the four groups. In particular, for the temperature, we consider a finite mixture of Gaussian distributions, see e.g. Schär et al. (2004),

$$f^{jk}_t(t) = \sum_{i=1}^{K^{jk}_t} \omega^{jk}_{ti} \cdot \frac{1}{\sqrt{2\pi[\sigma^{jk}_{ti}]^2}} \cdot \exp\left(-\frac{[t-\mu^{jk}_{ti}]^2}{2[\sigma^{jk}_{ti}]^2}\right),$$

where we allow for a different number of mixture components, K^{jk}_t, in each group (j,k), for $j = 0,1$ and $k = 0,1$. The mixture size, K^{jk}_t, will be selected using model selection criteria, as will be explained in Sect. 4.4. Given the mixture size, the marginal model parameters are $\left\{\omega^{jk}_{ti},\mu^{jk}_{ti},\sigma^{jk}_{ti}\right\}$, for $i = 1,\ldots,K^{jk}_t$; for $j = 0,1$ and $k = 0,1$.

Similarly, for the variable defining the humidity, we consider a finite mixture of Beta distributions, see e.g. Yao (1974) and Yang et al. (2015), since these are defined between 0 and 1,

$$f_h^{jk}(h) = \sum_{i=1}^{K_h^{jk}} \omega_{hi}^{jk} \cdot \frac{\Gamma(\alpha_{hi}^{jk} + \beta_{hi}^{jk})}{\Gamma(\alpha_{hi}^{jk}) \cdot \Gamma(\beta_{hi}^{jk})} \cdot h^{\alpha_{hi}^{jk}-1} \cdot (1-h)^{\beta_{hi}^{jk}-1},$$

where the marginal parameters are $\left\{\omega_{hi}^{jk}, \alpha_{hi}^{jk}, \beta_{hi}^{jk}\right\}$, for $i = 1, \ldots, K_h^{jk}$; for $j = 0, 1$ and $k = 0, 1$, and where the mixture size, K_h^{jk}, will be obtained using model selection criteria.

Finally, finite mixtures of Gamma distributions are assumed for the radiation, precipitation (Scholzel and Friederichs 2008) and discharge (Favre et al. 2004 and Gomez et al. 2017), since they are defined for positive values. In particular, for the radiation,

$$f_r^{jk}(r) = \sum_{i=1}^{K_r^{jk}} \omega_{ri}^{jk} \cdot \frac{[\lambda_{ri}^{jk}]^{\kappa_{ri}^{jk}}}{\Gamma(\kappa_{ri}^{jk})} \cdot r^{\kappa_{ri}^{jk}-1} \cdot \exp(-\lambda_{xi}^{jk} \cdot r), \tag{15}$$

where the marginal parameters are $\left\{\omega_{ri}^{jk}, \kappa_{ri}^{jk}, \lambda_{ri}^{jk}\right\}$, for $i = 1, \ldots, K_r^{jk}$; for $j = 0, 1$, and $k = 0, 1$. For the precipitation, we only need to define continuous densities for two groups since,

$$f_p(p) = \begin{cases} \delta(p), & \text{with} \quad p^{00} + p^{10}, \\ f_p^{01}(p), & \text{with} \quad p^{01}, \\ f_p^{11}(p), & \text{with} \quad p^{11}, \end{cases} \tag{16}$$

where $\delta(p)$ denotes the Dirac delta function which is zero everywhere except at $p = 0$, where it is infinite. Thus, we assume a different Gamma mixture for each $f_p^{01}(p)$ and $f_p^{11}(p)$, analogous to the one given in (15), with marginal parameters $\left\{\omega_{pi}^{j1}, \kappa_{pi}^{j1}, \lambda_{pi}^{j1}\right\}$, for $i = 1, \ldots, K_p^{j1}$; for $j = 0, 1$, where K_p^{j1} is the number of terms in the mixture. Finally, analogous to (16), the discharge is equal to zero with probability $p^{00} + p^{01}$, and follows a different Gamma mixture, as given in (15), for each $f_d^{10}(d)$ and $f_d^{11}(d)$, with probabilities p^{10} and p^{11}, respectively, and marginal parameters $\left\{\omega_{di}^{1k}, \kappa_{di}^{1k}, \lambda_{di}^{1k}\right\}$, for $i = 1, \ldots, K_p^{1k}$; for $k = 0, 1$, where K_d^{1k} is the number of terms in the mixture.

In practice, once we have an estimated mixture density, f_x^{jk}, for each variable, $x = t, h, r, p, d$, and group, $j = 1, 0$ and $k = 0, 1$, we may apply the corresponding distribution function, F_x^{jk}, to the subsample of data in each group in order to obtain the pseudo-u data, which will be approximately uniformly distributed for each group and for each variable.

4.3 Conditional probability

Once we have defined the complete multivariate model (5) in terms of copulas and marginal densities, we may obtain many quantities of interest. For example, we may derive the conditional probability of zero discharge for one particular day whose meteorological variables have been observed. Assume first that zero precipitation has been observed, then the conditional probability of zero discharge given the values of the remaining meteorological variables is,

$$\Pr(D = 0 \mid T = t, H = h, R = r, P = 0) = \frac{p^{00} f_{thr}^{00}(t, h, r)}{p^{00} f_{thr}^{00}(t, h, r) + p^{10} f_{thr}^{10}(t, h, r)}, \tag{17}$$

where p^{00}, p^{10} and $f_{thr}^{00}(t, h, r)$ are directly obtained from the model (5) and

$$f_{thr}^{10}(t, h, r) = c_{hr|t}^{10}\left(u_{h|t}^{10}, u_{r|t}^{10}\right) \cdot c_{th}^{10}\left(u_t^{10}, u_h^{10}\right) \cdot c_{tr}^{10}\left(u_t^{10}, u_r^{10}\right) \cdot f_t^{10}(t) \cdot f_h^{10}(h) \cdot f_r^{10}(r),$$

which is the marginal joint density of the triple $(t,\ h,\ r)$ from the joint density $f_{thrd}^{10}(t, h, r, d)$, given by (27). Figure 3 illustrates how to visualize the marginal density, $f_{thr}^{10}(t, h, r)$, from the joint density $f_{thrd}^{10}(t, h, r, d)$ by deleting all the nodes related with the variable discharge. Note that these can be viewed as nested c-vine copulas which allow to derive the dependence structure of marginal distributions given the c-vine structure of the joint distribution.

Now, consider the case where a positive precipitation, $p > 0$, has been observed. Then, the conditional probability of zero discharge given this positive precipitation and the values of the remaining meteorological variables is,

$$\Pr(D = 0 \mid T = t, H = h, R = r, P = p) = \frac{p^{01} f_{thrp}^{01}(t, h, r, p)}{p^{01} f_{thrp}^{01}(t, h, r, p) + p^{11} f_{thrp}^{11}(t, h, r, p)}, \tag{18}$$

where p^{01}, p^{11} and $f_{thrp}^{01}(t, h, r, p)$ are again directly obtained from the model (5) and

$$f_{thrp}^{11}(t, h, r, p) = c_{rp|th}^{11}\left(u_{hr|t}^{11}, u_{hp|t}^{11}\right) \cdot c_{hr|t}^{11}\left(u_{h|t}^{11}, u_{r|t}^{11}\right) \cdot c_{hp|t}^{11}\left(u_{h|t}^{11}, u_{p|t}^{11}\right) \cdot c_{th}^{11}\left(u_t^{11}, u_h^{11}\right) \cdot c_{tr}^{11}\left(u_t^{11}, u_r^{11}\right) \cdot c_{tp}^{11}\left(u_t^{11}, u_p^{11}\right) \cdot f_t^{11}(t) \cdot f_h^{11}(h) \cdot f_r^{11}(r) \cdot f_p^{11}(p),$$

which is the marginal joint density of the quadruple $(t,\ h,\ r,\ p)$ from the joint density $f_{thrpd}^{11}(t, h, r, p, d)$, given by (14). As before, we may visualize the marginal density $f_{thrp}^{11}(t, h, r, p)$ from the joint density $f_{thrpd}^{11}(t, h, r, p, d)$ by deleting the nodes related with the variable discharge in Fig. 2.

Furthermore, given the complete model (5), we may also obtain the whole conditional distribution function of the discharge given observed values of the remaining meteorological variables. As before, we may distinguish two cases according to having observed zero or positive precipitation. In the first case, the conditional distribution function is given by,

$$F(d \mid T=t, H=h, R=r, P=0) = \begin{cases} 0, & \text{with } \Pr(D=0 \mid T=t, H=h, R=r, P=0) \\ F^{10}_{thrd}(d \mid t, h, r), & \text{with } 1-\Pr(D=0 \mid T=t, H=h, R=r, P=0) \end{cases} \tag{19}$$

where $\Pr(D=0 \mid t, h, r, P=0)$ is given in (17) and $F^{10}_{thrd}(d \mid t, h, r)$ is the conditional cdf of the discharge obtained from the density, $f^{10}_{thrd}(t, h, r, d)$, given in (27), such that,

$$F^{10}_{thrd}(d \mid t, h, r) = \frac{\partial C^{10}_{rd|th}\left(u^{10}_{r|th}, u^{10}_{d|th}\right)}{\partial u^{10}_{r|th}}$$

which can be obtained directly from the bivariate copula in the last tree of Fig. 3.

Finally, the whole conditional distribution function of the discharge given a positive precipitation, $p > 0$, and observed values for the remaining meteorological variables is given by,

$$F(d \mid T=t, H=h, R=r, P=p) = \begin{cases} 0, & \text{with } \Pr(D=0 \mid T=t, H=h, R=r, P=p) \\ F^{11}_{thrdp}(d \mid t, h, r, p), & \text{with } 1-\Pr(D=0 \mid T=t, H=h, R=r, P=p) \end{cases} \tag{20}$$

where $\Pr(D=0 \mid t, h, r, P=p)$ is given in (18) and $F^{11}_{thrdp}(d \mid t, h, r, p)$ is the conditional cdf of the discharge in $f^{11}_{thrdp}(t, h, r, p, d)$, whose density is given in (14), such that,

$$F^{11}_{thrdp}(d \mid t, h, r, p) = \frac{\partial C^{11}_{pd|thr}\left(u^{11}_{p|thr}, u^{11}_{d|thr}\right)}{\partial u^{11}_{p|thr}}$$

which can be obtained directly from the bivariate copula in the last tree of Fig. 2.

4.4 Inference and prediction methods

Here, we provide a description of the methods that we propose for the estimation of the model parameters, model selection and prediction of discharge values. Assume that we have a data set from the five variables (T, H, R, P, D) and we want to estimate the parameters for our proposed model (5). As commented before, we firstly divide the sample in four subsamples according to the presence or not of discharge and/or precipitation and estimate the probabilities of each group, p^{jk}, for $j = 0, 1$ and $k = 0, 1$, using empirical frequencies. Then, for each subsample, we consider the IFM method, which is reviewed in Sect. 2. Thus, we firstly estimate the parameters of the marginal distributions for each variable in each group based on mixture models, as described in Sect. 4.2, using the MLE method. Note that the marginal parameters can be different in each group, but the estimation procedure is the same. In order to select the best number of components in each mixture, we make use of the Bayesian Information Criterion (BIC) which requires to estimate the parameters by MLE for different number of mixture components. We select the one with the minimum BIC value for each variable in each group. Finally, we apply the estimated marginal distribution functions, F^{jk}_x, for each variable, $x = t, h, r, p, d$, and for each group, $j = 0, 1$, and $k = 0, 1$ to the observed data, such that an estimation of the u-data values u^{jk}_x, are obtained for each variable and group.

The next step is to fit a c-vine copula for each u-data subsample in each group. Firstly, it is required to set an order for the variables. As described in Sect. 4.1, we have chosen the same order for the variables in the four groups so that the structure of the trees is the same. Figure 7 shows this structure. We have set the variables with strongest dependencies in the first nodes of the trees. In particular, the temperature has been set as the main variable since it shows the strongest Kendall's tau dependence with the discharge. Also, the discharge has been set as the last variable for convenience in the computation of conditional probabilities. Clearly, there are many other possible orders and some of them will be analyzed and compared in the application. Some discussion about the order issue is also included in Sect. 6.

Given the structure of the c-vines, appropriate pair-copula families are selected and estimated sequentially as follows. Parameters are estimated by maximum likelihood and the best copula family is selected using the BIC value.

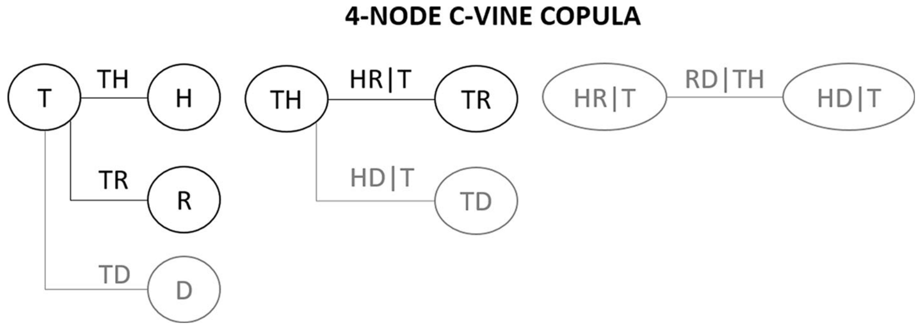

Fig. 3 Structure of a c-vine copula with 3 nodes inherited from a c-vine copula with 4 nodes

Consider the u-data values, u_x^{jk}, for each variable, $x = t, h, r, p, d$, and for each group, $j = 0, 1$, and $k = 0, 1$.

1. Fit bivariate copulas for u_t^{jk} and u_x^{jk}, for $x = h, r, p, d$, for all the edges in the first tree.
2. Generate the series $u_{x|t}^{jk} = \frac{\partial C_{tx}^{jk}(u_t^{jk}, u_x^{jk})}{\partial u_t^{jk}}$, for $x = h, r, p, d$, using the fitted copula from the previous step.
3. Fit bivariate copulas for $u_{h|t}^{jk}$ and $u_{x|t}^{jk}$, for $x = r, p, d$ for all the edges in the second tree.
4. Using the same procedure, generate series from the edges and fit copulas between the nodes for the remaining trees.

In the application, we have chosen among elliptical copulas (Gaussian, t-copula), one-parameter Archimedean copulas (Gumbel, Frank, Joe, Clayton) and their rotated versions. Before selecting a copula, the Kendall's tau independence test is performed for each pair of variables using a 0.05 significance level. All these estimations have been made using the functions available in the R package *VineCopula* (Schepsmeier et al. 2017; R Core Team 2018).

We consider and compare three different estimators to obtain predictions for values of the discharge given observed values of the remaining meteorological variables. The first estimator is the median of the conditional distribution of the discharge. This can be calculated as the value, $\hat{d}$, such that,

$$0.5 = Pr(D = 0 \mid t, h, r, p) + (1 - Pr(D = 0 \mid t, h, r, p)) \cdot F_d(\hat{d} \mid t, h, r, p),$$

where F_d is the conditional distribution function given in (19), for the case that the observed precipitation is zero, $p = 0$, or F_d is given by (20), for the case that the observed precipitation is positive, $p > 0$. As a second estimator, we consider the expected value of the conditional distribution of the discharge. This can be approximated using a Monte Carlo simulation by taking the sample mean of a set of simulated values from (19) or (20), respectively, depending if zero or positive precipitation has been observed. This procedure is detailed in Appendix 1. Finally, we propose a third prediction method based on a decision rule where we first obtain the conditional probability of zero discharge, (17) or (18), according to the observed precipitation. If this probability is larger than 0.5, our prediction for the discharge in that day is zero. Otherwise, we estimate the expected value of the conditional distribution for the discharge using a Monte Carlo simulation as before.

In order to examine the performance of our predictions, we first consider the Brier Score (Brier 1950) to evaluate the accuracy of the estimated probability of zero discharge. The Brier score measures the distance between the estimated probability and the true observed value of an event,

$$BS = \frac{1}{n}\sum_{i=1}^{n}(p_i - o_i)^2, \tag{21}$$

where n the prediction horizon, p_i is the probability that the event will happen and o_i takes value 1 if the event happens and 0 otherwise. We also examine the performance of the predicted amount of discharge with the three different prediction methods using the Mean Squared Error (MSE) and the Mean Absolute Error (MAE),

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(\hat{d}_i - d_i)^2, \tag{22}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n} \mid \hat{d}_i - d_i \mid, \tag{23}$$

where $\hat{d}_i$ is the estimated value and d_i is the true observed value.

5 Application

In this section, our proposed vine copula model is applied to the data provided by GLACKMA from their catchment area in glacier Collins within King George Island. First, the database is divided in groups according to four different hydrological periods to capture seasonality. Second, the model parameters, both the marginal distribution parameters and vine copula parameters, are estimated. Third, the conditional probability of having no discharge and the

Table 1 Definition of discharge periods in King George Island

Period	Dates	Description
1	26th November–30th December	Discharge start period. Since the last weeks of spring to early summer. The glacier discharge can be positive or zero
2	31st December–7th April	Main discharge period. Most of the summer. The glacier discharge is positive almost every day
3	8th April–15th June	Discharge end period. Since the end of summer and most of autumn. The glacier discharge can be zero or positive
4	16th June–25th November	Zero discharge period. Late autumn, all the austral winter and early spring. The glacier discharge is always zero

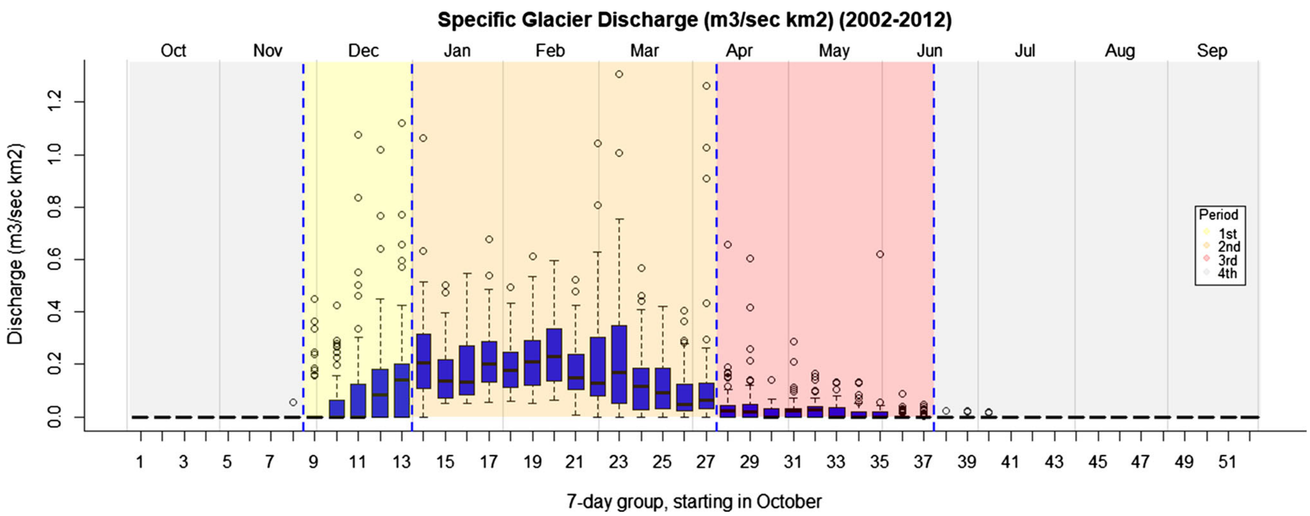

Fig. 4 Boxplots of the glacier discharge in each week from 2002 to 2012. Different periods are separated by vertical lines and different color shadows

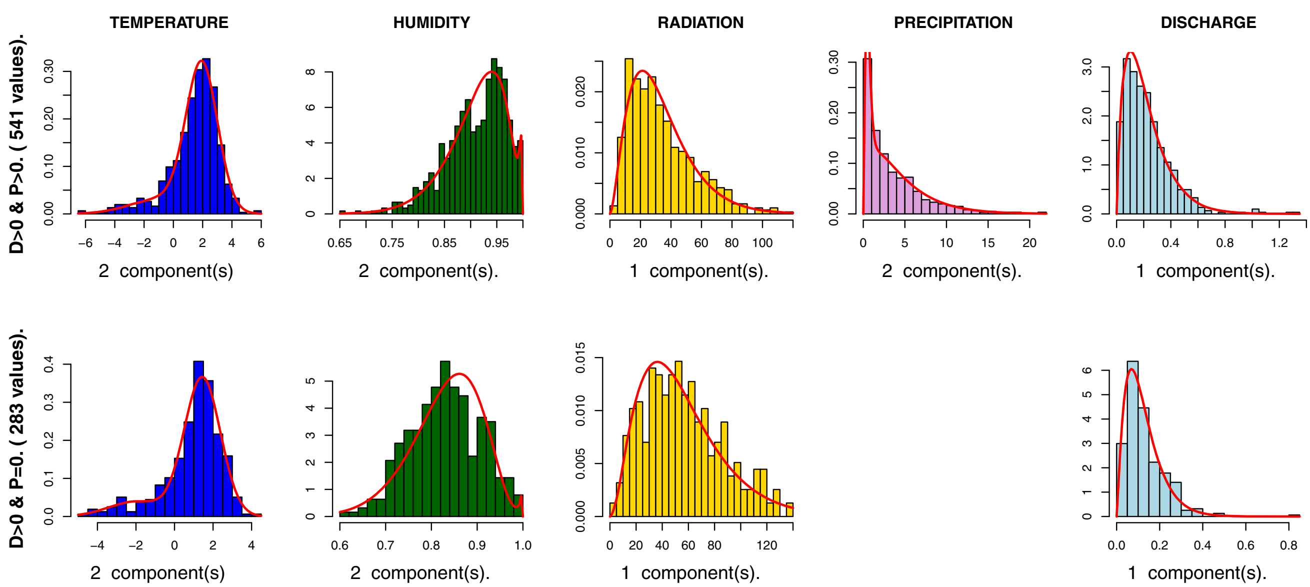

Fig. 5 Histograms of the observed data compared with the estimated mixture densities for each marginal variable, obtained for the second period and for the group with positive discharge and positive precipitation (top) and the group with positive discharge and zero precipitation (bottom). The number of mixture components is shown at the bottom of each plot

Table 2 Mixture parameter estimation

Per	Gr. DP	N	Temperature ω_1	μ_1	σ_1	μ_2	σ_2
1	00	40		−0.809 (0.191)	1.206 (0.135)		
	01	94		−0.163 (0.111)	1.074 (0.078)		
	10	68	0.220 (0.247)	−0.862 (1.513)	1.133 (0.651)	0.904 (0.170)	0.651 (0.129)
	11	83		0.915 (0.142)	1.295 (0.101)		
2	00	21	0.621 (0.106)	−2.370 (0.234)	0.829 (0.176)	0.665 (0.164)	0.460 (0.115)
	01	16		−1.298 (0.486)	1.944 (0.344)		
	10	283	0.145 (0.056)	−1.954 (0.790)	1.460 (0.430)	1.444 (0.101)	0.937 (0.073)
	11	541	0.177 (0.070)	−0.693 (0.814)	2.189 (0.266)	1.941 (0.082)	1.071 (0.077)
3	00	102		−5.375 (0.378)	3.817 (0.267)		
	01	225	0.665 (0.067)	−6.636 (0.499)	3.523 (0.262)	−0.992 (0.271)	1.230 (0.205)
	10	69		−2.363 (0.288)	2.392 (0.204)		
	11	234	0.517 (0.053)	−3.633 (0.377)	2.759 (0.208)	0.359 (0.107)	0.833 (0.086)

Per	Gr. DP	N	Humidity w_1	α_1	β_1	α_2	β_2
1	00	40		23.527 (5.314)	4.582 (0.99)		
	01	94		31.247 (4.716)	2.421 (0.332)		
	10	68		15.254 (2.648)	3.472 (0.570)		
	11	83		24.086 (3.847)	2.589 (0.379)		
2	00	21		40.476 (12.524)	10.439 (3.173)		
	01	16		16.284 (5.840)	3.326 (1.123)		
	10	283	0.029 (0.019)	6.016 (6.587)	0.240 (0.129)	19.277 (2.065)	3.957 (0.442)
	11	541	0.967 (0.012)	24.565 (1.859)	2.504 (0.192)	305.374 (195.601)	1.829 (0.723)
3	00	102		19.921 (2.829)	3.662 (0.492)		
	01	225		22.324 (2.15)	3.027 (0.271)		
	10	69		14.021 (2.431)	2.783 (0.449)		
	11	234		18.336 (1.761)	2.022 (0.174)		

Per	Gr. DP	N	Radiation ρ_1	α_1	β_1	α_2	β_2
1	00	40	0.455 (0.09)	69.832 (33.129)	0.605 (0.279)	86.14 (34.047)	1.146 (0.466)
	01	94		12.000 (1.726)	0.184 (0.027)		
	10	68		13.508 (2.287)	0.148 (0.026)		
	11	83		6.229 (0.942)	0.104 (0.016)		
2	00	21		6.980 (2.104)	0.236 (0.074)		
	01	16		2.543 (0.846)	0.103 (0.038)		
	10	283		2.913 (0.232)	0.053 (0.005)		
	11	541		2.721 (0.156)	0.081 (0.005)		

Table 2 (continued)

Per	Gr. DP	N	Radiation ρ_1	α_1	β_1	α_2	β_2
3	00	102	0.591 (0.071)	6.542 (1.492)	2.707 (0.729)	4.308 (1.864)	0.381 (0.142)
	01	225	0.536 (0.150)	4.347 (1.553)	1.945 (0.879)	1.858 (0.456)	0.284 (0.052)
	10	69		1.608 (0.251)	0.184 (0.034)		
	11	234		1.838 (0.157)	0.339 (0.033)		

Per	Gr. DP	N	Precipitation ρ_1	α_1	β_1	α_2	β_2
1	01	40	0.607 (0.188)	2.06 (0.597)	1.753 (0.896)	1.381 (0.513)	0.277 (0.093)
	11	94		1.242 (0.173)	0.583 (0.099)		
2	01	21		0.731 (0.221)	0.213 (0.090)		
	11	16	0.235 (0.057)	3.487 (0.960)	5.825 (2.337)	1.482 (0.192)	0.326 (0.032)
3	01	102	0.219 (0.051)	7.817 (2.848)	17.953 (7.749)	1.490 (0.186)	0.548 (0.067)
	11	225	0.245 (0.079)	3.503 (1.205)	5.018 (2.611)	1.406 (0.196)	0.354 (0.045)

Per	Gr. DP	N	Discharge ρ_1	α_1	β_1	α_2	β_2
1	10	40		1.788 (0.283)	11.055 (2.015)		
	11	94		1.640 (0.233)	8.372 (1.390)		
2	10	21		2.247 (0.177)	18.094 (1.593)		
	11	16		1.871 (0.105)	8.525 (0.549)		
3	10	102	0.936 (0.035)	4.029 (0.787)	127.662 (28.612)	18.01 (16.619)	139.819 (121.816)
	11	225	0.355 (0.054)	0.986 (0.143)	9.555 (1.899)	6.744 (1.214)	221.226 (43.461)

The first column indicates the period, the second refers to the group (j, k), where $j = 0, 1$, respectively, for zero or positive discharge and $k = 0, 1$, respectively, for zero or positive precipitation. Third column shows the number of observed values in each group. Standard errors are shown in parentheses

complete predictive discharge distribution is obtained using the estimated vine copula model. Finally, the obtained results are compared with those obtained with the bivariate copula model of our previous work (Gomez et al. 2017).

5.1 Parameter estimation

Recall that the GLACKMA database consists of five time series of data collected during eleven years. Here, the first ten years are used for parameter identification and data from 01/10/2011 to 30/09/2012 are used for model verification.

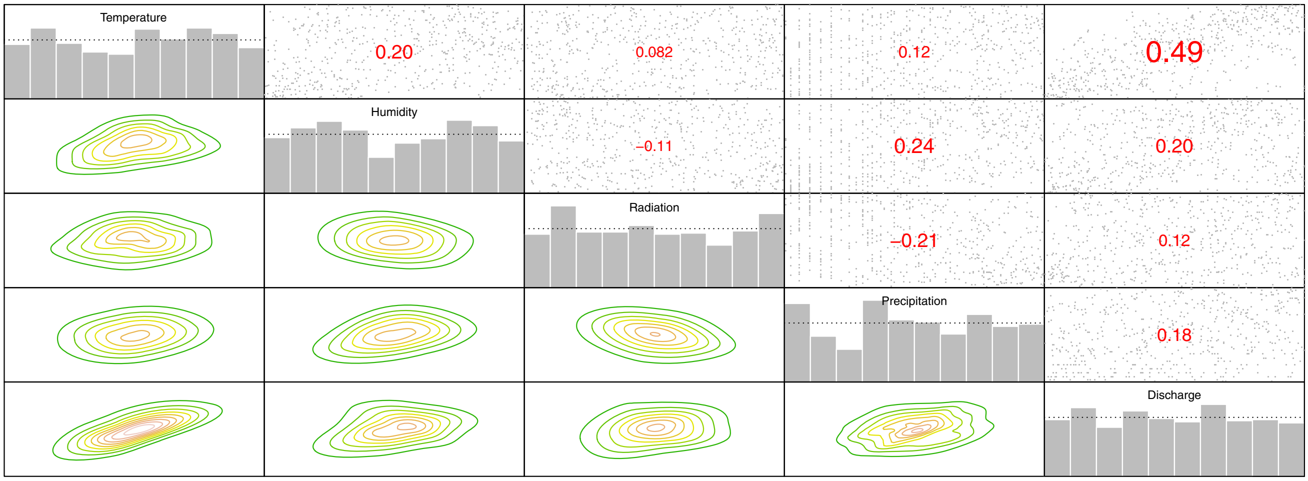

Fig. 6 Histograms of the pseudo-u data obtained with the estimated mixture cdf for each marginal variable, scatter plots for each pair of u-data and contour plots of the fitted bivariate copula densities obtained for the second period and for the group with positive discharge and positive precipitation

First of all, we want to capture the seasonal behavior of the discharge. In our previous work (Gomez et al. 2017), we try to capture it using partial sums of Fourier terms, but this procedure increases rapidly the number of parameters when we add more meteorological variables. On the other side, Braun (2001) has found three major ablation phases plus a non-ablation phase for each year in glacier behavior. This suggests us to divide the data in four different periods in order to capture the changes in the relationship between the variables. Table 1 shows the different periods selected for this study. As a justification of this division, Fig. 4 shows the boxplots of the average daily glacier, grouped by weeks, in the different periods. Apparently, there are different behaviors in the discharge regime. Remember that the fourth period has zero discharge in the observed values. Thus, the model will always predict zero discharge in this period, that is, equations (17) and (18) will always be one independently of the values of the other variables because the estimated probabilities of positive discharge, p^{10} and p^{11}, which are based on empirical frequencies, will be both equal to zero during this period.

Firstly, we determine the number of components and the mixture parameters of the marginal models for the first three periods in each of the four groups according to the presence or not of discharge and precipitation. As an example, Fig. 5 shows the adjustment of the mixtures to the observations of the five variables in the second period for the two groups with positive discharge. The number of selected mixture components is shown at the bottom of each plot. An apparently good adjustment between the mixture models and the empirical distributions is observed for all variables in both groups. The mixture marginal distribution parameter values are listed in Table 2. The first column indicates the period, the second refers to the group (j, k), where $j = 0, 1$, respectively, for zero or positive discharge and $k = 0, 1$, respectively, for zero or positive precipitation, and the third shows the number of observations available and used to fit the mixtures. Finally, using the marginal parameter estimates, we apply the mixture cdf to the subsample of data for each variable and group to obtain the pseudo-u data, which should be approximately uniformly distributed. Figure 6 (diagonal) shows the histograms of the obtained pseudo-u data for the second period and for the group with positive discharge and positive precipitation. We may observe that these are approximately uniformly distributed. Similar results are obtained for other groups and periods. Figure 6 (upper diagonal) also shows the scatter plots for each pair of u-data which offer an idea about the underlying dependence structure. For example, we may observe that there exists lower tail dependence between temperature and humidity for this group and period. Figure 6 (lower diagonal) also shows the contour plots of the fitted bivariate copulas obtained for each case using the following analysis.

The next step is to select the copula family and estimate its parameters for each edge in the vine copula structures. Figure 7 shows the structure of the c-vine copulas with the value of the parameter for every edge; each row in each edge corresponds to one of the first three periods. As you can see, that some edges have the independence copula, denoted by the letter **I**, this means that no significant dependence is found between the variables associated to this edge. BIC has been considered to compare different variable orders in the c-vine structure. Very close values were found. Also, we have used the Vuong test (Vuong 1989) to look for differences between different orders, but no significative difference has been found. Table 3 shows some of the results obtained for the c-vine copula for days with positive discharge and precipitation (5 nodes) in the

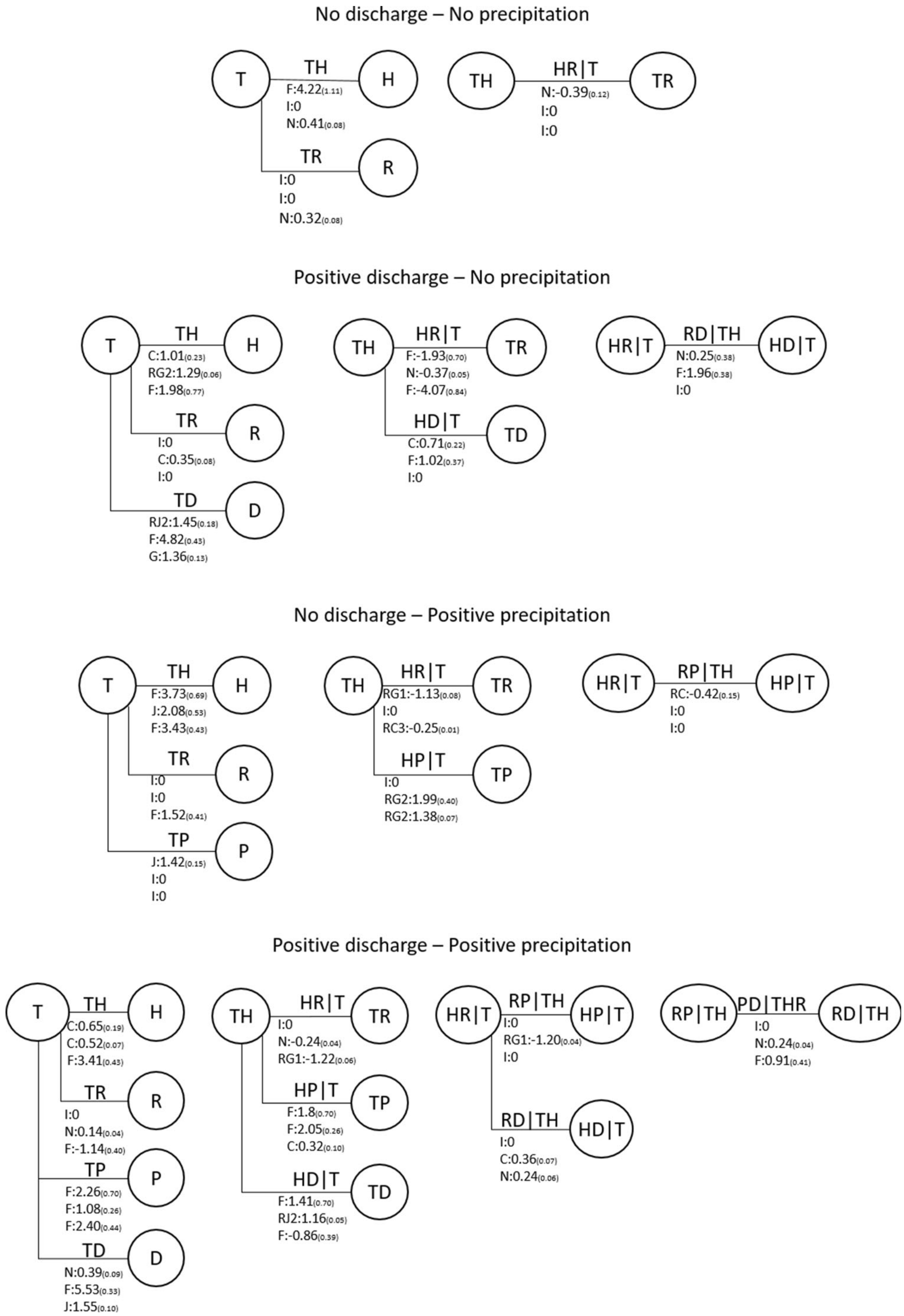

Fig. 7 Estimation of c-vine copula parameters for all periods and groupsEach row in each edge corresponds to one of the first three periods. Selected copulas are *I* , Independent, *N* Gaussian, *C* Clayton, *G* Gumbel, *F* Frank, *J* Joe, RC3 = Clayton rotated 270°, RG1 = Gumbel rotated 90°, RG2 = Gumbel rotated 180°, RJ2 = Joe rotated 180°. Standard errors are shown in parentheses

first period. The results for other groups and periods are similar. We have selected the temperature as the main node since it shows the strongest dependence with the discharge (as considered in e.g. Pham et al. 2016) and the discharge in the last place since this is the more convenient choice to facilitate the evaluation of the probability of discharge and the predicted discharge. Also, a goodness-of-fit test has been performed for each of the twelve c-vine copulas obtained with the proposed order. This test is based on the information matrix equality of White, as detailed in Schepsmeier (2016). Table 4 shows the White statistic and the corresponding *p* value for each c-vine copula. Both the model and the parameters seem to be appropriate.

Further, in order to examine the goodness of fit of the estimated bivariate copulas in each c-vine structure, we make use of the $\lambda-$function (Genest and Rivest 1993). The

Table 3 BIC value of different order combinations for the 5-cvine copula in the first period

Order	BIC	Vuong statistic	p value
THRPD	−20.331	0	1
TDPRH	−20.172	0.05	0.96
HTRPD	−16.079	1.121	0.262
HDRPT	−16.085	0.829	0.407
RPDTH	−16.641	0.797	0.426
RTDPH	−21.057	−0.19	0.849
PTRHD	−15.842	0.904	0.366
PDTRH	−15.427	1.008	0.314
DPRHT	−14.579	0.999	0.318
DTHRP	−16.275	0.746	0.456

Vuong test of comparison with the selected order (THRPD) and the corresponding p value

Table 4 White statistic and p value to test the goodness-of-fit over the twelve c-vine copulas for the selected order

Group	Period 1		Period 2		Period 3	
	White	p value	White	p value	White	p value
00	02.13	0.18	09.25	0.32	08.51	0.60
01	21.54	0.17	19.09	0.58	17.80	0.37
10	19.20	0.82	16.00	0.33	21.76	0.45
11	60.83	0.88	72.91	0.81	73.62	0.06

λ-function is characteristic for each copula family and is defined as:

$$\lambda(v,\theta) = v - K(v,\theta), \tag{24}$$

where $K(v,\theta) = P\big(C(u_1,u_2|\theta) \le v\big)$ is the Kendall's distribution function for the copula C with parameter θ, $v \in [0,1]$ and (u_1,u_2) is distributed according to C. Comparing empirical to theoretical λ-functions gives a method to examine if the selected copula might be appropriate to describe the observed dependence. As an illustration, Fig. 8 shows the comparison between the empirical and theoretical $\lambda-$function for each of the ten edges of the selected c-vine copula for the second period and for the group of data with positive values of discharge and precipitation. Apparently, there is a good fit between the selected and the empirical copula in all edges. Similar results have been obtained for all the other selected vine structures, which are not shown for the sake of brevity.

Table 5 Comparison of predicted with observed number of days with zero and positive discharge. On the left, for the in-sample data (2002–2011) used to estimate the model. On the right, for the out-of-sample data (2011–2012) used to validate the model

		Predicted					
		2002–2011			2011–2012		
		$D=0$	$D>0$	Total	$D=0$	$D>0$	Total
Observed	$D=0$	1886	146	2032	204	23	227
	$D>0$	150	1145	1295	13	126	139

5.2 Conditional probability of zero discharge

Once we have obtained all the model parameters, we are interested in estimating the probability of having zero discharge conditioned on the observed values of temperature, humidity, radiation and precipitation on the corresponding day. As described in Sect. 4.4, we will predict positive discharge for a particular day if the estimated conditional probability of zero discharge, see (17) and (18), is smaller than 0.5. Table 5 compares the predicted with the observed number of

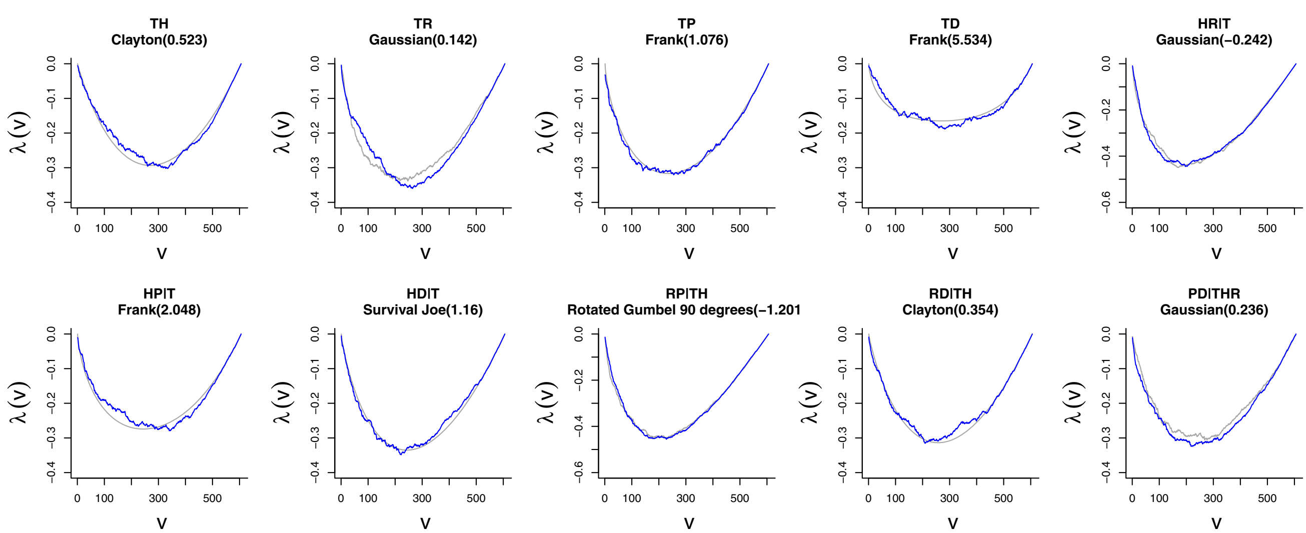

Fig. 8 Empirical λ-function for the ten edges of the c-vine for the second period and for the group with positive discharge and precipitation. The blue and grey lines correspond, respectively, to the empirical and theoretical functions

Table 6 Comparison of the Brier scores obtained with a logistic regression and with the proposed vine copula model

	2002–2011		2011–2012	
	Logistic model	Vine model	Logistic model	Vine model
Global	0.0643	0.0605	0.0815	0.0772
Period 1	0.1401	0.1309	0.2474	0.2984
Period 2	0.0359	0.0317	0.0254	0.0235
Period 3	0.1999	0.1901	0.1861	0.1476

On the left, for the in-sample data (2002–2011) used to estimate the model. On the right, for the out-of-sample data (2011–2012) used to validate the model

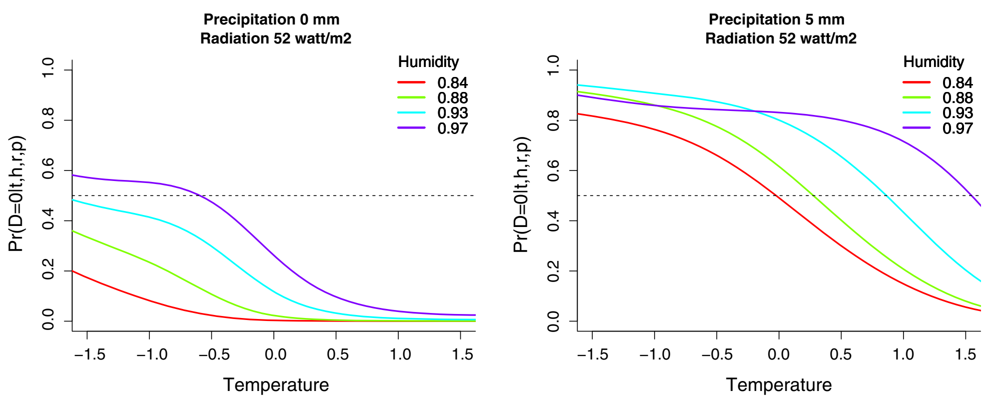

Fig. 9 Estimated probability of zero discharge as a function of the temperature for different values of the humidity, in the presence or absence of precipitation and a fixed value for the radiation, obtained for the first period

days with zero and positive discharge. For the in-sample data (2002–2011), we obtain that 92.8% of days with observed zero discharge are correctly predicted with the c-vine model, whereas 88.4% of days with observed positive discharge are correctly predicted. The performance of the copula model is even better for the out-of-sample data from the last hydrological year, used to validate the model. Our model has a 89.8% and 90.6% of correctly predicted days for days with observed zero and positive discharge respectively. We have compared these probabilities with the ones obtained with a logistic regression, which has been developed under the same conditions as the vine model, that is, using a different specification for each group and period. Table 6 shows the Brier scores (21) for both models. We can see that the vine copula model outperforms the logistic regression, globally (except for the first period for the out-of-sample data) and for each period, and for the in-sample and the out-of-sample data. The smaller the Brier Score, the better the predictions.

Additionally, we are also interested in studying the behavior of the conditional probability of zero discharge in terms of the meteorological variables. As an illustration, Fig. 9 shows the estimated probability of zero discharge as a function of the temperature for different values of the humidity, in the presence or absence of precipitation and a fixed value for the radiation. Note that the positive precipitation increases the probability of zero discharge, especially when the temperatures are below zero. In both plots we can also see that higher temperatures cause a decay in the probability of having zero discharge and that an increase of the percentage of humidity increases the probability of having zero discharge. Similar plots can be done to compare how the same meteorological conditions in different periods modify the probability of zero discharge. This kind of plots provide an interesting tool for analysing the influence of meteorological conditions in glacier discharge under different meteorological scenarios.

5.3 Predicted discharge

Finally, as described in Sect. 4.3, a predicted value of the discharge can be obtained for all days using the predictive discharge distribution, see (19) and (20). We use the three prediction methods explained in Sect. 4.4 based on the predictive mean, median and a decision rule. These predictions have been compared with those obtained with our previous bivariate copula model (Gomez et al. 2017). Table 7 shows the MSE, see (22), and the MAE, see (23), obtained for both models. We may observe that in all cases the mean errors obtained for the c-vine copula model are smaller than those obtained with the bivariate copula model. Then, clearly, the c-vine copula model gives more accurate predictions of the discharge. Finally, the proposed model is validated with the observed values of the last year (2011–12). The two last columns of Table 7 show the MSE and MAE for the out-of-

Table 7 Mean squared and absolute errors for the predicted value for the discharge obtained for the proposed vine copula model and the bivariate copula from a previous work, for the in-sample (first two columns) and out-of-sample (last two columns) data

		2002–2011		2011–2012	
Model	Method	MSE	MAE	MSE	MAE
Vine copula model	Median	0.00620	0.02800	0.01214	0.04685
	Mean	0.00606	0.03179	0.01074	0.04858
	Decision rule	0.00607	0.03019	0.01078	0.04574
Bivariate copula model		0.00718	0.03317	0.03753	0.05362

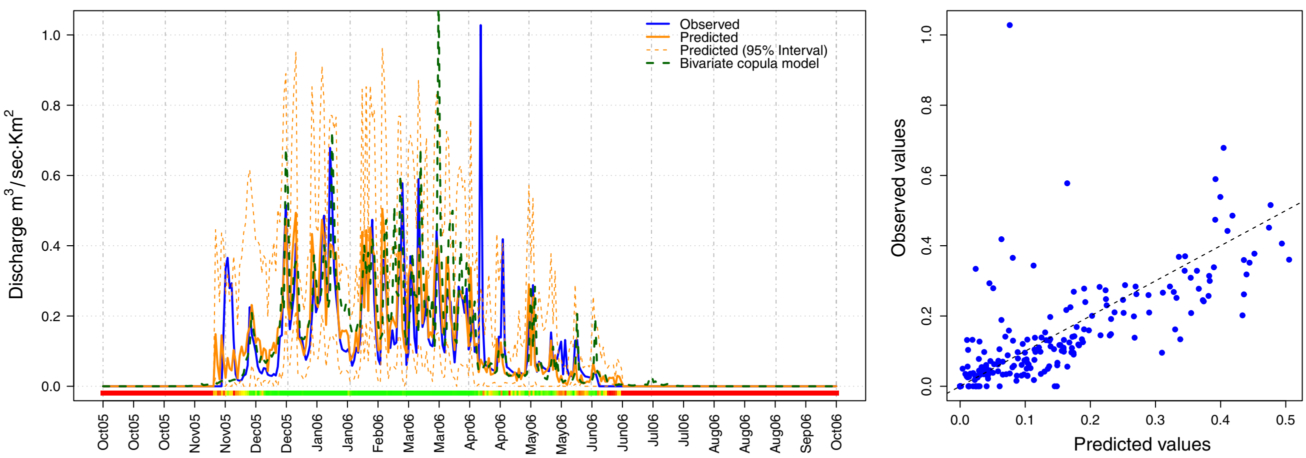

Fig. 10 Left panel shows the time series of the observed and predicted values for the discharge with the bivariate and vine copula models for the year 2005–06. Also shown are 95% predictive intervals for the vine copula model. At bottom of the plot, the conditional probability of zero discharge for each day in a scale from red (probability one) to green (probability zero). Right panel contains the comparison between predicted and observed values for the proposed vine copula model

sample period. Again, we can observe that the mean errors of the proposed model are smaller than those produced by the bivariate copula model. Therefore, it can be concluded that the use of more meteorological variables in the proposed vine copula model provides more accurate predictions then using only the temperature as in our previous model.

As an example, Fig. 10 (left panel) shows the observed values of the discharge for the year 2005–06 compared with the predictive discharge obtained with the proposed c-vine copula model, together with the corresponding 95% confidence intervals, and the predictions obtained with the bivariate copula model. Also, it is illustrated the conditional probability of zero discharge from red, for probability one, to green, for probability zero. Figure 10 (right panel) shows the scatter plot for the predicted and the observed values of the discharge for that year. We can conclude that the c-vine copula model provides a good fit to the observations.

6 Discussion

In this section, we interpret and discuss the present findings, limitations and possible future extensions. In this work, we have constructed a multivariate model based on c-vine copulas to describe the relationship between the specific glacier discharge per unit area and other meteorological variables such as temperature, humidity, radiation and precipitation. We have observed that this model outperforms considerably the time-varying bivariate copula model proposed in a previous study that only considered temperature and glacier discharge (Gomez et al. 2017), where it was only considered the temperature and specific glacier discharge per unit area. Another important difference with our previous work is that here we have considered a zero-inflated model, following Erhardt and Czado (2012) and Brechmann et al. (2014), to describe the observed zero values for the discharge and precipitation. The present proposed method also allows for the estimation of the conditional probability of zero discharge given observed meteorological variables. We have found that these estimations are better than those obtained with a simple logistic regression using the same covariates. Finally, our approach also provides an estimation for the predictive discharge distribution conditional on the observed meteorological variables. Three different point estimations for the discharge have been used, based on the conditional predictive distribution, which have shown a good prediction performance for out-of-sample observations.

However, the proposed model presents some limitations that could be addressed in future extensions. One important issue is the temporal autocorrelation of the data. In addition to the seasonal effect, which we have already addressed, there exists a clear temporal autocorrelation of environmental daily observations such as humidity, temperature, radiation or precipitation. One possible solution would be to consider an autoregressive structure in each marginal variable and model the dependence of the residuals through vine copulas. This approach is considered in e.g. Cong and Brady (2012) where the relationship between temperature and precipitation is studied using the residuals obtained after fitting an AR(1) model in each marginal series and also a temporal trend. More generally, a possible better approach would be to follow the procedure by Pereira and Veiga (2017)) where vine copulas are used to model the temporal dependence in the univariate marginal series such that lags that are greater than one can be also incorporated. Note that using this approach we could also capture non-linear dependencies with past observations. Finally, we could also consider other dependence structures, like spatial dependence, using data from other pilot experimental watersheds installed by GLACKMA at different latitudes. This could be addressed using spatial vine copulas as described in Musafer and Thomson (2017).

Regarding the proposed vine copula model, various extensions are also possible. It would be interesting to compare the obtained results with those using other regular vine copulas, such as, for example, d-vine copulas. Also, we could allow for a different order in each c-vine for each group and period. Thus, we could have used for example a sequential procedure such as the so-called Dißmann algorithm, see Dißmann et al. (2013), where an automated selection method is developed based on empirical Kendalls tau values. Finally, the imposed "simplifying assumption" (Haff et al. 2010) could be relaxed to study if there are important changes in the conditional probability of zero discharge and predictive discharge distribution. The R GamCopula package could help us with these calculations (Vatter and Chavez-Demoulin 2015, Vatter and Nagler 2016).

Finally, there is also a notable issue regarding the sample size. Note that since we have divided the data set into various subsamples according to the seasonal period and also according to the presence or not of precipitation and discharge, there are some group-period combinations where the number of observations is very small, which leads to large standard errors in the parameter estimation in those subgroups. Therefore, we are currently working on a more general hierarchical Bayesian model where it is possible to introduce dependencies among the different groups including additional layers in the model structure. Using this approach, it is possible to model for example the relationship between the parameters that define the dependence between temperature and humidity in different groups and periods.

7 Conclusion

In this paper, we have proposed a vine copula model for modelling the relationship between the glacier discharge and other meteorological variables, such as, temperature, humidity, solar radiation and precipitation. The probability of zero discharge for each day is estimated given the observed values of the meteorological variables. Also, the predictive value of the discharge is obtained from its conditional distribution given the observations of the meteorological variables. The proposed approach has been applied to the data collected by GLACKMA from the glacier Collins between 2002 and 2012. The database has been divided into four periods according to the different hydrological seasons and the parameters have been adjusted to obtain the joint distribution of the five variables in each one of these periods. The monitor station in King George island have been registering data which have not been already collected by the GLACKMA association. Our intention is to validate our proposed model with these new data whenever they are available.

Acknowledgements We would like to thank the anonymous reviewers for their careful reading and very helpful comments. We are very grateful to the GLACKMA association. The first and second authors acknowledge financial support by MEC project MINECO2015-66593-P from the Spanish Government. The third author would like to thank the Russian, Argentinean, German, Uruguayan and Chilean Antarctic Programs for their continuous logistic support over the years. The crews of Bellingshausen, Artigas, and Carlini station as well as the Dallmann Laboratory provided a warm and pleasant environment during fieldwork. GLACKMA's contribution was also partially financed by the European Science Foundation, ESF project IMCOAST (EUI2009-04068) and the Ministerio de Educación y Ciencia (CGL2007-65522-C02-01/ANT).

A Appendix A: Density functions in terms of vine copulas

The joint density functions in (8a), (8b) and (8d) can be expressed respectively in terms of a vine copulas as,

$$f^{00}_{thr}(t,h,r) = f^{00}_t(t)\cdot f^{00}_h(h)\cdot f^{00}_r(r)\cdot c^{00}_{th}(u^{00}_t,u^{00}_h) \cdot c^{00}_{tr}(u^{00}_t,u^{00}_r)\cdot c^{00}_{hr|t}(u^{00}_{h|t},u^{00}_{r|t}), \quad (25)$$

$$f^{01}_{thrp}(t,h,r,p) = f^{01}_t(t)\cdot f^{01}_h(h)\cdot f^{01}_r(r)\cdot f^{01}_p(p)\cdot c^{01}_{th}(u^{01}_t,u^{01}_h) \cdot c^{01}_{tr}(u^{01}_t,u^{01}_r)\cdot c^{01}_{tp}(u^{01}_t,u^{01}_p) \cdot c^{01}_{hr|t}(u^{01}_{h|t},u^{01}_{r|t})\cdot c^{01}_{hp|t}(u^{01}_{h|t},u^{01}_{p|t})\cdot c^{01}_{rp|th}(u^{01}_{r|th},u^{01}_{p|th}), \quad (26)$$

and,

$$
\begin{aligned}
f^{10}_{thrd}(t,h,r,d) = {} & f^{10}_t(t)\cdot f^{10}_h(h)\cdot f^{10}_r(r)\cdot f^{10}_d(d)\cdot c_{th}(u^{10}_t,u^{10}_h)\\
& \cdot c^{10}_{tr}(u^{10}_t,u^{10}_r)\cdot c^{10}_{td}(u^{10}_t,u^{10}_d)\\
& \cdot c^{10}_{hr|t}(u^{10}_{h|t},u^{10}_{r|t})\cdot c^{10}_{hd|t}(u^{10}_{h|t},u^{10}_{d|t})\cdot c^{10}_{rd|th}(u^{10}_{r|th},u^{10}_{d|th}),
\end{aligned}
\tag{27}
$$

where

$$
\begin{aligned}
u^{jk}_x &= F^{jk}_x(x), \quad \text{for } x=t,h,r,p,d,\\
u^{jk}_{x|t} &= F^{jk}_{tx}(x\mid t) = \frac{\partial C^{jk}_{tx}(u^{jk}_t,u^{jk}_x)}{\partial u^{jk}_t}, \quad \text{for } x=h,r,p,d,\\
u^{jk}_{x|th} &= F^{jk}_{thx}(x\mid t,h) = \frac{\partial C^{jk}_{hx|t}(u^{jk}_{h|t},u^{jk}_{x|t})}{\partial u^{jk}_{h|t}}, \quad \text{for } x=r,p,d,\\
u^{jk}_{x|thr} &= F^{jk}_{thrx}(x\mid t,h,r) = \frac{\partial C^{jk}_{rx|th}(u^{jk}_{r|th},u^{jk}_{x|th})}{\partial u^{jk}_{r|th}}, \quad \text{for } x=p,d.
\end{aligned}
$$

for $j=0,1$, and $k=0,1$.

Figure 7 shows the different structures of the c-vine copulas used in this paper.

B Appendix B: Discharge prediction algorithms

In this appendix, we explain the algorithm to obtain the predictive values of the discharge with the conditional probability given in (19) and (20). Algorithm 1 details the estimation procedure to obtain the predictive mean of the specific glacier discharge per unit area given the temperature, humidity, radiation and precipitation.

Algorithm 1 Predictive discharge (using the mean)

Require: $t, h, r, p, \boldsymbol{\theta}, F^{1k}_t, F^{1k}_h, F^{1k}_r, F^{1k}_p, F^{1k}_d, (k=0,1)$

procedure
- **if** p=0 **then**
  - Compute $u^{10}_t = F^{10}_t(t)$, $u^{10}_h = F^{10}_h(h)$ and $u^{10}_r = F^{10}_r(r)$
  - Compute $u^{10}_{h|t} = C^{10}_{th}(u^{10}_h \mid u^{10}_t;\theta^{10}_{th})$ and $u^{10}_{r|t} = C^{10}_{tr}(u^{10}_r \mid u^{10}_t;\theta^{10}_{tr})$
  - Compute $u^{10}_{r|th} = C^{10}_{hr|t}(u^{10}_{r|t} \mid u^{10}_{h|t};\theta^{10}_{thr})$
  - Simulate $u^{10}_{d|thr} \sim U(0,1)$
  - Obtain the value $u^{10}_{d|th}$ that verify $u^{10}_{d|thr} = C^{10}_{rd|th}(u^{10}_{d|th} \mid u^{10}_{r|th};\theta^{10}_{thrd})$
  - Obtain the value $u^{10}_{d|t}$ that verify $u^{10}_{d|th} = C^{10}_{hd|t}(u^{10}_{d|t} \mid u^{10}_{h|t};\theta^{10}_{thd})$
  - Obtain the value u^{10}_d that verify $u^{10}_{d|t} = C^{10}_{td}(u^{10}_d \mid u^{10}_t;\theta^{10}_{td})$
  - Obtain $\widehat{d} = (F^{10}_d)^{-1}(u^{10}_d)$
- **else**
  - Compute $u^{11}_t = F^{11}_t(t)$, $u^{11}_h = F^{11}_h(h)$, $u^{11}_r = F^{11}_r(r)$ and $u^{11}_p = F^{11}_p(p)$
  - Compute $u^{11}_{h|t} = C_{th}(u^{11}_h \mid u^{11}_t;\theta^{11}_{th})$, $u^{11}_{r|t} = C_{tr}(u^{11}_r \mid u^{11}_t;\theta^{11}_{tr})$ and $u^{11}_{p|t} = C^{11}_{tp}(u^{11}_p \mid u^{11}_t;\theta^{11}_{tp})$
  - Compute $u^{11}_{r|th} = C^{11}_{hr|t}(u^{11}_{r|t} \mid u^{11}_{h|t};\theta^{11}_{thr})$ and $u^{11}_{p|th} = C^{11}_{hp|t}(u^{11}_{p|t} \mid u^{11}_{h|t};\theta^{11}_{thp})$
  - Compute $u^{11}_{p|thr} = C^{11}_{rp|th}(u^{11}_{p|th} \mid u^{11}_{r|th};\theta^{11}_{thp})$
  - Simulate $u^{11}_{d|thrp} \sim U(0,1)$
  - Obtain the value $u^{11}_{d|thr}$ that verify $u^{11}_{d|thrp} = C^{11}_{pd|thr}(u^{11}_{d|thr} \mid u^{11}_{p|thr};\theta^{11}_{thrpd})$
  - Obtain the value $u^{11}_{d|th}$ that verify $u^{11}_{d|thr} = C^{11}_{rd|th}(u^{11}_{d|th} \mid u^{11}_{r|th};\theta^{11}_{thrd})$
  - Obtain the value $u^{11}_{d|t}$ that verify $u^{11}_{d|th} = C^{11}_{hd|t}(u^{11}_{d|t} \mid u^{11}_{h|t};\theta^{11}_{thd})$
  - Obtain the value u^{11}_d that verify $u^{11}_{d|t} = C^{11}_{td}(u^{11}_d \mid u^{11}_t;\theta^{11}_{td})$
  - Obtain $\widehat{d} = (F^{11}_d)^{-1}(u^{11}_d)$
- **end if**

end procedure

For the case that we want to estimate the predictive median of the discharge, we may replace in Algorithm 1 instructions (6) and (16) by "Compute $u_{thrpd} = 1 - \frac{0.5}{\Pr(D=0|t,h,r,p)}$" and "Compute $u_{thrd} = 1 - \frac{0.5}{\Pr(D=0|t,h,r)}$" respectively.

Finally, for the third prediction method, the conditional probability, $\Pr(D = 0 \mid t, h, r, p)$ is firstly estimated and then, it is predicted that $\widehat{d} = 0$ if the estimated probability of zero discharge is greater than 0.5 or obtained with Algorithm 1 if it is smaller.

References

Aas K, Czado C, Frigessi A, Bakken H (2009) Pair-copula construction of multiple dependence. Insur Math Econ 44:182–198

Acar EF, Genest C, NešLehová J (2012) Beyond simplified pair-copula constructions. J Multivar Anal 110:74–90

Barrand NE, Vaughan DG, Steiner N, Tedesco M, Kuipers Munneke P, Broeke MR, Hosking JS (2013) Trends in Antarctic Peninsula surface melting conditions from observations and regional climate modeling. J Geophys Res Earth Surf 118(1):315–330

Bedford T, Cooke RM (2001) Probability density decomposition for conditionally dependent random variables modeled by vines. Ann Math Artif Intell 32:245–268

Braun M (2001) Ablation on the ice cap of King George Island (Antarctica). (Doctoral thesis, University of Freiburg, Germany)

Brechmann E, Czado C, Paterlini S (2014) Flexible dependence modeling of operational risk losses and its impact on total capital requirements. J Bank Financ 40:271–285

Brier GW (1950) Verification of forecasts expressed in terms of probability. Mon Weather Rev 78(1):1–3

Carnicero JA, Ausín MC, Wiper MP (2013) Non-parametric copulas for circular-linear and circular-circular data: an application to wind directions. Stoch Environ Res Risk Assess 27(8):1991–2002

Cogley JG, Hock R, Rasmussen LA, Arendt AA, Bauder A, Braithwaite RJ, Jansson P, Kaser G, Mller M, Nicholson L, Zemp M (2011) Glossary of glacier mass balance and related terms, IHP-VII Technical Documents in Hydrology No. 86, IACS Contribution No. 2, UNESCO-IHP, Paris

Cong R, Brady M (2012). The interdependence between rainfall and temperature, Copula analyses. Sci World J 2012:405675. https://doi.org/10.1100/2012/405675

De Michele C, Salvadori G (2003) A generalized Pareto intensity duration model of storm rainfall exploiting 2-copulas. J Geophys Res Atmos 108(D2):1–11

Dißmann J, Brechmann EC, Czado C, Kurowicka D (2013) Selecting and estimating regular vine copulae and application to financial returns. Comput Stat Data Anal 59:52–69

Domínguez MC, Eraso A (2007) Substantial changes happened during the last years in the icecap of King George, Insular Antactica. In: Tyk A, Stefaniak K (eds) Karst and Cryokarst, Studies of the Faculty of Earth Sciences, vol 45. University of Silesia, Katowice, pp 87–110

Erhardt V, Czado C (2012) Modeling dependent yearly claim totals including zero claims in private health insurance. Scand Actuar J 2012(2):106–129

Favre AC, El Adlouni S, Perreault L, Thimonge N, Bobe B (2004) Multivariate hydrological frequency analysis using copulas. Water Resour Res. https://doi.org/10.1029/2003WR002456

Genest C, Rivest LP (1993) Statistical inference procedures for bivariate archimedean copulas. J Am Stat Assoc 88:1034–1043

Genest C, Favre AC (2007) Everything you always wanted to know about copula but were afraid to ask. J Hydrol Eng 4:347–368

Gomez M, Ausin MC, Dominguez MC (2017) Seasonal copula models for the analysis of glacier discharge at King George Island, Antarctica. Stoch Environ Res Risk Assess 31:1107–1121. https://doi.org/10.1007/s00477-016-1217-7

Gyasi-Agyei Y, Melching CS (2012) Modelling the dependence and internal structure of storm events for continuous rainfall simulation. J Hydrol 464:249–261

Gyasi-Agyei Y (2013) Evaluation of the effects of temperature changes on fine timescale rainfall. Water Resour Res 49(7):4379–4398. https://doi.org/10.1002/wrcr.20369

Haff IH, Aas K, Frigessi A (2010) On the simplified pair-copula construction—simply useful or too simplistic? J Multivar Anal 101(5):1296–1310

Hamlet AF, Lettenmaier DP (1999) Effects of climate change on hydrology and water resources in the Columbia River Basin. J Am Water Resourc Assoc 35(6):1597–1623

Hock R (1999) A distributed temperature-index ice-and snowmelt model including potential direct solar radiation. J Glaciol 45(149):101–111

Hock R (2003) Temperature index melt modelling in mountain areas. J Hydrol 282(1):104–115

Joe H (1996) Families of m-variate distributions with given margins and $m(m-1)/2$ bivariate dependence parameters. In: Ruschendorf L, Schweizer B, Taylor MD (eds) Distributions with fixed marginals and related topics. Institute of Mathematical Statistics, Haywood, pp 120–141

Joe H (1997) Multivariate models and multivariate dependence concepts. CRC Press, Boca Raton

Killiches M, Kraus D, Czado C (2017) Examination and visualisation of the simplifying assumption for vine copulas in three dimensions. Aust N Z J Stat 59(1):95–117

Marsh P (1999) Snowcover formation and melt: recent advances and future prospects. Hydrol Process 13:2117–2134

Musafer GN, Thomson MH (2017) Non-linear optimal multivariate spatial design using spatial vine copulas. Stoch Environ Res Risk Assess 31:551–570

Nelsen RB (2006) An introduction to Copulas, 2nd edn. Springer, New York

Pham MT, Vernieuwe H, De Baets B, Willems P, Verhoest NEC (2016) Stochastic simulation of precipitation-consistent daily reference evapotranspiration using vine copulas. Stoch Environ Res Risk Assess 30(8):2197–2214

Pellicciotti F, Brock B, Strasser U, Burlando P, Funk M, Corripio J (2005) An enhanced temperature-index glacier melt model including the shortwave radiation balance: development and testing for Haut Glacier d'Arolla, Switzerland. J Glaciol 51(175):573–587

Pereira G, Veiga A (2017) PAR (p)-vine copula based model for stochastic streamflow scenario generation. Stoch Environ Res Risk Assess 32(3):833–842

R Core Team (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/

Rückamp M, Braun M, Suckro S, Blindow N (2011) Observed glacial changes on the King George Island ice cap, Antarctica, in the last decade. Glob Planet Change 79:99–109

Sarhadi A, Burn DH, Ausín MC, Wiper MP (2016) Time varying nonstationary multivariate risk analysis using a dynamic

Bayesian copula. Water Resour Res 52:2327–2349. https://doi.org/10.1002/2015WR018525

Schär C, Vidale PL, Lüthi D, Frei C, Häberli C, Liniger MA, Appenzeller C (2004) The role of increasing temperature variability in European summer heatwaves. Nature 427(6972):332–336

Scholzel C, Friederichs P (2008) Multivariate non-normally distributed random variables in climate research - introduction to the copula approach. Nonlinear Process Geophys 15:761–772

Schepsmeier U (2016) A goodness-of-fit test for regular vine copula models. Econom Rev 35:1–22. https://doi.org/10.1080/07474938.2016.1222231

Schepsmeier U, Stoeber J, Brechmann EC, Graeler B, Nagler T, Erhardt T (2017). VineCopula: Statistical Inference of Vine Copulas. R package version 2.1.2. https://CRAN.R-project.org/package=VineCopula

Sicart JE, Hock R, Six D (2008) Glacier melt, air temperature, and energy balance in different climates: The Bolivian Tropics, the French Alps, and northern Sweden. J Geophys Res Atmos 113(D24). https://doi.org/10.1029/2008JD010406

Sklar A (1959) Fonctions de répartition à n dimensions et leurs marges. Publications de l'Institut de Statistique de l'Université de Paris 8:229–231

Spanhel F, Kurz MS (2015). Simplified vine copula models: Approximations based on the simplifying assumption. arXiv preprint arXiv:1510.06971

Turner J, Colwell SR, Marshall GJ, Lachlan-Cope TA, Carleton AM, Jones PD, Lagun V, Reid PA, Iagovkina S (2005) Antarctic climate change during the last 50 years. Int J Climatol 25(3):279–294

Vatter T, Nagler T (2016) Generalized additive models for pair-copula constructions. arXiv:1608.01593v2

Vatter T, Chavez-Demoulin V (2015) Generalized additive models for conditional dependence structures. J Multivar Anal 141:147–167. https://doi.org/10.1016/j.jmva.2015.07.003

Vaughan DG, Marshall GJ, Connolley WM, Parkinson C, Mulvaney R, Hodgson DA, Turner J (2003) Recent rapid regional climate warming on the Antarctic Peninsula. Clim Change 60(3):243–274

Vuong QH (1989) Likelihood ratio tests for model selection and non-nested hypotheses. Econom J Econom Soc 57:307–333

Xiong L, Jiang C, Xu CY, Yu KX, Guo S (2015) A framework of change-point detection for multivariate hydrological series. Water Resour Res 51:8198–8217. https://doi.org/10.1002/2015WR017677

Yao AY (1974) A statistical model for the surface relative humidity. J Appl Meteorol 13(1):17–21

Yang W, Gardelin M, Olsson J, Bosshard T (2015) Multi-variable bias correction: application of forest fire risk in present and future climate in Sweden. Nat Hazards Earth Syst Sci 15(9):2037–2057