

A nested multivariate copula approach to hydrometeorological simulations of spring floods: the case of the Richelieu River (Québec, Canada) record flood

Christian Saad · Salaheddine El Adlouni ·
André St-Hilaire · Philippe Gachon

Published online: 14 October 2014

© The Author(s) 2014. This article is published with open access at Springerlink.com

Abstract Floods have potentially devastating consequences on populations, industries and environmental systems. They often result from a combination of effects from meteorological, physiographic and anthropogenic natures. The analysis of flood hazards under a multivariate perspective is primordial to evaluate several of the combined factors. This study analyzes spring flood-causing mechanisms in terms of the occurrence, frequency, duration and intensity of precipitation as well as temperature events and their combinations previous to and during floods using frequency analysis as well as a proposed multivariate copula

approach along with hydrometeorological indices. This research was initiated over the Richelieu River watershed (Quebec, Canada), with a particular emphasis on the 2011 spring flood, constituting one of the most damaging events over the last century for this region. Although some work has already been conducted to determine certain causes of this record flood, the use of multivariate statistical analysis of hydrologic and meteorological events has not yet been explored. This study proposes a multivariate flood risk model based on fully nested Archimedean Frank and Clayton copulas in a hydrometeorological context. Several combinations of the 2011 Richelieu River flood-causing meteorological factors are determined by estimating joint and conditional return periods with the application of the proposed model in a trivariate case. The effects of the frequency of daily frost/thaw episodes in winter, the cumulative total precipitation fallen between the months of November and March and the 90th percentile of rainfall in spring on peak flow and flood duration are quantified, as these combined factors represent relevant drivers of this 2011 Richelieu River record flood. Multiple plausible and physically founded flood-causing scenarios are also analyzed to quantify various risks of inundation.

Electronic supplementary material The online version of this article (doi:10.1007/s00477-014-0971-7) contains supplementary material, which is available to authorized users.

C. Saad (✉) · A. St-Hilaire
Centre Eau-Terre-Environnement, Institut National de la
Recherche Scientifique, 490, de la Couronne Street,
Quebec G1K 9A9, Canada
e-mail: christian.saad@ete.inrs.ca

A. St-Hilaire
e-mail: andre.st-hilaire@ete.inrs.ca

C. Saad · P. Gachon
Centre ESCER, Université du Québec à Montréal, 201 President-
Kennedy Avenue, Montreal, QC H2X 3Y7, Canada
e-mail: philippe.gachon@ec.gc.ca

S. El Adlouni
Département de Mathématiques et Statistique, Université de
Moncton, 18 Antoine-Maillet Avenue, Moncton, NB E1A 3E9,
Canada
e-mail: salah-eddine.el.adlouni@umoncton.ca

P. Gachon
Canadian Centre for Climate Modelling and Analysis, Climate
Research Division, Environment Canada, 800, de la Gauchetière
West Street, Montreal, QC H5A 1L9, Canada

Keywords Multivariate copula · Spring flood risks ·
Hydrometeorology · Frequency analysis

Abbreviations

AIC	Akaike information criterion
BIC	Bayesian information criterion
FNA	Fully nested Archimedean
GEV	Generalised extreme value
T	Return period
EC	Environment Canada
NCDC	National Climatic Data Center

MAM	March, April and May or spring season
JJA	June, July and August or summer season
SON	September, October and November or fall season
DJF	December, January and February or winter season
Fldur	Flood duration
Qmax _{spr}	Maximum spring flow
CWD	Consecutive wet days
PrecTOTNM	Total precipitation fallen between November and March
PrecSolDJF	Solid precipitation fallen in winter
PrecSolM	Solid precipitation in March
PrecLiqMAM	Liquid precipitation fallen in spring
R5DLiqMAM	Maximum cumulative liquid precipitation fallen in 5 consecutive days in spring
SDIILiqMAM	Simple daily liquid precipitation in spring
P90pLiqMAM	90th percentile of liquid precipitation in spring
DDmelt	Degree days of melt
FrThSeqDJF	Number of days with frost/thaw cycles in winter

1 Introduction

In the spring of 2011, the Richelieu River (Canada) and the Lake Champlain drainage basin (straddling the U.S.-Canada border) were flooded for 67 days, with lake levels reaching record heights of 31.47 m of water above mean sea level; an event which had yet to be seen in the past century. Major environmental consequences followed damage to residential, commercial and agricultural properties throughout the watershed with more than 4,000 homes flooded and a total estimated cost of approximately 88 million dollars in damages as well as emergency and recovery operations for Canada and the United-States of America combined. The lands adjacent to the Richelieu River were most affected by this event; about 70 % of that total estimated cost was expensed to the Quebec province (GTILCR 2013). Consequently, concerned decision makers, business owners and residents of the watershed demanded to know what could have caused such an extreme flood. In response, an early study by Environment Canada put forward the hypothesis that the joint occurrence of particular meteorological events played a major role in the observed high water levels and record flows (Cantin 2011). Beniston (2009) and Hao et al. (2013) have shown that analyzing the joint occurrence of temperature and precipitation quantile exceedance allowed for a more comprehensive understanding of trends in cold/dry, cold/wet,

warm/dry and warm/wet extremes for both past and future periods around the globe, which have a major incidence on hydrological systems. The main goal of this work is to explore the joint effect of hydrometeorological conditions using a proposed multivariate statistical approach. The multi-dimensional aspect of this study is essential since most spring floods are consequences of multiple cumulative and simultaneous events/conditions, including primarily snowmelt conditions. The severity of spring floods in North-East America depends mainly on the rate of snowmelt and the snow accumulation from the previous winter, but intense spring rainfall events may also amplify the phenomenon (Lawford et al. 1995; Turcotte et al. 2010; Villarini et al. 2011; Mazouz et al. 2012). Such meteorological events can be quantified using indicators developed by the Expert Team on Climate Change Detection Monitoring and Indices (ETCCDMI) and/or by the STATistical and Regional dynamical Downscaling of EXtremes for European regions project (STARDEX), which were used in previous studies in eastern Canada in Gachon et al. (2005). Hence, the joint occurrences of such meteorological events must be taken under consideration when analyzing hydro-meteorological hazards.

In practice, the concurrent consideration of various hydrologic variables in flood prediction systems is a common approach. Lavallée et al. (2000) describe a flood forecasting system for summer and fall flash floods, which has proven efficient for a neighboring watershed of the Richelieu River. It takes into account the number of humid days that recently passed in combination with past and predicted occurrences and intensities of rainfall events in order to best assess future flood risks. Turcotte et al. (2010) describe the current operational hydrologic forecasting system used by the “*Centre d’Expertise Hydrique du Québec*” of the Quebec Ministry of Sustainable Development, Environment, Wildlife and Parks (Canada) as an assimilation of previous hydrometric, temperature, precipitation and snow accumulation observations as well as short (0–72 h) and medium (0–90 days) term temperature and precipitation forecasts used as inputs to hydrologic models. Hence, understanding the combined effects of various past and/or forecasted hydrometeorological events is essential to proper flood predictions and risk management. Although copulas are not commonly used for short term flood predictions, this statistical technique can be used for flood risk management as a risk assessment tool that takes into account the joint effect of inter-dependant hydrological and meteorological variables on floods.

The term copula was introduced to the scientific community by Sklar (1959), as an n -dimensional cumulative distribution function used to describe the dependence between random variables. This statistical technique connects multivariate probability distributions to the marginal

probability distributions of each variable considered (Nelsen 2006). The latter differs from more common statistical approaches used to study or model extreme values in hydrology, which consists of associating a unique probability distribution to all interdependent variables while assuming that those variables are independent from each other (Zhang et al. 2006). In recent years, copulas have generated predominant interest, development and applications in the finance sector (Genest et al. 2009). In fact, the book by Nelsen (2006; originally published in 1999) was particularly important in the vast spreading and understanding of copula theory. According to Genest et al. (2009), an interest in extreme-value theorem and copula applications in hydrology grew following several worldwide conferences on the subject. Several types of copulas exist and have been used in environmental studies, but bivariate Archimedean copula has grown popular in hydrological research due to its associative, flexible and symmetric properties. The latter property does not hold true for all Archimedean copulas, as some have asymmetric constructions (Liebscher 2009). Moreover, bivariate Archimedean copulas can represent a large variety of dependence structures between variables both in the positive and negative range via the Gumbel (1958), Clayton (1978) and Frank (1979) copula. The common approach uses a single parameter to model the link between pairs of variables. The copula parameter can directly be derived from the Kendall's correlation coefficient (i.e. Kendall's Tau) associated to the variables considered (Favre et al. 2004). There are several ways to adapt bivariate Archimedean copulas to account for 3 or more variables. One way is to adopt a classical multivariate Archimedean copula using a single parameter to represent the dependence between all the variables (Fischer et al. 2007). Another uses a nested approach, which consists of joining two or more bivariate Archimedean copulas using parent Archimedean copulas (Hofert 2008). The advantages of such an approach are numerous and include a good flexibility in the representation of dependence structures of hydrometeorological variables. In other words, not only can copulas represent a large range of positive and negative correlations, but the nested approach takes into account various degrees of dependencies for different pairs of variables (Huard et al. 2004). Copulas have been used to calculate the conditional and joint probability of occurrence of particular hydrological events as well as to estimate the return period of such events (Salvadori and De Michele 2007). El Adlouni and Ouarda (2008) demonstrated the added value gained by using copulas to evaluate the combined effect of river flow and downstream lake levels when estimating hydrologic hazards. Copula modeling was also used to assess hydrologic characteristics of ungauged watersheds via the relationship between hydrometeorological indices such as cumulative annual evapotranspiration

and maximum annual water deficit determined from nearby gauged watersheds (Nazemi and Elshorbagy 2012). In addition, Zhang and Singh (2006) used entropy theory in combination with copula theory to estimate the joint and conditional return periods of rainfall and runoff events of various watersheds over North America. Jeong et al. (2013) used bivariate Archimedean copulas in order to estimate the projected changes of 3 pairs of flood characteristics for future stream flow scenarios (2041–2070) driven by five members of the Canadian Regional Climate Models using ensemble simulations.

This study explores hydrometeorological risks over the Richelieu River and Lake Champlain (LCR) watershed using a proposed multivariate copula approach. The main objective consists of gaining a more comprehensive understanding of flood risks for the region of interest. Such knowledge can help improve the detection of future flood events and ultimately serve in the development of flood warning and risk management systems as well as engineering designs. In that perspective, a nested multivariate copula model is developed and evaluated on the region of interest. The joint effect of particular meteorological events on the recurrence of extreme floods such as the 2011 spring incident of the Richelieu River are investigated using the proposed approach. Joint and conditional return periods of extreme peak flow, flood duration and meteorological events of the study region are estimated. A characterization of the events preceding and during severe floods, as well as under less extreme conditions is conducted to complement this information. Furthermore, statistically plausible and physically founded scenarios of notable flood events are elaborated.

The methodology is first introduced by describing the hydrometeorological indices that are used to characterize spring flood hazards (Sect. 2.1.1). The copula theory (Sect. 2.2.1) and the proposed simulation algorithm (Sect. 2.2.2) will also be covered in the methodology. This is followed by a description of the study region and data (Sect. 3). A brief hydrometeorological characterization of the study region is introduced for context purposes (Sect. 4.1). The model evaluation (Sect. 4.2.1), its application to a record flood (Sect. 4.2.2) and the elaborated hydrometeorological scenarios (Sect. 4.2.3) are then presented. The discussion and conclusion (Sect. 5) investigate further research opportunities on the subject.

2 Methodology

2.1 Analysis of hydrometeorological events

2.1.1 Hydrometeorological indicators

Hydrometeorological characteristics that best describe spring flood hazards are explored by analyzing the total

daily precipitation and temperature observations in conjunction with the flow regime of floods. This is complemented with the use of indices calculated for different time-scales (i.e.: monthly, seasonal, annual) and for various locations over the region of interest in order to spatio-temporally describe hydrometeorological events in terms of occurrence, intensity, duration, quantity and variability. The seasonal timescale consists of the 4 temperate seasons: Winter, composed of the months of December, January and February (DJF), Spring (MAM), Summer (JJA) and Fall (SON). In addition, extended seasons comprised of several months (from November to March), are also considered in order to account for cumulative precipitation (mainly solid, i.e. snow) fallen prior to the flood seasons during cold months (with negative daily temperatures). In fact, the amount of snow cover accumulated during the winter months constitutes the main factor for the strong majority of spring flood events in Nordic countries such as Canada.

The flood peak and duration variables are used to measure flood severity in terms of intensity and length respectively, while evaluating the performance of the proposed approach. Those variables are quantified using the maximum flow index and a peak-over-threshold flood duration index adapted specifically for the river of interest. Although the maximum flow index can be calculated for the four seasons, an emphasis is put on snowmelt floods that occur in spring; hence, peak flow analyses of the current work are conducted using the maximum spring flow ($Q_{max,spr}$, in m^3/s). The flood duration index (Fl_{dur} , in days) consists of the consecutive number of days for which the daily flow values exceed a certain onset/offset threshold. The latter is calculated by multiplying a scale factor to the climatological mean annual flow. An appropriate threshold needs to be neither too high nor too low to avoid underestimation or overestimation of flood duration, respectively. Furthermore, it must be selected in consideration of high threshold effects on the correlation between peak flow and flood duration. As the threshold value increases, the estimated flood duration decreases and the absolute value of the correlation between both indices increases in return (Karmakar and Simonovic 2009). The flood duration threshold is also selected with the objective of maximizing the number of years for which a flood is detected since high thresholds may result in omitting certain floods. Finally, the duration index is calculated between the 45th and 250th Julian day of each year in order to specifically account for spring floods that may start before the end of winter or extend into the summer.

This study also considers climate indices developed by the ETCCDMI and the STARDEX project to characterize meteorological events as observed by the weather stations of the region of interest for a common period. The quantity of precipitation fallen per station ($Prec_{TOT}$, in mm) is

represented in terms of total, liquid and solid cumulative precipitations for various timescales (i.e. monthly or seasonal). The liquid and solid precipitations are derived from the daily total precipitation and the mean temperature values. When the daily mean temperature is below $0\text{ }^{\circ}\text{C}$, the precipitation fallen on that same day is considered solid, whereas liquid precipitation is assumed when mean temperature is above that threshold value. The duration of precipitation events are determined with the number of consecutive wet days (CWD , in days; i.e. number of days with precipitation $\geq 1\text{ mm/day}$). The intensity of precipitation events are calculated using several indicators including the simple daily intensity index ($SDII$, in mm/wet day), which consists of the ratio of cumulated precipitation fallen per period (i.e. month, season, year) to the number of wet days (using the threshold of 1 mm/day to define wet/dry days). In addition, the maximum cumulative precipitation fallen during 5 consecutive days ($R5D$, in mm) is calculated to detect persistent and intense precipitation events for the period of interest. In order to determine the intensity of heavy daily precipitation events, the 90th percentile of precipitation ($P90p$, in mm/day) is calculated per period using the Cunnane formula (1978), commonly used by Canadian hydrologists for flood frequency curves (Helsel and Hirsch 2002). The number of days per period with daily frost/thaw episodes ($FrThSeq$, in days) is also calculated for different timescales in order to assess the snow accumulation potential prior to the flood season. The criteria selected to differentiate between frost and thaw events consists of reaching a daily minimum temperature below $0\text{ }^{\circ}\text{C}$ for frost events, and a maximum temperature above $0\text{ }^{\circ}\text{C}$ for thaw events on the same date. A degree day of melt index is used to estimate the potential water equivalent of snowmelt available per day (DD_{melt} , mm). The latter is described in Anctil et al. (2005) as an empirically developed index for mostly forested regions, and is calculated according to the following formula: $h = 0.05 * (T_{mean} - 32)$, where h is the potential water equivalent of snowmelt (converted in mm by a ratio of 1 inch/25.4 mm), and T_{mean} is the mean temperature of the studied period in degrees Fahrenheit. This definition of DD_{melt} is selected for the purpose of the study because of the limited data required for its calculation. The validity of this index is confined to mean temperatures between 1 and $19\text{ }^{\circ}\text{C}$, which are typically satisfied during the months of March and April, when snowmelt is predominant in North America.

2.1.2 Interannual variability analysis

The inter-annual standardised anomalies of each index are calculated for a common period in order to compare the yearly hydrometeorological events as recorded by the main

observation stations (i.e. hydrometric and meteorological). In addition, the dependence between the different indices are detected and quantified for every observation station as well as relevant timescales and periods using Kendall’s correlation test with 90 and 95 % confidence intervals in order to describe the relationship between the hydrometeorological events.

2.1.3 Univariate frequency analysis

The univariate risks of occurrence of extreme events are quantified using frequency analysis. This statistical technique estimates the probability of occurrence of a particular event according to observations of past events of similar nature. The model used for such estimates consists of a mathematical equation that determines the probability that a specific value is exceeded [i.e. $P(X > x)$, where x is the quantile and X the random variable for which we are estimating the exceedance probability]. The recurrence (or return period) of the event is calculated using the inverse value of the exceedance probability. Consequently, the probability of occurrence of a particular event is expressed in terms of the return period; a high return period for a specific event means that this event is rarely observed. The semi-automated approach developed by El Adlouni et al. (2008) is used to select and adjust an adequate probability model to each variable. The particularity of this method consists of verifying a series of statistical criteria while executing various tests, classifying the probability distribution of the variable of interest according to the weight of the upper-tail of that distribution. Since this work focuses on extreme flood events, two types of heavy-tailed distributions are considered: The generalised extreme value (GEV) and the Gamma distributions. In the case where multiple probability distributions fit the observation series, the Akaike and Bayesian information criteria (AIC, Akaike 1974; BIC, Schwarz 1978, respectively) are used to select an optimal probability model. Therefore, using a univariate frequency analysis, the return period of extreme events are estimated considering the observed time series of a single variable. However, spring flood hazards more often result from a combination of multiple meteorological events, which is why the consideration of joint meteorological conditions for the estimation of flood return periods is also conducted.

2.2 Multivariate analysis

2.2.1 Copula theory

Multivariate frequency analysis allows for greater understanding of flood risks by considering more than one variable that influences such hazards. A copula approach is

used to conduct this analysis. Hence, it is important to define the terms and notation as well as to describe the theory behind this method. A copula is a statistical technique used to connect multivariate to univariate probability distributions (Nelsen 2006). Following this description, the multivariate probability distribution is expressed in terms of the uniform marginal distributions $[F_X(x), F_Y(y)]$ of the continuous random variables $[X, Y]$ in conjunction with the dependence function (C) between those variables as shown in Eq. 1.

$$F(X, Y) = C(F_X(x), F_Y(y)) \tag{1}$$

Several types of copula functions exist, but Archimedean copula families are more commonly used in hydrology due to their properties, which allow for greater flexibility and simplicity of use. In fact, a single parameter per copula is adjusted according to the marginal probabilities and Kendall’s tau in order to best represent the positive or negative dependence between the variables. Some common Archimedean copula families include the Frank and Clayton copulas, which have the following general form in a bivariate case:

$$C(F_X(x), F_Y(y)) = \varphi^{-1}(\varphi(F_X(x)) + \varphi(F_Y(y))) \tag{2}$$

where φ , the “generating function”, is a convex [$\varphi''(u) > 0$], monotonically decreasing [$\varphi'(u) < 0$] function with $\varphi(1) = 0$, which differs according to the copula family (Table 1), and φ^{-1} the pseudo-inverse of that function. where $D_k(x) = \frac{k}{x^k} \int_0^x \frac{t^k}{e^t - 1} dt$ is the «Debye» function of first order ($k = 1$).

These bivariate copulas are used to calculate the probability of occurrence of joint and conditional events (Eqs. 3 and 4) as well as their return periods (Eqs. 5 and 6). Interpreting conditional return periods requires an understanding of the relationship between the univariate and multivariate estimations of recurrences and associated hazards. The results of univariate frequency analysis will indicate the estimated recurrence of the studied phenomenon (T), which can also be expressed in terms of probability of occurrence (i.e. $1/T \times 100$ %). The bivariate conditional return period indicates the probability that the studied phenomenon occurs if and only if a prior condition takes place. In other terms, the conditional return period depends on the recurrences of both the studied phenomenon and the covariate. An example of this interpretation is given in Sect. 4.2.2. These conditional probabilities can be used to assess particular flood occurrences using observed or predicted meteorological conditions. They can also be used to elaborate various statistically founded flood-causing scenarios, while determining which meteorological conditions are most influential to the flood characteristics of interest. This is conducted by comparing the conditional

return period of flood characteristics under meteorological conditions of the same order of magnitude. Those orders of magnitude consist of common return periods such as 2, 20, 50 and 100 years estimated from their respective marginal distributions. Therefore, under similar marginal conditions, a lower conditional return period reflects a more influential meteorological phenomenon for the flood characteristic in question. Similarly, in a multivariate case, when interpreting the effect of various orders of magnitude of a covariate on a flood characteristic, a lower recurrence signifies a more influential flood-causing condition.

$$P(X > x, Y > y) = [1 - P(X \leq x)] + [1 - P(Y \leq y)] - 1 + C(F_x(x), F_y(y)) \tag{3}$$

$$P(X > x|Y > y) = \frac{P(X > x, Y > y)}{P(Y > y)} \tag{4}$$

$$T_{X > x, Y > y} = \frac{1}{P(X > x, Y > y)} \tag{5}$$

$$T_{X > x|Y > y} = \frac{1}{P(X > x|Y > y)} \tag{6}$$

The performance of such statistical models can be improved by considering additional information on flood characteristics from more than one significant meteorological condition. In the case where three or more interdependent variables are considered in a multivariate frequency analysis, a fully nested Archimedean copula (FNA) can be adopted. This consists of separating the n variables in $n-1$ mutually exclusive sub-groups with a single copula parameter adjusted per sub-group (Fig. 1; Fischer et al. 2007). In other words, a first copula is adjusted for the first pair of variables and a second copula is adjusted according to that initial copula and the third variable as shown in Eq. 7.

$$C(F_X(x), F_Y(y), F_Z(z)) = \varphi_2^{-1}[\varphi_2(\varphi_1^{-1}[\varphi_1(F_X(x)) + \varphi_1(F_Y(y))]) + \varphi_2(F_Z(z))] \tag{7}$$

This approach allows for some flexibility in the representation of dependence structures between the different groups of variables. However, with this configuration, a similar dependence between the first and third variables will be associated to the second and third variables. Consequently, a less accurate estimation of the probability of occurrence may be expected with an increasing number of variables. Furthermore, several constraints arise in order to respect the essential monotonicity condition described by Eq. 8.

$$\varphi_2(\varphi_1^{-1}[\varphi_1(F_X(x)) + \varphi_1(F_Y(y))]) \tag{8}$$

Huard et al. (2004) have derived those multivariate constraints for various Archimedean copulas as shown by the trivariate constraint of the Clayton copula parameter in

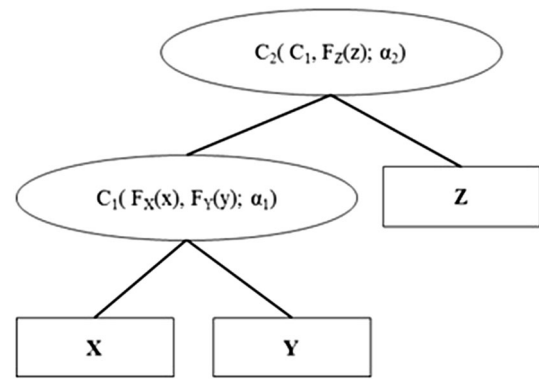


Fig. 1 Schematic representation of a fully nested trivariate Archimedean copula

Table 1 and that of the Frank copula parameter illustrated in Fig. 2.

2.2.2 Proposed simulation algorithm

The multivariate FNA approach is applied to the Clayton and Frank copulas under trivariate conditions in order to determine the probability of occurrence of joint and conditional hydrometeorological extremes associated with various events. Two simulation algorithms are proposed in order to estimate these probabilities of occurrence since no explicit formulations are available. The steps undertaken to estimate these probabilities from interdependent hydro-meteorological simulations are described in this section and a corresponding MATLAB function is provided in appendix. The parameters of the Clayton and Frank copulas are estimated for (1) a first pair of variables and (2) those copulas are calculated. The most appropriate copula used to model the dependence between those variables is selected using the k-plot and the quadratic distance criteria (3; see below). In the trivariate case, (4) steps 1 and 2 are repeated using the copula selected in step 3 in combination with the third variable of interest; however, the parameter of this parent copula is fitted to the new combination of variables. An adaptation of the Marshall and Olkin (1988) bivariate simulation algorithm is then applied (5) to generate quantiles of the 3 variables. (6) The joint and conditional probability of occurrence and the return periods of particular events are estimated.

The estimation of the parameters of the bivariate Archimedean copulas is conducted with the method of moments, which was found adequate for estimating parameters of distribution functions for small sample sizes. Furthermore, separately calculating marginal and joint distribution parameters provides more accurate estimates than methods used to approximate all parameters simultaneously. The method of moments consists of estimating the

Table 1 Description of the generating functions, parameter constraints as well as of the relationships between copula parameter and Kendall’s Tau for Archimedean copulas (contents extracted from Huard et al. 2004; El Adlouni et al. 2008)

Copula family	$\varphi(k)$, with $k = F_X(x)$ or $F_Y(y)$	Bivariate copula $C(F_X(x), F_Y(y))$	Copula parameter versus Kendall’s tau	Bivariate constraint on α	Trivariate constraint on α
Clayton	$k^{-\alpha} - 1$	$(F_X(x)^{-\alpha} + F_Y(y)^{-\alpha})^{-1/\alpha}$	$\tau = \frac{\alpha}{\alpha+2}$	$\alpha > 0$	$\alpha_2 \leq \alpha_1$
Frank	$\ln\left(\frac{\exp(\alpha k) - 1}{\exp(\alpha) - 1}\right)$	$\frac{1}{\alpha} \ln\left\{1 + \frac{[\exp(\alpha F_X(x)) - 1][\exp(\alpha F_Y(y)) - 1]}{\exp(\alpha) - 1}\right\}$	$\tau = 1 - \frac{4}{\alpha} \{1 - D_1(\alpha)\}$	$\alpha \geq 0$	Fig. 2

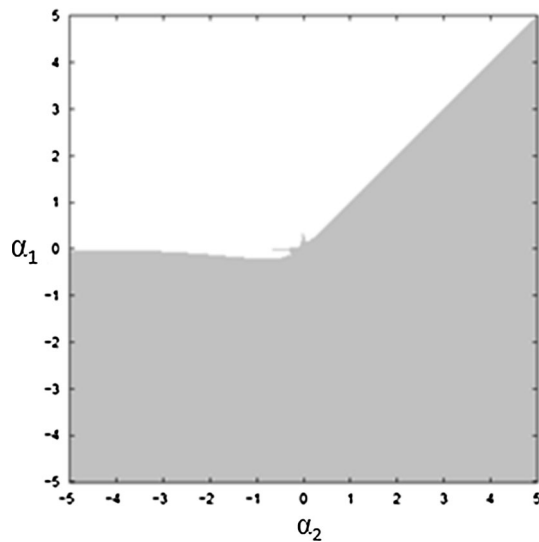


Fig. 2 Graphical representations of the trivariate fully nested Frank copula parameter constraint, where values within white spaces satisfy the constraints (Huard et al. 2004)

Clayton and Frank copula parameter according to its relationship with Kendall’s Tau as shown in Table 1 (El Adlouni and Ouarda 2008).

The most appropriate copula is selected according to the k -plot and the minimum quadratic distance (L^2) between the empirical (\hat{K}) and theoretical ($K\psi$) values of the K criterion calculated for the various copulas. The theoretical values are determined using the generator function respective to each copula and a vector of uniform values z for which the criterion is evaluated using Eq. 9.

$$K_\psi = z - \frac{\varphi(z)}{\varphi'(z)} \tag{9}$$

On the other hand, the empirical values of K are calculated using the sample size (S) as well as both variables of interest (x, y) as shown in Eq. 10.

$$\hat{K}(z) = \frac{1}{S} \sum_{s=1}^T 1_{[w_i \leq z]},$$

where $w_i = \frac{1}{S-1} \sum_{s=1}^S 1_{[x_1^s < x_1^i, y_2^s < y_2^i]}$ and $s = 1, \dots, S$

$$\tag{10}$$

Graphical comparisons of those theoretical and empirical values are conducted using k -plots. The copula which minimizes the surface area between the theoretical and empirical curves is selected.

The Marshall and Olkin (1988) bivariate simulation algorithm for the Clayton copula consists of the following:

1. Generating 2 independent values (y_1, y_2) from the exponential distribution with mean parameter $\mu = 1$.
2. Generating an independent value (z_1) from the gamma distribution with shape and scale parameters $1/\alpha_1$ and 1, respectively, which will be used in combination with the copula parameter to link both variables.
3. Calculating the simulated variables using $u_j = (1 + y_j/z_1)^{(-1/\alpha_1)}$ with $j = 1, 2$; where random variable u_j is distributed according to the Clayton copula.

In the trivariate case, generalizing this algorithm for 3 variables consists of generating a third randomly distributed variable u_3 in order to have $[C(u_1, u_2; \alpha_1), u_3] \sim C(\alpha_2)$. A second linking value z_2 can be calculated from $C(u_1, -u_2; \alpha_1) = V_1$, which is then used to estimate u_3 . In this case, the trivariate simulation algorithm of the Clayton copula is as follows:

1. Executing the same three steps as the bivariate Clayton simulation.
2. Calculating the first copula $V_1 = C(u_1, u_2; \alpha_1)$.
3. Generating two more exponentially distributed independent values (y_3, y_4) with mean parameter $\mu = 1$.
4. Calculating the second linking value z_2 from V_1 using $z_2 = \frac{y_3}{(V_1^{-\alpha_2} - 1)}$.
5. Calculating the simulated value u_3 similarly to step 3 of the bivariate Clayton algorithm stated above, but with $j = 3$; where random variable u_3 is distributed according to the Clayton copula.

Finally, the quantiles associated to those probabilities are obtained by applying the marginal inverse cumulative probability distributions of the respective variables.

In the case of the Frank copula, the algorithm is simpler since the variables are simulated using the inverse of the Frank generating function. The bivariate algorithm goes as follow:

1. Generating two uniformly distributed random numbers (y_1, y_2) .
2. Associating the first random number y_1 to the first variable u_1 .
3. Calculating the second simulated variable by taking the inverse of the Frank generating function (φ^{-1}) between y_2 and u_1 as $u_2 = -\alpha_1^{-1} \ln\left(1 + \frac{(y_2 e^{-\alpha_1} - 1)}{(y_2 + (1 - y_2)e^{-\alpha_1 u_1})}\right)$.

In the trivariate case, the third variable u_3 is simulated in order to have $[C(u_1, u_2; \alpha_1), u_3] \sim C(\alpha_2)$ similarly to the Clayton copula. To do so, one has to generate a third uniformly distributed random number and calculate the inverse of the Frank generating function of that number with the initial copula value of the first 2 simulated variables. In this case, the trivariate nested simulation algorithm for the Frank copula is as follow:

1. Executing the same three steps as the bivariate Frank simulation.
2. Generating a third uniformly distributed random value y_3 .
3. Calculating the first copula $V_1 = C(u_1, u_2; \alpha_1)$ between the first 2 simulated variables.
4. Calculating the third simulated variable u_3 by taking the inverse of the Frank copula between y_3 and u_1 similarly to step 3 of the bivariate Frank copula algorithm stated above.

The quantiles of these probabilities are determined similarly to the approach taken for the Clayton copula. It is important to note that the inverse of the Frank generating function cannot be solved analytically. A numerical method is used to estimate the copula parameter.

An evaluation of the proposed adaptations of the fully-nested Archimedean multivariate copula simulation algorithms is conducted by comparing the distribution of the simulated values to that of the observed values for the different variables considered with the use of box-and-whisker plots. In addition, 2 and 3-dimensional graphic representations of the simulated values of various combinations of variables are examined for expected dependence structures. Moreover, the scatterplots of observed events are superimposed to these simulations in order to qualitatively evaluate the suitability of each copula as suggested by Chowdhary et al. (2011). Finally, the univariate return periods of the simulated variables are compared to the results of the frequency analysis obtained from observations in order to verify that the marginal distributions are respected.

For the purpose of this study, a large sample of 100,000 simulations of inter-dependent variables is produced using the before-mentioned algorithm. The joint and conditional probability of occurrence of particular events are estimated from samples of 50,000 multivariate simulations randomly

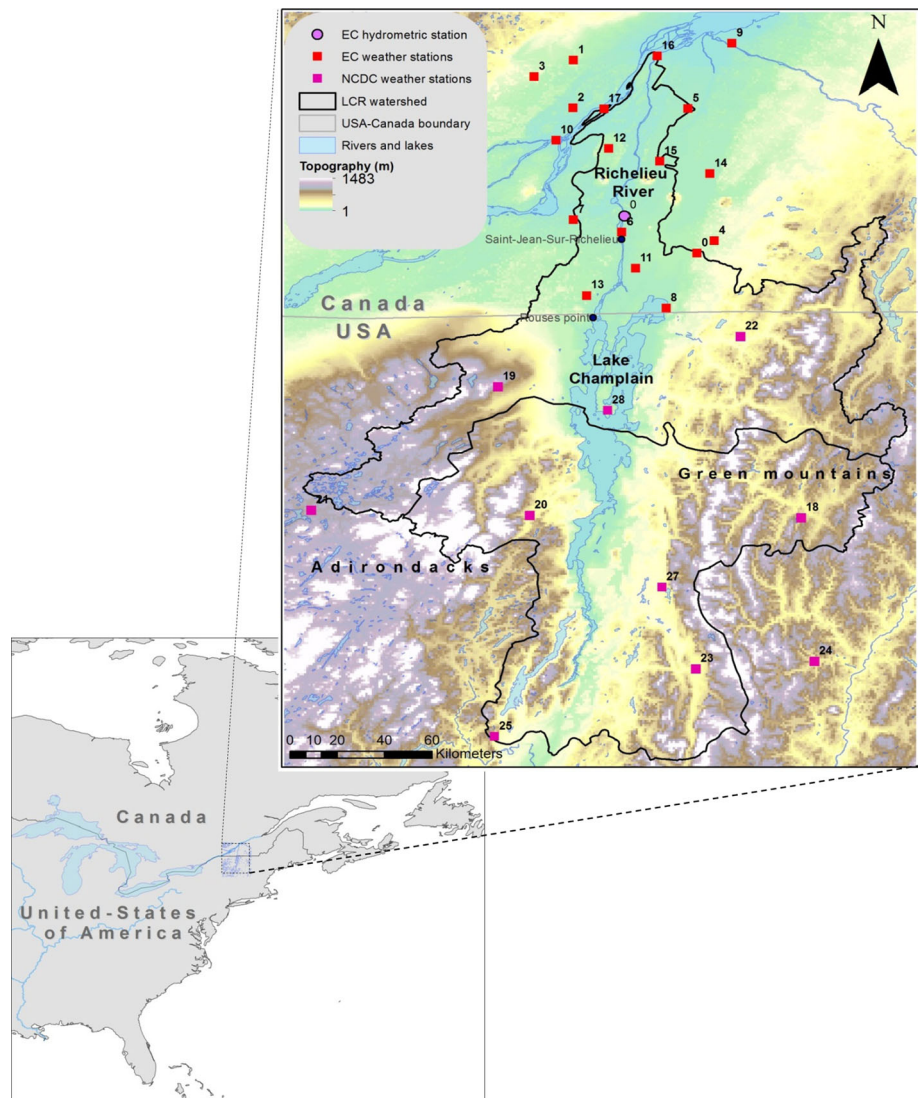
extracted from the total number of simulations generated. Various combinations of 2–3 dependent hydrometeorological indices are primarily used to characterise a record breaking flood of the river of interest in terms of intensity and duration. The same method is used to statistically elaborate plausible flood-causing hydrometeorological scenarios by selecting various combinations of meteorological thresholds for which the probability of occurrence of a particular flood event approaches unity. Those scenarios are analysed in terms of their return periods and relative to their 95 % confidence interval, which is obtained via the before-mentioned bootstrapping technique (Efron and Tibshirani 1993).

The advantages of implementing such a nested multivariate copula model are multiple. Foremost, they lay in the flexibility to represent various dependence structures between several variables even with relatively small sample sizes in addition to preserving the marginal distribution which best represent each variable. In addition, the computational resources required to run such simulations are relatively low, accessible and easy to implement in an operational context. Furthermore, the uncertainties associated to the estimated risks are illustrated via the 95 % confidence interval of the simulations.

3 Study region and data

Partially located in a moderate continental sub-humid climate characterized by cold winters, warm summers and a long season favorable to agricultural growth (Comité de Concertation et de Valorisation du Bassin de la Rivière Richelieu 2011), the study region stretches 125 km along the Richelieu River from its outlet at the St. Lawrence River in Quebec (Canada) and extends another 193 km to the southern tip of Lake Champlain. This lake is situated between the states of New York (NY) and Vermont (VT) in the United States of America (USA). This area is delineated by two watersheds with only a small extent in Canada (16 %), more specifically in the flat low lands of southern Quebec and the St. Lawrence valley. The remaining region of interest (84 %) falls between the Adirondacks (NY) and the Green mountains (VT) in the USA, which amount to a total watershed area of approximately 23,899 km². About 90 % of the Richelieu River flow originates from the mostly forested (85 % of the total study area) Lake Champlain watershed near Rouses Point (NY). The difference in height from that point to the outlet is more than 33 m, with a shallower river bed adjacent to Saint-Jean-Sur-Richelieu (Quebec) and no structures regulating the natural flow of the river. The daily flow-measuring hydrometric station used in this study is approximately 10 km downstream of Saint-Jean-Sur-Richelieu at Rapides Fryer (Fig. 3). Twenty eight (28)

Fig. 3 Maps showing the Lake Champlain and Richelieu River watershed in a continental (lower left) and regional (top right) perspective including physiographic and boundary features in addition to the network of observation stations used in this study



meteorological stations spread through Canada and USA are also considered to measure daily total precipitation, minimum and maximum air temperature for the 31-year common period between January 1st, 1981 and December 31st, 2011. From this data bank, the South Hero (NY) station, identified as station #28 in Fig. 3, is selected for univariate and multivariate analysis because of its proximity to the hydrometric station and due to shared homogeneous meteorological characteristics with most of the stations in the region upstream of Rapides Fryer (Fig. 3). Further information on these observation stations are detailed in Table 2.

4 Results

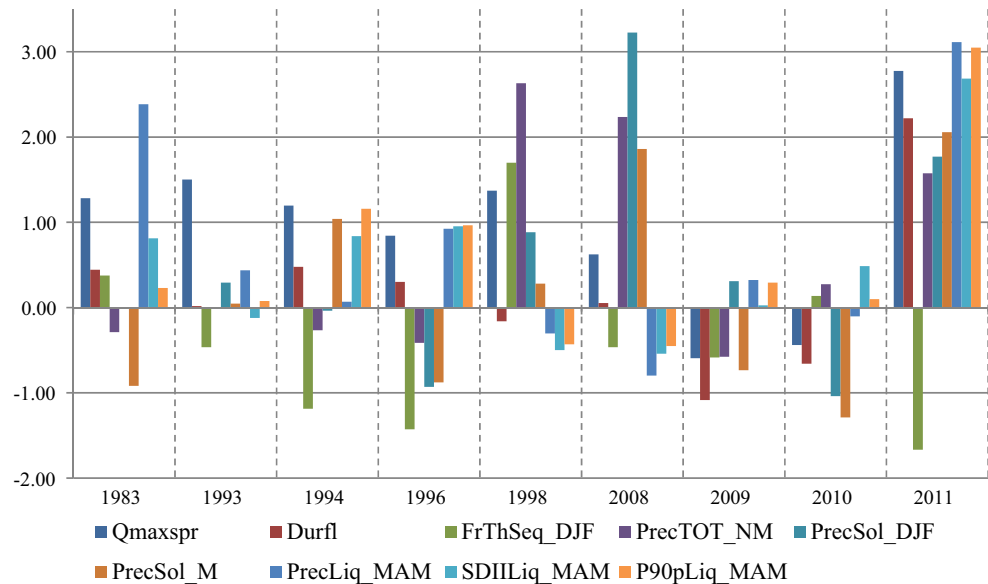
4.1 Hydrometeorological analysis

The interannual standardized anomalies of hydrometeorological indices calculated for years with the highest

maximum spring flow (i.e. 1983, 1993, 1994, 1996 and 1998) and more recently, less abnormal years (i.e. 2008, 2009 and 2010) that preceded the 2011 record spring flood are compared relative to the 1981–2011 period in Fig. 4. The maximum spring flow recorded at Rapides Fryer was abnormally high in 2011 with a flow of approximately $1,550 \text{ m}^3/\text{s}$. A positively high anomaly can also be observed for the flood duration ($Q_{\text{flood_threshold}} \geq 672.32 \text{ m}^3/\text{s}$) that lasted 111 days and extended into the summer with a maximum flow of $1,470 \text{ m}^3/\text{s}$. The univariate return periods for such maximum spring flow and flood duration events are of ~ 90 and 22 years according to the lognormal and gamma distributions, respectively. Table 3 enumerates the marginal distributions used for all indices and stations studied. Assumptions of homogeneity, stationarity and independence for the time series of all the indicators are verified with a 90 % confidence interval and in some cases with 95 % confidence. The 1993 $Q_{\text{max_spr}}$ anomaly is the second highest for the 31 year record with a

Table 2 Details of the hydrometric (H) and meteorological (M) observation stations of particular interest with location relative to the hydrometric station and altitude in meters above mean sea-level (see their locations in Fig. 3)

Station name	Station identifier	Station type	Latitude (in °N)	Longitude (in °W)	Altitude (in m)	Location	Ownership
South Hero	28	M	44.626	73.303	34	Upstream	National Climatic Data Center (NCDC)
Rapides Fryer	0	H	45.398	73.241	–	Reference	Environment Canada (EC)

Fig. 4 Graphical representation of the standardized interannual anomalies of hydrometeorological indices obtained from observations at the Rapides Fryer and South Hero stations

spring flow of 1,260 m³/s. However, the flood duration for this year is close to normal with only 49 days of flood. Similar anomalies of $Q_{max,spr}$ and $Fldur$ are observed for most years characterized by high maximum spring flows (i.e. 1983, 1994, 1996 and 1998). In fact, the latter instances have maximum spring flows (Q) equal or superior to the minor flood threshold (i.e. $Q \geq 1,064$ m³/s), which is set empirically by the public security authority of the province of Quebec. Such a flood entails inundations of nearby roads, fields and possibly small damages to homes situated along the river. In contrast, the years 1993 and 1998 exceeded the medium flood severity threshold ($Q \geq 1,221$ m³/s), which typically involves damages of several homes and bridges, poor road conditions, risks of overflowing sewers as well as possible evacuations. Furthermore, the 2011 spring flood of the Richelieu River exceeded the major flood threshold of $Q \geq 1,335$ m³/s, which brings about the same consequences as all the lower flood severity levels in addition to causing damage to large inhabited sectors while threatening people and infrastructures in the extended area due to high water levels and possible electrical power failures.

Meteorological events observed at the South Hero station show interesting characteristics, particularly for 2011.

Observations at the center of the watershed, near Lake Champlain (i.e.: South Hero station), indicate that the 2010–2011 winter and spring seasons were relatively abnormal in terms of the intensity and quantity of precipitation in addition to the variability of temperature. In fact, the cumulative total precipitation fallen between November and March was positively abnormal for 2011 with quantities of 354.7 mm, corresponding to a 14-year recurrence event. Generally, lower cumulative winter and spring precipitations are observed throughout the 31 year record for the South Hero weather station relative to observations taken in the south of the watershed (not shown in this analysis). The 2010–2011 cumulative solid precipitation fallen in *DJF* at South Hero was abnormally high with 199.2 mm, which corresponds to recurrences of approximately once every 12 years. The solid precipitation fallen in March was abnormally high with 64.2 mm, corresponding to a return period of 25 years. The intensity of those snowfall events were particularly important in 2011 for the South Hero station with a maximum cumulative solid precipitation in 5 consecutive days of 63 mm, which corresponds to a return period of 12 years. Out of the 31 year record, 2008 is the only other year with a more abnormal cumulative solid precipitation fallen in winter

Table 3 Marginal distributions and parameters of hydrometeorological indices and stations

Station/ Index	$Q_{max_{spr}}$	Fldur	FrThSeq (DJF)	PrecTOT (NM)	PrecSol (DJF)	PrecSol (M)	PrecLiq (MAM)	SDIILiq (MAM)	P90pLiq (MAM)
Rapides Fryer	Lognormal (6.8, 0.2)	Gamma (0.05, 2.4)	-	-	-	-	-	-	-
South Hero	-	-	Gamma (0.7, 37.9)	GEV (51, 0.2, 225)	Gamma (0.09, 6.8)	Gumbel (15.7, 15)	Inverse gamma (1163, 8.1)	GEV (1.6, 0.2, 6.7)	Gamma (0.9, 14.3)

and abnormal intensity of snowfall similar to that of March 2011. However, the $Q_{max_{spr}}$ and $Fldur$ of 2008 are not as extreme as that of the 2011 near centennial flood. These differences could be attributed in part to the contributing effects of high snowmelt rate and liquid precipitation fallen in spring (see below) in addition to the before-mentioned explicative variables.

Quantities of liquid precipitation fallen in the spring of 2011 were abnormally high at the South Hero station when compared to 2008 and most of the years with high maximum spring flows except for 1983. In fact, the 2011 spring liquid precipitation quantities are characterized by return periods of about 80 years with 375.2 mm at South Hero. The intensity of those rainfall events is positively abnormal and particularly high for South Hero with mean daily intensities of 12.1 mm/wet day and a 90th percentile of liquid precipitation of 31.1 mm/day. These events are extremely rare with recurrences in the order of once every 120 and 260 years, respectively. The combination of these precipitation events with particular temperature variability may increase flood risks. However, the snowmelt rates as calculated by the DD_{melt} index for the month of March and the 2011 spring season were less than what is normally observed for that period of the year. On the other hand, the number of days with complete daily frost/thaw episodes was abnormally low for the 2010–2011 winter season over the entire watershed with only 19 days of $FrThSeq$ at the South Hero station, which corresponds to return periods of 20 years. Such a phenomenon favors snow accumulation during the DJF months, which may result in greater amounts of available snowmelt in spring. A comparison of the 2011 hydrometeorological events with the 31-year mean values of each indicator is summed up in Table 4.

The Kendall’s rank correlation test (1938) is conducted between the various indices for all the stations in order to detect and quantify the dependence between the different hydrometeorological events. Table 5 shows the correlation coefficients between some of these indices for the South Hero station. These Kendall tau coefficients (τ) fall within the range of acceptable correlation coefficients for the use of Clayton and Frank copulas as defined by Michiels and Schepper (2008; i.e. $-1 \geq \tau \leq 1$). The maximum spring flow and the flood duration indicators are significantly positively correlated with each other and that correlation increases with higher flood detection thresholds for the flood duration index. Certain meteorological indices are also positively correlated with each other for all the weather stations such as the 90th percentile and simple daily liquid precipitation in spring. In other words, an increase in intense precipitation events leads to an increase in mean daily precipitation per wet day for the same period. The correlations of the other combinations of indices are not significant for every station. Moreover, pairs of

Table 4 Hydrometeorological events of 2011 compared to their 1981–2011 mean values

Hydrometeorological index	2011	1981–2011 mean value
$Q_{max,spr}$ (m ³ /s)	1,550	918
$Fldur$ (days)	111	49
$PrecTOTNM$ (mm)	355	253
$PrecSolDJF$ (mm)	199.2	76.6
$PrecSolM$ (mm)	64.2	24.7
$PrecLiqMAM$ (mm)	375.2	163.3
$SDIILiqMAM$ (mm/wet day)	12.1	7.3
$P90pLiqMAM$ (mm/day)	31.1	16.8
$FrThSeqDJF$ (days)	19	33

Table 5 Kendall's rank correlation coefficient of hydrometeorological indices for the South Hero meteorological station and the Rapides Fryer hydrometric station under 90 and 95 % confidence intervals in normal and bold fonts, respectively

Indices	$T_{Kendall}$
$Q_{maxspr}-Fldur$	0.483
$Q_{maxspr}-P90precLiq MAM$	0.352
$Q_{maxspr}-SDII precLiq MAM$	0.242
$Q_{maxspr}-PrecSol DJF$	0.258
$Q_{maxspr}-FrThSeq DJF$	0.308
$Q_{maxspr}-PrecLiq MAM$	0.326
$Q_{maxspr}-PrecTOT NM$	0.360
$Q_{maxspr}-R5D precSol March$	0.235
$Q_{maxspr}-R5D precLiq MAM$	0.249
$Q_{maxspr}-R5D precLiq April$	0.258
$Fldur-P90precLiq MAM$	0.440
$Fldur-SDII precLiq MAM$	0.357
$Fldur-PrecLiq MAM$	0.441
$Fldur-R5D precLiq MAM$	0.301
$Fldur-R5D precLiq April$	0.389
$P90precLiq MAM-SDIIprecLiq MAM$	0.666
$P90precLiq MAM-PrecTOT NM$	0.287

indices for which the South Hero weather station has significant correlations are given particular attention for the multivariate analysis, especially if they can be used to anticipate flood occurrences. For example, it can be observed that the maximum spring flow is negatively correlated with the frequency of daily frost/thaw episodes in winter as recorded by the South Hero station. Hence, the number of days for which the temperature fluctuates above and below 0 °C in winter appears to have an inverse effect on the intensity of flow. In other words, less/more thawing days during winter may increase/decrease the quantity of snowmelt water available in spring that contributes to the flood. However, the index for degree days of melt is not

correlated to the peak spring flow nor flood duration indices for any weather station. The latter is unexpected and raises doubts about the suitability of the $DDmelt$ index chosen for the study area.

In summary, the cumulative total precipitation fallen between the months of November and March is positively correlated with the maximum spring flow for the station of interest. Intensity indices of liquid precipitation in spring such as the 90th percentile of rainfall are also positively correlated with $Q_{max,spr}$ and $Fldur$ for several observation stations including that of South Hero. As a result, the spring peak flow increases at Rapides Fryer as the cumulative total precipitation fallen from November through March and the 90th percentile of liquid precipitation in spring increase as well as when the frequency of daily frost/thaw episodes in winter decreases for several weather stations across the LCR watershed. Multivariate analyses of these correlated variables are conducted in the next section for a more in depth understanding of their effect on the Richelieu River flood risks.

4.2 Multivariate analysis

4.2.1 Evaluation of simulation algorithm

In Fig. 5, the performance of the generalized nested Archimedean copula algorithm is analyzed with the use of box-and-whisker plots. The algorithms for the Clayton and Frank copula have proven capable of representing central distributions of the events observed by the South Hero and Rapides Fryer stations while respecting the dependence structure between the variables tested; namely the maximum spring flow, cumulative solid precipitation in winter and the 90th percentile of rainfall in spring. In fact, the inter-quartile ranges (IQR) are quite similar for the observed and simulated values of the three variables; with a slightly underestimated median for the 90th percentile of rainfall in spring for both copulas (Fig. 5). The whiskers, representing values within 1.5 times the IQR of Q_1 and Q_3 , are longer for the simulated meteorological variables, indicating that the simulated extremes and variations are higher than what is observed. The slight skewness in the higher quantiles of the observed $Q_{max,spr}$ index is not represented in the simulated time-series of both copulas due to the low estimated value of the scale parameter for the lognormal distribution associated to that simulated variable. There is a larger number of outliers for the three simulated variables relative to the respective observed data, which is possibly due to the limited number of observations available for the 1981–2011 period compared to the 50,000 simulations generated for this evaluation as argued by Chowdhary et al. (2011).

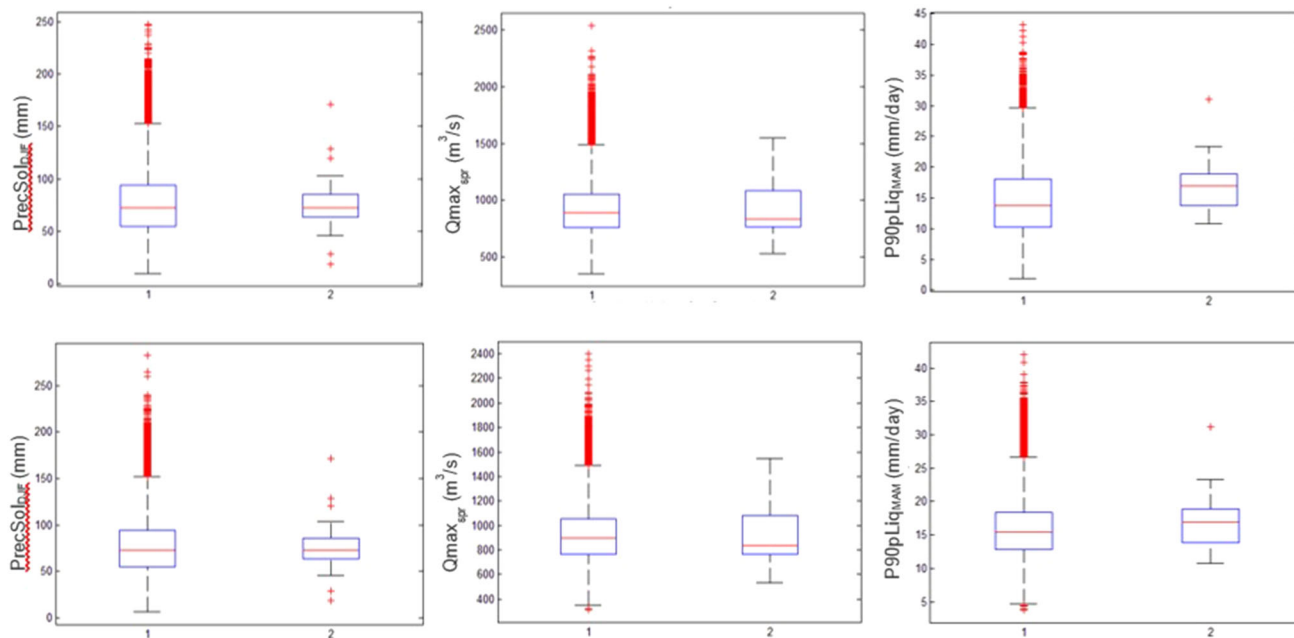


Fig. 5 Box-and-whisker plots of the (1) simulated and (2) observed cumulative solid precipitation in winter (in mm on column 1), maximum spring flow (in m³/s on column 2) and 90th percentile of spring rainfall (in mm/day on column 3) indices for the Clayton (row 1) and Frank (row 2) copulas. The scales on the y-axis are not similar for all indices

Table 6 Simulated and observed univariate exceedance probabilities of the 2011 hydrometeorological events using the Clayton and Frank copulas

Copula family	Index and threshold	Observations	Simulations
Clayton	$Q_{max_{spr}} \geq 1,550 \text{ m}^3/\text{s}$	0.0103	0.0107
	$PrecSolDJF \geq 128 \text{ cm}$	0.0556	0.0577
	$P90pLiqMAM \geq 31 \text{ mm/day}$	0.0038	0.0038
Frank	$Q_{max_{spr}} \geq 1,550 \text{ m}^3/\text{s}$	0.0103	0.0110
	$PrecSolDJF \geq 128 \text{ cm}$	0.0556	0.0556
	$P90pLiqMAM \geq 31 \text{ mm/day}$	0.0038	0.0038

The univariate probabilities of exceeding the 2011 thresholds as estimated from the 3 simulated variables are similar to those obtained by univariate frequency analysis of the observed events of 2011 (Table 6). Therefore, the marginal probability distributions of the variables considered in the proposed copula models are adequately preserved and respected. Figure 6 represents the simulations on 3 and 2-dimensional scatter plots, which are used to analyze the dependence structure between the three variables for the Clayton or Frank copula (Fig. 6a, b). The superimposed scatter plots of the observations and simulations indicate that both sets of simulations (i.e. via Clayton and Frank copulas) appear suitable for this exercise. For low values of the cumulative solid precipitation in winter as well as of the 90th percentile of rainfall in spring, a low value of maximum spring flow is jointly simulated. As expected, when the values of the meteorological indices increase, so does the joint value of

$Q_{max_{spr}}$. However, this increase is not without important variations as can be seen by the scattering of points under more extreme conditions. A similar tendency is observed in the scatter plots of the $Q_{max_{spr}}-PrecSol$ and $Q_{max_{spr}}-P90p$ pairs of simulations. Using the Frank copula (Fig. 6b), the dependence structures are not as clear to interpret as those obtained via the Clayton copula. According to the k -plot and the quadratic distance criteria ($L_{Frank}^2 = 0.0019$ vs $L_{Clayton}^2 = 0.0016$), the Clayton copula is better suited to represent the interdependence of the variables being tested in this exercise. However, further evaluations of the trivariate Frank copula simulation algorithm demonstrated appropriate dependence structures. Overall, the generalized nested simulation algorithm for both copulas is considered adequate for the current study, but the Clayton copula performs best when combined with the method of moments for parameter estimation.

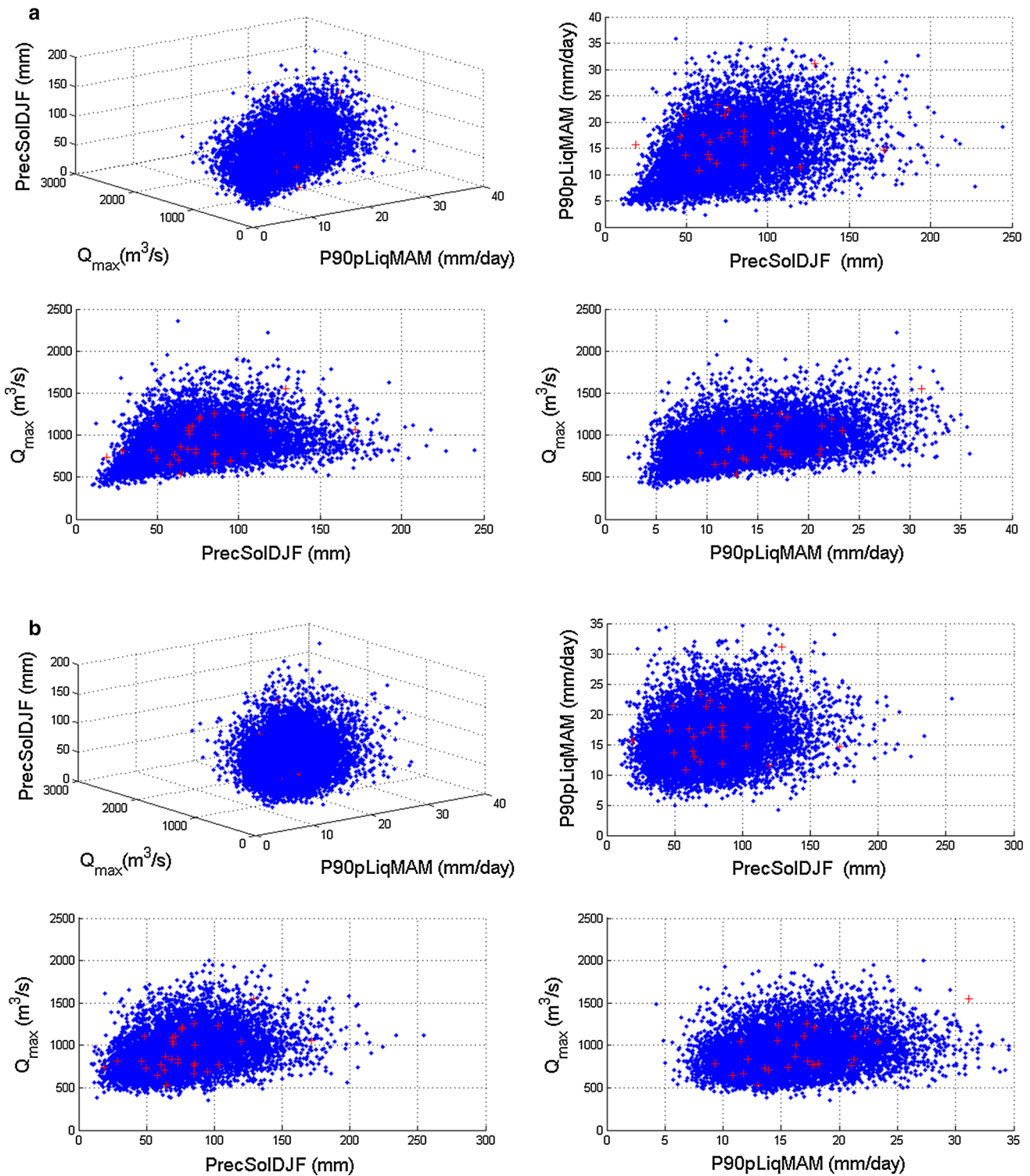


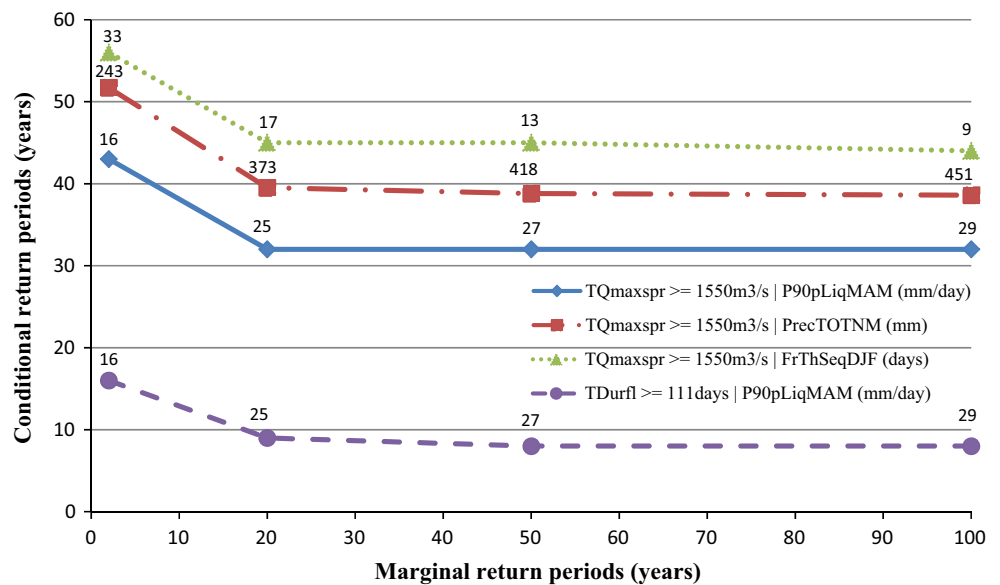
Fig. 6 Graphical representations of $Q_{max_{spr}}$, $PrecSol\ DJF$ and $P90pLiqMAM$ values as simulated by the trivariate nested Clayton (a) and Frank (b) algorithms (dots) in 3-dimensional and 2-dimensional perspectives superimposed by the observed values of these indices (crosses)

4.2.2 Multivariate characterisation of Richelieu River record flood

Bivariate analyses of joint and conditional $Q_{max_{spr}}$ and $Fldur$ return periods for the 2011 spring flood of the

Richelieu River are first conducted in consideration of $PrecTOT$, $P90p$ and $FrThSeq$ meteorological indices for the South Hero and Rapides Fryer stations. The interpretation of conditional return periods requires an understanding of the relationship between the univariate and multivariate

Fig. 7 Conditional return periods for the occurrence of the 2011 flood peak and duration under meteorological conditions corresponding to the 2, 20, 50 and 100-year (T) events obtained from the marginal distributions; the quantiles of the meteorological conditions in their respective units are displayed above each point



estimations of recurrences and risks. The example of the conditional return period of the maximum spring flow in consideration of the 90th percentile of rainfall in spring as a covariate is used to illustrate that interpretation.

The probability that the peak spring flow observed in 2011 occurs is of 0.011 or 1.1 %, which corresponds to a return period of 90 years (i.e. $1/0.011$) according to a univariate frequency analysis. By taking into account additional information from dependent meteorological observations such as the 90th percentile of rainfall in spring, the probabilities that the same spring peak flow occurs can be estimated for different extreme precipitation intensities. For example, by considering the occurrence of a $P90pLiqMAM$ of 16.41 mm/day, which has a return period of only 2 years, the probability that 2011’s $Q_{max,spr}$ is exceeded is of 0.0233 or 2.3 %, corresponding to a return period of 43 years (i.e. $1/0.0233$). In other words, when 2-year extreme spring rainfall intensities occur, there is ~2.3 % probability that 2011’s maximum spring flow also occurs.

In order to determine which meteorological variable has the most influence on the flood characteristics, the conditional return period of the hydrologic events under meteorological conditions of 2, 20, 50 and 100-year return periods are compared. The effect of these meteorological events on the recurrence of the 2011 spring peak flow and flood duration is quantified (i.e. $T_{Q_{max,spr}|X}$ and $T_{Durfl|X}$, respectively, where X is a meteorological condition). In a first case, the conditional return periods of a record flood event is quantified in consideration of the meteorological indices of interest presented in Fig. 7. The 90th percentile of rainfall in spring appears to have the most influence on flood peak occurrences between the meteorological indices and marginal conditions evaluated. The cumulative total precipitation fallen between November and March and the frequency of frost/thaw episodes in winter

also appear to have an effect on the conditional return period of $Q_{max,spr}$, but to a lesser degree. Similarly, the 2011 flood duration is affected by extreme conditions of rainfall intensity in spring since the $Fdur$ conditional return period is lowest under unusual $P90pLiqMAM$ events. In all the evaluated cases, the separate effect by those meteorological phenomena on the conditional return periods of either flood characteristic diminishes considerably for meteorological conditions of marginal return period’s superior to 20 years. The latter may indicate that more than one meteorological condition must be met in order to increase the risk of a spring peak flow and flood duration similar to 2011.

Following the same approach, the 2011 hydrometeorological events are analysed in a bivariate case using the proposed copula algorithm and the observations of the South Hero and Rapides Fryer stations. The joint occurrence of the 2011 maximum spring flow and flood duration (i.e. $1,550 \text{ m}^3/\text{s}$ and 111 days, respectively) is infrequent with an estimated return period of the order of 960 years according to the Clayton copula. Note that those estimates are based on a maximum sample size of 31 years. Consequently, important uncertainties are expected for events with relatively high return periods when using a limited sample size of a few decades. Therefore, estimates of joint return periods should be interpreted as ordinal scales of measurement varying from frequent, normal, rare and extremely rare events and not as precise estimates. On the other hand, the conditional return period of flood duration of 111 days is of 11 years if a maximum spring flow similar to the 2011 record is reached.

Moreover, the maximum spring flow of 2011 has return periods of approximately 51 and 40 years when the observed frequency of daily frost/thaw episodes in winter or the cumulative total precipitation fallen between November and March are similar to those of 2011. Whereas $Q_{max,spr}$ has a

Table 7 Joint and conditional trivariate return periods (in years) of the 2011 hydrometeorological events observed at the South Hero and Rapides Fryer stations

Variables	LL	UL	Copula
$T_{2011}(P90pLiqMAM, PrecTOTNM)$	830	1,470	
$T_{2011}(P90pLiqMAM, Q_{max_{spr}}, PrecTOTNM)$	7, 140	50,000	Clayton
$T_{2011}(Q_{max_{spr}} P90pLiqMAM, PrecTOTNM)$	7	52	
$T_{2011}(PrecTOTNM, FrThSeqDJF)$	133	162	
$T_{2011}(PrecTOTNM, Q_{max_{spr}}, FrThSeqDJF)$	1,852	5,000	Frank
$T_{2011}(Q_{max_{spr}} PrecTOTNM, FrThSeqDJF)$	13	36	

LL and UL correspond to lower and upper limits (respectively) of the 95 % confidence interval for the estimated conditional and joint return periods

lower conditional return period of about 31 years when the 90th percentile of rainfall in spring is similar to the 2011 event observed at South Hero. In other words, a $Q_{max_{spr}} \geq 1,550 \text{ m}^3/\text{s}$ has higher probabilities of occurring when $P90pLiqMAM \geq 31.1 \text{ mm/day}$ then when either $PrecTOTNM \geq 355 \text{ mm}$ or $FrThSeqDJF \geq 19 \text{ days}$. This indicates that the intensity of spring precipitation fallen at the center of the LCR watershed in 2011 had an important influence on its spring peak flow. This can be explained by the abnormally high 90th percentile of rainfall that spring, which had an approximate marginal return period of 260 years. The occurrence of the 2011 flood duration was influenced by these intense precipitation events. In fact, $P90pLiqMAM$ had a prominent effect on the probability of occurrence of the 2011 flood duration since a conditional return period of 8 years is estimated under such extreme spring rainfall intensities.

In order to determine how the combined occurrence of the three meteorological conditions affected the record flood, a trivariate analysis of the conditional $[T_{2011}(a,b,c)]$ and joint $[T_{2011}(a,b,c)]$ return periods of the 2011 maximum spring flow is conducted. Table 7 sums up the lower (LL) and upper (UL) limits of the 95 % confidence interval for the estimated conditional and joint return periods of the hydrometeorological events observed by the South Hero and Rapides Fryer stations. It must be noted that the generalized simulation algorithm for the Frank copula has difficulties converging estimates of joint and conditional trivariate return periods when extreme or very rare events are considered. It is not necessarily the case for the Clayton copula. Hence, when both copulas are suitable, the Clayton copula is used in an attempt to remediate convergence issues.

The results show that the estimated joint return periods are larger than the conditional recurrences as expected from their respective definitions. The bivariate return

period of the joint occurrence of the 90th percentile of spring rainfall and the cumulative total precipitation fallen between November and March 2011 is high, but not as important as the return period of the trivariate joint occurrence of these 2 meteorological conditions with the 2011 maximum spring flow. Whereas the estimated trivariate conditional return period indicates that the occurrence of such a spring peak flow can be very probable when cumulative and intense precipitation conditions similar to those of 2011 are observed. However, the 95 % confidence intervals of these trivariate recurrence estimates are relatively large and must be interpreted with caution.

On the other hand, the occurrence of the frequency of frost/thaw episodes recorded in winter 2011 in conjunction with the cumulative total precipitation fallen in the months from November to March 2011 may have a lesser, but more certain influence on the recurrence of the spring peak flow for that year. The confidence interval of the return period for the joint bivariate occurrence of the 2011 $FrThSeqDJF$ and $PrecTOTNM$ events is much lower than that of the $P90pLiqMAM$ and $PrecTOTNM$ combination of events, which can be explained by the lower marginal recurrences of the former two variables. Overall, those considered indices characterizing the cumulative and punctual intensity of precipitation as well as the variability of temperature in the months prior and during flood-prone seasons appear to be appropriate indicators of extreme peak flow occurrences over the Richelieu River.

4.2.3 Hydrometeorological scenarios

The proposed nested multivariate Archimedean copula approach is used to statistically elaborate multiple plausible and physically founded scenarios of hydrometeorological events causing minor and major floods according to the peak flow thresholds set by the Quebec provincial public security authority. The trivariate conditional and joint return periods of flow occurrences associated with these minor and major flood thresholds (1,064 and 1,335 m^3/s , respectively) are evaluated according to a variety of meteorological events extracted from the observed range of $PrecTOTNM$, $P90pLiqMAM$ and $FrThSeqDJF$ of the available 31-year record. Since no flood duration is associated to the flood conditions set by the public security authority, scenarios are only generated according to flow thresholds. The 95 % confidence interval of the estimated return periods for the joint and conditional trivariate minor and major flood scenarios are presented in Tables 8 and 9.

The return periods of the joint occurrences of the $FrThSeqDJF$, $PrecTOTNM$ and either flood threshold are generally lower than those of the $P90pLiqMAM$, $PrecTOTNM$ and respective flood threshold. The reason for this is that the latter meteorological conditions have much higher extreme values with more scarce occurrences than

Table 8 Lower and upper estimated joint and conditional return periods (years) of trivariate scenarios of minor (left cells) and major (right cells) floods under observed *P90LiqMAM* (mm/day) and *PrecTOTNM* (mm) conditions

	Minor flood					Major flood				
	<i>P90pLiqMAM</i> → <i>PrecTOTNM</i> ↓					<i>P90pLiqMAM</i> → <i>PrecTOTNM</i> ↓				
	9.3	20.2	31.1	9.3	20.2	31.1	9.3	20.2	31.1	
Joint return period	158.2	12–13	515–806	4–4	12–13	515–806	21–22	54–62	1,429–3,125	
	290.3	12–13	694–1,190	12–13	21–23	694–1,190	46–51	75–88	1,515–333	
	422.4	71–83	1,724–4,545	71–83	102–122	1,724–4,545	196–259	249–338	2,500–7,143	
Conditional return period	158.2	4–4	2–3	4–4	2–3	2–3	19–20	11–12	6–12	
	290.3	2–2	1–2	2–2	2–2	1–2	8–9	6–7	3–6	
	422.4	2–2	1–1	2–2	1–1	1–1	4–6	3–4	1–3	

the former combination of events. Moreover, the return periods of trivariate joint occurrences for the minor flood threshold are lower than for the major flood threshold since the latter are in fact, less common. In general, the uncertainties attributed to the estimated return periods of joint occurrences increases substantially with higher values of meteorological conditions. However, that increase is less important for the estimated return periods of the conditional occurrences of flood thresholds. Overall, these scenarios indicate that commonly observed *PrecTOTNM*, *P90pLiqMAM* and *FrThSeqDJF* conditions by the South Hero station can generate minor floods approximately every 4 years or 25 % of the time, while the risk for major flood events is of at most 4 % (i.e. return period of 25 year) under similar conditions. Those risks increase with more rare and extreme meteorological conditions to maximum flood risks ranging from about 50 to 100 %, respectively for major and minor floods. Although such joint trivariate events are rare, these meteorological conditions are appropriate indicators of flood causing scenarios, and can be used for design purposes or for hydrologic sensitivity analysis of the river of interest.

5 Discussion and conclusion

The fully nested multivariate Archimedean copula approach suggested in this study has proven capable of representing the inter-dependence of hydrometeorological phenomena while preserving the marginal distributions of the considered variables. A generalisation of the bivariate simulation algorithms of the Frank and Clayton copulas to the trivariate case was conducted in a nested framework in order to satisfy this objective. This technique preserves the most suitable marginal distribution for each variable instead of assigning a common distribution to all variables.

The evaluation of the approach showed that both the Clayton and Frank simulation algorithms were capable of representing the inter-quartile range of the variables tested as well as the univariate exceedance probability of extreme hydrometeorological observations. However, differences in the number of outliers and extreme values were found between the simulations and observations due in part to the sample size difference of the two datasets. In fact, more abundant observations would benefit the performance of marginal and copula parameter estimations by considering a larger range of the phenomenon’s natural variability. In the absence of a large sample size, the use of the method of moments for parameter estimation was opted since it performs as well as, if not better than, other methods such as the maximum likelihood approach under similar conditions. In addition, the separate estimation of the marginal and copula parameters via the method of moments,

Table 9 Same as Table 8, but for observed *FrThSeqDJF* (days) and *PrecTOTNM* (mm) conditions

	Minor flood				Major flood			
	<i>FrThSeqDJF</i> → <i>PrecTOTNM</i> ↓		<i>PrecTOTNM</i> ↓		<i>FrThSeqDJF</i> → <i>PrecTOTNM</i> ↓		<i>PrecTOTNM</i> ↓	
	53	36	19	19	53	36	19	19
Joint return period	158.2	4–4	5–5	27–30	158.2	24–26	21–23	106–129
	290.3	16–17	17–18	66–77	290.3	74–85	69–79	234–309
	422.4	148–185	158–197	500–781	422.4	625–1,020	588–962	1,667–3,846
Conditional return period	158.2	4–4	3–3	2–2	158.2	15–16	20–22	8–9
	290.3	2–3	2–2	1–2	290.3	9–10	10–12	5–7
	422.4	2–3	2–2	1–2	422.4	7–12	8–13	4–9

although tedious, is more effective than techniques which estimate the entirety of the model’s parameters simultaneously.

Moreover, the simulation algorithms of both copulas represent adequately the dependence structures between variables, while performing particularly better for the Clayton copula in this study. The difference in performance with the Frank copula may lie in the use of a numerical method to solve for the inverse function of the Frank generator, while the inverse of the Clayton generator can be found analytically. In fact, convergence issues were observed for the Frank copula during simulations of extreme values for certain variables. In this case, the Clayton copula was used when both copulas were found similarly suitable, thereby reducing the range of the 95 % confidence interval for the estimated joint and conditional return periods. Nonetheless, the algorithm developed for the Frank copula may be more appropriate when simulating dependencies between variables of less extreme nature and more central values (Genest and Favre 2007). Thus, the fully nested approach presented in this study may be applicable for other Archimedean copulas if desired.

The confidence interval was calculated with a bootstrapping technique developed to maximise the sampling distribution by selecting with replacement 1,000 samples of 50,000 values of each variable from a total of 100,000 simulations. This approach ensured that random simulations are selected for every sample in large enough numbers to estimate high return periods, thereby allowing an adequate assessment of the accuracy of the simulation algorithms. By doing so, practical estimates of the risks and return periods of joint and conditional occurrences of common and scarce bivariate and trivariate hydrometeorological events are successfully produced using the Clayton and Frank copulas. Larger confidence intervals translate in elevated uncertainties and smaller confidence intervals into higher certainty. Therefore, uncertainties associated to return periods of extreme/rare events were larger than to those of more normal/frequent conditions due to the limited availability of observations in the study region and, in turn, due to the limited representation of the natural variability and extremes within the sample. In this case, it is recommended that estimates of joint or conditional return periods be interpreted as ordinal scales of measurements varying from frequent to extremely rare events. Furthermore, the multivariate analysis was limited to the consideration of three variables due to the constraints imposed by the sample size and the nested approach (Table 1). A general rule of thumb in statistical analysis is to have about 4 times more observations than the number of parameters to estimate. In the case of the Richelieu River, the sample size of observations was limited to 31 years hence only 3 variables were considered simultaneously since a minimum of 8

parameters have to be estimated. Hence, the applicability of the proposed methodology and uncertainties associated to its estimates are strongly influenced by the sample size of the observations and by the nature of dependence between the variables of interest (i.e. extreme/central values; magnitude of dependence).

In one application of this tool, a record breaking flood event of the Richelieu River has been characterized in terms of spring peak flow and flood duration risks. It was shown that the combined effect of cumulative total precipitation fallen between November and March with the 90th percentile of rainfall in spring and the frequency of daily frost/thaw episodes in winter explained part of the 2011 extreme spring peak flow and flood duration of the river. However, when extreme meteorological and hydrological events such as the 90th percentile of rainfall and the spring peak flow of 2011 are considered, the proposed simulation algorithm has difficulties estimating the joint and conditional return periods even with the use of the Clayton copula. The limited sample size of observations may be one of the reasons behind this performance issue since relatively small ranges of the natural variability of the phenomena of interest are considered for the calibration of the model. In a subsequent study, it would also be interesting to include a more sophisticated degree days of melt index, which takes into account the snow density or accumulation, elevation and wind speeds as suggested by the National Engineering Handbook of the Natural Resource Conservation Service (2004), since spring floods are often associated to particular snowmelt rates in North America. Moreover, certain variables were stationary with only 90 % confidence. Therefore, it would be interesting to explore a non-stationary approach to the proposed fully nested multivariate Archimedean copula thereby eliminating the effect of any possible trend in the dataset.

In a second application, physically founded and probable flood-causing hydrometeorological scenarios are elaborated. The latter showed that risks of minor and major flood events increased with more extreme meteorological conditions of cumulative precipitation and of frequency of frost/thaw episodes in the months previous to the flood season, especially when extreme spring precipitation intensities also occurred. However, it was also shown that the combination of less abnormal meteorological events can still expose the river to minor and even major flood risks when several particular conditions are met. Such scenarios will be generated for the LCR watershed in a subsequent study and will be given as input to a semi-distributed hydrologic model calibrated over the Richelieu River in order to determine flood risks along the river according to the flow response.

The following conclusions are drawn from this study: (1) Trivariate hydrometeorological flood analysis can be conducted using the proposed simulation algorithms developed

for nested multivariate Archimedean Frank and Clayton copulas if adequate datasets are available; (2) Joint and conditional return periods obtained from such a model can provide valuable information for decision makers in order to elaborate adaptation and flood risk management strategies when appropriate indicators are selected; (3) The consideration of more than one meteorological condition is required to explain extreme flood events since the effect of a single meteorological event is, in some cases, not sufficient to cause record breaking floods such as the 2011 spring flood of the Richelieu River; (4) Cumulative total precipitation and frequency of daily frost/thaw episodes in the months preceding flood-season combined with the extreme intensities of rainfall during the flood-prone spring season constitute major meteorological factors of maximum spring flow for the Richelieu River; And (5) the effect of different combinations of meteorological conditions on flood characteristics can be compared with the use of copulas. In fact, as demonstrated in our study, the 90th percentile of rainfall in spring has an important influence on extreme spring peak flow. This is especially true when particular preconditioned memory of water storage along the Richelieu River and Lake Champlain drainage basin is satisfied through the cumulative total precipitation fallen between the months of November and March or the frequency of frost/thaw cycles in winter. Thus, these three combined meteorological factors should be considered or serve as indicators of high impact floods for the region of study in any intermediate or long-term alert system or risk management scheme. This is of particular importance for the survey of spring flood occurrence in Canada, as the current practice in meteorological services is to use univariate meteorological information or to assume a linear combination of multiple variables, mainly related to precipitation, in order to assess spring flood risks within river basins.

Acknowledgments The authors acknowledge the financial support of the Natural Science and Engineering Research Council of Canada, Environment Canada, and the Institut National de la Recherche Scientifique (centre Eau-Terre-Environnement). Other financial support from the IRIACC (The International Research Initiative on Adaptation to Climate Change)-FACE project (see <http://face.ete.inrs.ca/>) is also acknowledged. IRIACC is managed by the IDRC (International Development Research Center), co-financed by the IDRC, the Canadian Health Research Institutes, the SSHRC (Social Sciences and Humanities Research Council of Canada) and the NSERC (Natural Sciences and Engineering Research Council of Canada). Our gratitude also goes to the ESCER center of Université du Québec À Montréal as well as to the department of Mathematics and Statistics of Université de Moncton for their support. Thank you to the «Data Access Interface» (DAI) project, Environment Canada and to the National Climatic Data Center of the National Oceanic and Atmospheric Administration for access to the hydrometeorological data used in this work.

Open Access This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

References

- Akaike H (1974) A new look at the statistical model identification. *IEEE Trans Autom Control* AC 19(6):716–722
- Antil F, Rouselle J, Lauzon N (2005) Hydrologie : cheminements de l'eau, 2e edn. Presses Internationales Polytechniques, Montreal
- Beniston M (2009) Trends in joint quantiles of temperature and precipitation in Europe since 1901 and projected for 2100. *Geophys Res Lett* 36(L07707):1–6. doi:10.1029/2008GL037119
- Cantin J-F (2011) The Richelieu River Flood Event—Spring 2011: Presentation to the Provincial Flood Forecasting Warning Committee. MSC-Operations Quebec Region, Environment Canada. Black Creek Pioneer Village. October 12 and 13, 2011
- Chowdhary H, Escobar AL, Singh VP (2011) Identification of suitable copulas for bivariate frequency analysis of flood peak and flood volume data. *Hydrol Res* 42(2–3):193–216. doi:10.2166/nh.2011.065
- Clayton DG (1978) A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence. *Biometrika* 65:141–151
- Comité de Concertation et de Valorisation du Bassin de la Rivière Richelieu (2011) Plan Directeur de L'Eau—Portrait et Diagnostique du Bassin Versant de la Rivière Richelieu. COVABAR, Beloeil, March 2011
- Cunnane C (1978) Unbiased plotting positions—review. *J Hydrol* 37:205–222
- Efron B, Tibshirani JR (1993) An introduction to bootstrapping. Chapman & Hall, New York 456 p
- El Adlouni S, Ouarda TBMJ (2008) Étude de la loi conjointe débit-niveau par les copules: cas de la rivière Châteauguay. *Revue Canadienne de Génie Civil* 35:1128–1137
- El Adlouni S, Bobée B, Ouarda TBMJ (2008) On the tails of extreme event distributions in hydrology. *J Hydrol* 355(1–4):16–33
- Favre A-C, El-Adlouni S, Perrault L, Thiémond N, Bobée B (2004) Multivariate hydrological frequency analysis using copulas. *Water Resour Res* 40(W01101):1–12. doi:10.1029/2003WR002456
- Fischer MJ, Köck C, Schlüter S, Weigert F (2007) Multivariate copula models at work: outperforming the “desert island copula”? Department of Statistics and Econometrics, Erlangen-Nuremberg University, Germany. No. 79
- Frank MJ (1979) On the simultaneous associativity of $F(x, y)$ and $x + y - F(x, y)$. *Aequ Math* 19:194–226
- Gachon P, St-Hilaire A, Ouarda T, Nguyen VTV, Lin C, Milton J, Chaumont D, Goldstein J, Hessami M, Nguyen T, Selva F, Nadeau M, Roy P, Parishkura D, Major N, Choux M, Bourque A (2005) A first evaluation of the strength and weaknesses of statistical downscaling methods for simulating extremes over various regions of Eastern Canada. Sub-component, Climate Change Action Fund (CCAF), Environment Canada, Final Report, Montreal, Quebec, Canada, 209 pp
- Genest C, Favre A-C (2007) Everything you always wanted to know about copula modeling but were afraid to ask. *J Hydrol Eng* 12:347–368. doi:10.1061/(ASCE)1084-0699(2007)12:4(347)
- Genest C, Gendron M, Bourdeau-Brien M (2009) The advent of copulas in finance. *Eur J Financ* 15(7–8):609–618
- Groupe de Travail International du Lac Champlain et de la Rivière Richelieu (2013) Plan d'Étude Provisoire: Détermination des Mesures Visant à Atténuer les Inondations et leurs Répercussions dans le Bassin du Lac Champlain et de la Rivière Richelieu. Préparer pour la Commission mixte internationale. February 2013
- Gumbel EJ (1958) Distribution à plusieurs variables dont les marges sont données. *C R Acad Sci* 246:2717–2720
- Hao Z, AghaKouchak A, Phillips JT (2013) Changes in concurrent monthly precipitation and temperature extremes. *Environ Res Lett* 8:1–6. doi:10.1088/1748-9326/8/3/034014
- Helsel DR, Hirsch RM (2002) Statistical methods in water resources techniques of water resources investigations. Book 4, Chapter A3, U.S. Geological Survey, 522 pp
- Hofert M (2008) Sampling Archimedean copulas. *Comput Stat Data Anal* 52(12):5163–5174
- Huard D, Favre A-C, Poulin A, Évin G (2004) Faisabilité de l'analyse fréquentielle trivariée des pointes, volumes et durées de crue à l'aide de copules archimédiennes. Internal Report No. I-151, INRS-Eau, Terre et Environnement, September 10th, 2004, 37 pp
- Jeong ID, Sushama L, Khaliq MN, Roy R (2013) A copula-based multivariate analysis of RCM projected changes to flood characteristics for northeastern Canada. *Clim Dyn*. doi:10.1007/s00382-013-1851-4
- Karmakar S, Simonovic SP (2009) Bivariate flood frequency analysis. Part 2: a copula-based approach with mixed marginal distributions. *J Risk Manag* 2:32–34
- Kendall MG (1938) A new measure of rank correlation. *Biometrika* 30(1/2):81–93
- Lavallée D, Roy L, Marche C (2000) Un système de prévision appliqué aux crues subites de la rivière Châteauguay. *Can J Civ Eng* 27:1311–1315
- Lawford RG, Prowse TD, Hogg WD, Warkentin AA, Pilon PJ (1995) Hydrometeorological aspects of flood hazards in Canada. *Atmos Ocean* 33(2):303–328
- Liebscher E (2009) Construction of asymmetric multivariate copulas. *J Multivar Anal* 99:2234–2250
- Marshall WA, Olkin I (1988) Families of multivariate distributions. *J Am Stat Assoc* 83(403):834–841. doi:10.2307/2289314
- Mazouz R, Assani AA, Quessy JF, Lègaré G (2012) Comparison of the interannual variability of spring floods characteristics of tributaries of the St. Lawrence River in Quebec (Canada). *Adv Water Resour* 35:110–120
- Michiels F, Schepper AD (2008) A copula test space model : how to avoid the wrong copula choice. *Kybernetika* 44:864–878
- Natural Resources Conservation Service (2004) National engineering handbook—part 630 hydrology. United-States Department of Agriculture, July 2004
- Nazemi A, Elshorbagy A (2012) Application of copula modelling to the performance assessment of reconstructed watersheds. *Stoch Environ Res Risk Assess* 26:189–205
- Nelsen BR (2006) An introduction to copulas, 2nd edn. Springer, New York 272 p
- Salvadori G, De Michele C (2007) On the use of copulas in hydrology: theory and practice. *J Hydrol Eng* 12:369–380
- Schwarz G (1978) Estimating the dimension of a model. *Ann Stat* 6:461–464
- Zhang L, Singh VP, FASCE (2006) Bivariate flood frequency analysis using the copula method. *J Hydrol Eng* 11(2):150–164
- Sklar A (1959) Fonction de répartition à n dimensions et leurs marges. *Publ Inst Stat Univ Paris* 8:229–231
- Turcotte R, Fortier F, Fillion T-CF, Fortin V, Roy A, Royer A (2010) Simulation hydrologique des derniers jours de la crue de printemps: le problème de la neige manquante. *Hydrol Sci J* 55(6):872–882
- Villarini G, Smith JA, Baeck ML, Krajewski WF (2011) Examining flood frequency distributions in the Midwest U.S. *J Am Water Resour Assoc* 47(3):447–463
- Zhang L, Singh VP (2006) Bivariate rainfall and runoff analysis using entropy and copula theories. *Entropy* 14(9):1784–1812