mathematics

MDPI

Article

Estimation and Prediction of Record Values Using Pivotal Quantities and Copulas

Jeongwook Lee [1], Joon Jin Song [2], Yongku Kim [3] and Jung In Seo [4,*]

[1] Department of Statistics, Daejeon University, Daejeon 34519, Korea; gunzion12@gmail.com
[2] Department of Statistical Science, Baylor University, Waco, TX 76798, USA; Joon_Song@baylor.edu
[3] Department of Statistics, Kyungpook National University, Daegu 41566, Korea; kim.1252@knu.ac.kr
[4] Division of Convergence Education, Halla University, Wonju, Gangwon-do 26404, Korea
* Correspondence: leehoo1928@gmail.com

Received: 21 August 2020; Accepted: 24 September 2020; Published: 1 October 2020

Abstract: Recently, the area of sea ice is rapidly decreasing due to global warming, and since the Arctic sea ice has a great impact on climate change, interest in this is increasing very much all over the world. In fact, the area of sea ice reached a record low in September 2012 after satellite observations began in late 1979. In addition, in early 2018, the glacier on the northern coast of Greenland began to collapse. If we are interested in record values of sea ice area, modeling relationships of these values and predicting future record values can be a very important issue because the record values that consist of larger or smaller values than the preceding observations are very closely related to each other. The relationship between the record values can be modeled based on the pivotal quantity and canonical and drawable vine copulas, and the relationship is called a dependence structure. In addition, predictions for future record values can be solved in a very concise way based on the pivotal quantity. To accomplish that, this article proposes an approach to model the dependence structure between record values based on the canonical and drawable vine. To do this, unknown parameters of a probability distribution need to be estimated first, and the pivotal-based method is provided. In the pivotal-based estimation, a new algorithm to deal with a nuisance parameter is proposed. This method allows one to reduce computational complexity when constructing exact confidence intervals of functions with unknown parameters. This method not only reduces computational complexity when constructing exact confidence intervals of functions with unknown parameters, but is also very useful for obtaining the replicated data needed to model the dependence structure based on canonical and drawable vine. In addition, prediction methods for future record values are proposed with the pivotal quantity, and we compared them with a time series forecasting method in real data analysis. The validity of the proposed methods was examined through Monte Carlo simulations and analysis for Arctic sea ice data.

Keywords: C- and D-vine copulas; confidence interval; exponentiated Gumbel distribution; pivotal quantity; record values

1. Introduction

Extreme weather and air pollution have been received steadily increasing attention over the past decade. Examples include extreme temperatures, the exceedances of flood peaks, and pollutant concentrations deviating considerably from expected average levels. In such cases, predicting observations more extreme than the current extreme values is an important issue. The topic of record values was introduced by Chandler [1], and Balakrishnan et al. [2] established some recurrence relationships for single and double moments of lower record values from the Gumbel distribution. Coles and Tawn [3] analyzed a daily rainfall series for modeling the extremes of a rainfall process in

the context of the record values. Wang et al. [4] provided approaches to constructing exact confidence intervals (CIs) for unknown parameters in the family of proportional reversed hazard distributions based on lower record values. Seo and Kim [5] provided classical and Bayesian approaches to inference for the Gumbel distribution based on lower record values. Seo and Kim [6] presented an objective Bayesian analysis method for the two-parameter Rayleigh distribution based on record values. Seo and Kim [7] proposed an entropy inference method based on an objective Bayesian approach for when observed record values have a two-parameter logistic distribution.

There are two types of the record values. If the observation is greater than all the preceding observations it is called the upper record. On the other hand, if the observation is smaller than all the preceding observations then it is called the lower record. There are a few situations where lower record values are of special interest. For example, Arctic sea ice greatly affects climate change, and the reduction of Arctic sea ice is a very serious issue. In this case, only sea ice extent less than the previous one is of interest and recorded, which poses a problem of predicting the next sea ice extent. The lower record value is described as follows. Let $\{X_1, X_2, \ldots\}$ be a sequence of independent and identically distributed (iid) random variables with a probability density function (PDF) $f(x)$ and a cumulative distribution function (CDF) $F(x)$. Then, X_j is a lower record value if $X_j < X_i$ for every $i < j$. The indexes for which lower record values occur are given by the record times $\{L(k), k \geq 1\}$, where $L(k) = \min\left\{j | j > L(k-1), X_j < X_{L(k-1)}, \right\}, k > 1$, with $L(1) = 1$. Therefore, a sequence of lower record values is denoted by $\left\{X_{L(k)}, k = 1, 2, \ldots\right\}$ from the original sequence $\{X_1, X_2, \ldots\}$. These record values are heavily related to each other. Imani and Braga-Neto [8] proposed an efficient finite-horizon feedback controller similar to an optimal linear quadratic Gaussian estimator for partially-observed Boolean dynamical systems as a general class of nonlinear state-space model. Imani et al. [9] proposed an optimal Bayesian filter approach to the problem of recursive estimation in partially-observed Boolean dynamical systems. To establish the relationship between the record values, a copulas approach based on the canonical (C) and drawable (D) vine is proposed in this article.

Copulas have recently received much attention as a modeling tool for describing the dependency structure of multivariate data. The notion of a copula function can be found in Sklar [10]. One of the advantages of copula models is to build a variety of dependence structures based on existing parametric or non-parametric models of the marginal distributions. The copulas can be described as follows. Let F be the d-dimensional function of the random vector $\boldsymbol{X} = (X_1, \ldots, X_d)^T$ with marginal distributions $F_{X_1}(x), \ldots, F_{X_d}(x)$. Then there exists a copula C such that for all $\boldsymbol{x} = (x_1, \ldots, x_d)^T \in [-\infty, \infty]^d$,

$$F(\mathbf{x}) = C\left(F_{X_1}(x), \ldots, F_{X_d}(x)\right),$$

where the copula C is unique if $F_{X_1}(x), \ldots, F_{X_d}(x)$ are continuous by Sklar's theorem (Sklar [10], 1959). In this case, the copula C can be interpreted as the distribution function of a d-dimensional random variable on $[0, 1]^d$ with uniform marginal distributions. Tsung et al. [11] conducted a comprehensive literature review of statistical transfer learning methods focusing on statistical models and statistical methodologies, including a Gaussian copula. Rocher et al. [12] proposed a generative copula-based method that can elaborately estimate the likelihood that a particular person will be correctly re-identified, even in a very incomplete dataset. Vine copulas were introduced by Joe [13] to overcome limitations of standard multivariate copulas in higher dimensions, where standard multivariate copulas lack the flexibility of accurately modeling the dependence.

Aas et al. [14] described statistical inference techniques for the C- and D-vine copulas. Berg and Aas [15] and Fischer et al. [16] showed the excellence of the D-vine copula approach, compared to alternative copulas in constructing higher dimensional dependency structures.

To the best of our knowledge, modeling the dependence structure between record values numerically has been little explored. In this paper, we propose an approach with which to model the relationship between the record values based on C- and D-vine copulas and to predict future record values.

The remainder of the article is organized as follows. Section 2 introduces C- and D-vine copulas and provides pivotal-based approaches to estimate the model parameters and to predict future record values by proposing a new algorithm to deal with a nuisance parameter. Section 3 presents simulation studies to validate the proposed approaches. We applied the methods to Arctic sea ice data; see Section 4. Concluding remarks and some discussions are in Section 5.

2. Methods

Let $f(x;\theta)$ and $F(x;\theta)$ be the marginal density and distribution functions of X, respectively, where θ is an unknown parameter. Then a schematic diagram (Figure 1) of our method is given by

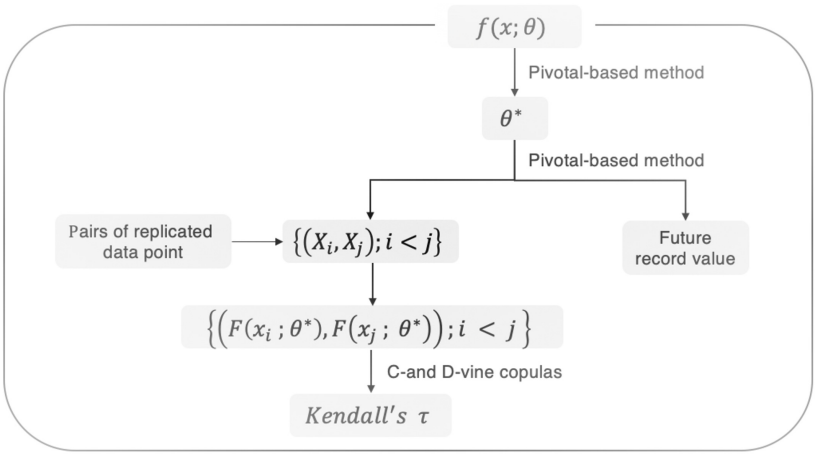

Figure 1. Schematic diagram.

The C -and D-vine copulas are first described in the following subsection.

2.1. C- and D-Vine Copulas

A vine is a flexible graphical model that decomposes a multivariate probability distribution into bivariate copulas, where each pair-copula can be chosen independently from the others [14]. This article considers C- and D-vine copulas to model the relationship between record values based on C- and D-vine copulas.

The C-vine decomposition is given by

$$\begin{aligned} f(x_1,\ldots,x_n) &= f_1(x_1)\prod_{i=2}^{n}\prod_{k=1}^{i-1} c_{i-k,i|1,\ldots,x_{i-k-1}} f_k(x_i) \\ &= \prod_{k=1}^{n} f_k(x_k)\left(\prod_{j=1}^{n}\prod_{i=1}^{k-1} c_{i-k,i|1,\ldots,i-k-1}\right) \\ &= \prod_{k=1}^{n} f_k(x_k)\left(\prod_{j=1}^{n-1}\prod_{i=1}^{n-j} c_{j,j+i|1,\ldots,j-1}\right). \end{aligned}$$

Then, we can specify the pairs of the d-dimensional C-vine copula model in the following order:

$$\begin{aligned} &(1,2),(1,3),(1,4),\ldots,(1,d), && \text{(tree 1)} \\ &(2,3|1),(2,4|1),\ldots,(2,d|1), && \text{(tree 2)} \\ &\cdots, \\ &(d-1,d|1,\ldots,d-2). && \text{(tree } d-1\text{)}, \end{aligned}$$

which has vectors of length $d(d-1)/2$, where d is the number of variables.

The D-vine decomposition is given by

$$\begin{aligned} f(x_1,\ldots,x_n) &= \prod_{i=2}^{n} f_{i|x_1,\ldots,x_{i-1}}(x_i) f_1(x_1) \\ &= \prod_{l=1}^{n} f_l(x_l) \left(\prod_{i=2}^{n} \prod_{j=1}^{i-2} c_{ji|j+1,\ldots,i-1} \right) \prod_{i=2}^{n} c_{i-1,i} \\ &= \prod_{l=1}^{n} f_l(x_l) \left(\prod_{i=2}^{n-1} \prod_{j=1}^{n-i} c_{j,i|j+1,\ldots,j+i-1} \right). \end{aligned}$$

Similarly, the pairs of the d-dimensional D-vine copula model are specified in the following order:

$$\begin{aligned} &(1,2),(2,3),(3,4),\ldots,(d-1,d), && \text{(tree 1)} \\ &(1,3|2),(2,4|3),\ldots,(d-2,d|d-1), && \text{(tree 2)} \\ &(1,4|2,3),(2,5|3,4),\ldots,(d-3,d|d-2,d-1), && \text{(tree 3)} \\ &\cdots, \\ &(1,d|2,\ldots,d-1). && \text{(tree } d-1\text{)} \end{aligned}$$

To measure the dependence of each pair-copula, we consider tree 1 that can be employed to obtain Kendall's τ (Nelsen [17], 2006) given by

$$\begin{aligned} \tau_C &= 4 \left(\int_0^1 \int_0^1 C(u_i,u_j) dC(u_i,u_j) \right) - 1 \\ &= 4E\left[C(U_i,U_j) \right] - 1, \end{aligned}$$

where $C(u_i,u_j)$ is a bivariate copula function for $u_i,u_j \in [0,1]$.

In a C- and D-vine, consider the exponentiated Gumbel distribution (EGD) with the CDF

$$F(x) = \left(e^{-e^{-x/\sigma}} \right)^{\lambda}, \quad -\infty < x < \infty,\ \sigma,\ \lambda > 0, \tag{1}$$

where σ and λ are the scale and shape parameters. Then, u_i is the value of the marginal distribution of $X_{L(i)}$ with

$$F_{X_{L(i)}}(x) = e^{-H(x)} \sum_{j=0}^{i-1} \frac{[H(x)]^j}{j!}, \tag{2}$$

where $H(x) = -\log F(x)$ (Ahsanullah [18] and Arnold et al. [19]). That is, for $d-1$ pairs of data points $\{(x_{L(i)}, x_{L(j)}); i<j\}$, the corresponding couples $\{(u_i,u_j); i<j\}$ can be computed from the marginal distribution (2). In addition, the marginal density function of $X_{L(i)}$ is given by

$$f_{X_{L(i)}}(x) = \frac{1}{\Gamma(k)} \left[-\log F(x) \right]^{i-1} f(x).$$

Note that it is necessary to estimate σ and λ for computing the values of u_i and u_j.

2.2. Pivotal-Based Approach

Here we present a pivotal-based method to estimate the parameters of the CDF (1). First, a lemma is introduced to deal with nuisance parameters in order to establish the relationship between record values.

Lemma 1. *Let $X_{L(1)}, \cdots, X_{L(k)}$ be the lower record values from the CDF (1). Then,*

(a) $T_k(\sigma, \lambda) = 2\lambda e^{-X_{L(k)}/\sigma}$ *has a* χ^2 *distribution with* $2k$ *degrees of freedom;*

(b) $V(\sigma) = \frac{e^{(X_{L(1)} - X_{L(k)})/\sigma} - 1}{k-1}$ *has a F distribution with* $2k-2$ *and* 2 *degrees of freedom;*

(c) $W(\sigma) = \frac{2\left(\sum_{j=1}^{k} X_{L(j)} - kX_{L(k)}\right)}{\sigma}$ *has a* χ^2 *distribution with* $2k-2$ *degrees of freedom.*

Proof. Let $X_{L(1)}, \cdots, X_{L(k)}$ be the lower record values from the CDF (1). Then, we have

$$Z_i = -\log F(x_{L(i)}),\ i = 1, \ldots, k$$

that have a standard exponential distribution, and that leads to the following spacings

$$S_i = Z_i - Z_{i-1},\ i = 1, \ldots, k (Z_0 \equiv 0),$$

which is independent and identically distributed as the standard exponential distribution with mean 1. From the spacings, the pivotal quantities

$$\begin{aligned} T_j(\sigma, \lambda) &= 2\sum_{i=1}^{j} S_i \\ &= 2\lambda e^{-X_{L(j)}/\sigma},\ j = 1, \ldots, k \end{aligned} \tag{3}$$

are derived, which are independent random variables such that there is an χ^2 distribution with $2j (j = 1, \ldots, k)$. From (3), the pivotal quantity (a) is easily proved. In addition, let $T_{-1,k}(\sigma, \lambda) = 2\sum_{i=2}^{k} S_i$. Then, the pivotal quantity (b) is proved as

$$\begin{aligned} V(\sigma) &= \frac{T_{-1,k}(\sigma, \lambda)/(2k-2)}{T_1(\sigma, \lambda)/2} \\ &= \frac{e^{(X_{L(1)} - X_{L(k)})/\sigma} - 1}{k-1} \end{aligned}$$

because $T_{-1,k}(\sigma, \lambda)$ and $T_1(\sigma, \lambda)$ are independent random variables, given the fact that S_i $(i = 1, \ldots, k)$ are independent and identically random values, as mentioned earlier; both have a χ^2 distribution with $2k-2$ and 2 degrees of freedom, respectively. On top of that, we can derive the following pivotal quantities by using Lemma 2 of Wang et al. [4] in (3)

$$\begin{aligned} U_j &= \left(\frac{T_j(\sigma, \lambda)}{T_{j+1}(\sigma, \lambda)}\right)^j \\ &= e^{j(X_{L(j+1)} - X_{L(j)})/\sigma},\ j = 1, \ldots, k-1 \end{aligned}$$

which are independent and identically distributed as the uniform distribution on the interval $(0, 1)$. Then, the pivotal quantity (c) is proved as

$$\begin{aligned} W(\sigma) &= -2\sum_{j=1}^{k-1} \log U_j \\ &= \frac{2\left(\sum_{j=1}^{k} X_{L(j)} - kX_{L(k)}\right)}{\sigma}. \end{aligned}$$

□

From Lemma 1(c), the unique solution σ^* is given by

$$\sigma^* = \frac{2\left(\sum_{j=1}^{k} X_{L(j)} - kX_{L(k)}\right)}{Q_W},$$

where Q_W follows a χ^2 distribution with $2k-2$ degrees of freedom. Then, for any $0 < \alpha < 1$, an exact $100(1-\alpha)\%$ CI for σ based on $W(\sigma)$ is given by

$$\left(\sigma^*_{([(\alpha/2)N])},\ \sigma^*_{([(1-\alpha/2)N])}\right),$$

where $\sigma_{([\alpha N])}$ is the $[\alpha N]$th smallest of $\{\sigma_l^*\}$. Note that the exact CI is the equal-tail CI because it splits the probability equally, putting $\alpha/2$ in each tail of the distribution. For any $0 < \alpha < 1$, an exact $100(1-\alpha)\%$ CI with the shortest-length based on σ^* is given by

$$\left(\sigma^*_{(l^*)},\ \sigma^*_{(l^*+[(1-\alpha)N])}\right),$$

where l^* is chosen so that

$$\sigma^*_{(l^*+[(1-\alpha)N])} - \sigma^*_{(l^*)} = \min_{1\le l\le N-[(1-\alpha)N]} \left[\sigma^*_{(l+[(1-\alpha)N])} - \sigma^*_{(l)}\right].$$

Similarly, the unique solution from $V(\sigma)$ in Lemma 1 is given by

$$\sigma^* = \frac{X_{L(1)} - X_{L(k)}}{\log\left[1 + (k-1)Q_F\right]},$$

where Q_F follows a F distribution with $2k-2$ and 2 degrees of freedom. Then, with the same argument, the exact $100(1-\alpha)\%$ equal-tailed and shortest CIs for σ based on $V(\sigma)$ can be constructed. In Section 3, it is found that $W(\sigma)$ provides a more efficient CI than $V(\sigma)$ in terms of average lengths (ALs) of the CIs through Monte Carlo simulations, as in the case of Seo and Kim [5].

For λ, we have that

$$\lambda = \frac{T_k}{2e^{-X_{L(k)}/\sigma}} \tag{4}$$

by putting $T_k = T_k(\sigma, \lambda)$ in Lemma 1. In addition, let $g(W, \boldsymbol{x_L})$ be the unique solution of $W(\sigma) = W$ for $W > 0$, where $\boldsymbol{x_L} = \left(x_{L(1)}, \cdots, x_{L(k)}\right)$. Then, by substituting $g(W, \boldsymbol{x_L})$ for σ in (4), the following generalized quantity is given by

$$\psi = \frac{T_k}{2e^{-X_{L(k)}/g(W,\boldsymbol{x_L})}}.$$

The existing literature (Wang et al. [4] and Wang et al. [20]) supposed that W has a χ^2 distribution with $2k-2$ degrees of freedom, and then obtained the percentiles of ψ generating W and T_k independently from the χ^2 distribution with $2k-2$ and $2k$ degrees of freedom, respectively, although W and T_k are not independent. As an alternative, the following algorithm is proposed to obtain the percentiles of ψ.

Step 1. Generate $Q_{\chi^2,1}, Q_{\chi^2,2}, \ldots, Q_{\chi^2,k}$ from a χ^2 distribution with two degrees of freedom.
Step 2. Compute $T_j = \sum_{i=1}^{j} Q_{\chi^2,i}$ for $i = 1, \ldots, k$.
Step 3. Compute $W = 2\sum_{j=1}^{k-1} \log\left(\frac{T_k}{T_j}\right)$ and solve the equation $W(\sigma) = W$ for σ to obtain $g(W, \boldsymbol{x_L})$.
Step 4. Compute ψ.
Step 5. Repeat $N(\geq 10,000)$ times.

From the algorithm, the equal-tailed and shortest CIs for λ based on ψ are given by

$$\left(\psi_{([(\alpha/2)N])},\ \psi_{([(1-\alpha/2)N])}\right)$$

and

$$\left(\psi_{(l^*)},\ \psi_{(l^*+[(1-\alpha)N])}\right),$$

respectively, where l^* is chosen so that

$$\psi_{(l^*+[(1-\alpha)N])} - \psi_{(l^*)} = \min_{1\le l\le N-[(1-\alpha)N]} \left[\psi_{(l+[(1-\alpha)N])} - \psi_{(l)}\right].$$

2.3. Prediction

Let $X_{L(r)}(r = k+1, k+2, \cdots)$ be the future lower record values. Then, for any $0 < \alpha < 1$, the conditional quantile is given by

$$\begin{aligned} X_{L(r),\alpha} \mid x_{L(k)} &= F^{-1}_{X_{L(r)}|X_{L(k)}}(\alpha) \\ &= \inf\{y \in \mathbb{R} | F_{X_{L(r)}|X_{L(k)}}(y) \ge \alpha\}, \end{aligned}$$

where $F_{X_{L(r)}|x_{L(k)}}(\cdot)$ is the conditional distribution of $X_{L(r)}$ given $x_{L(k)}$. However, the quantile cannot be obtained numerically because it does not have closed forms. Instead, we propose a pivotal approach based on the following lemma.

Lemma 2. *Let $Y_C = \log F(x_{L(k)}) - \log F(x)$ in the conditional density function of $X_{L(r)}$ given $x_{L(k)}$ be defined by Ahsanullah* [18] *as*

$$f_{X_{L(r)}|x_{L(k)}}(x) = \frac{\left[\log F(x_{L(k)}) - \log F(x)\right]^{r-k-1} f(x)}{\Gamma(r-k)F(x_{L(k)})}, \quad x_{L(r)} < x_{L(k)}. \tag{5}$$

Then, Y_C has a gamma distribution with the parameters $(r-k, 1)$.

Proof. Let $Y_C = \log F(x_{L(k)}) - \log F(x)$ in (5). Then, the Jacobian of the transformation is

$$\begin{aligned} J &= \frac{\mathrm{d}}{\mathrm{d}y_C} x \\ &= \frac{F(x_{L(k)})}{f(x)} e^{-y}, \end{aligned}$$

and the density function of Y_C is

$$f_{Y_C}(y) = \frac{1}{\Gamma(r-k)} y^{r-k-1} e^{-y},$$

which is the probability density function of a gamma distribution with parameters $(r-k, 1)$. □

With the same argument as Section 2, an algorithm for obtaining the Markov-chain Monte-Carlo (MCMC) samples $X^i_{L(r)}(i = 1, \ldots, N)$ based on the pivotal quantity Y_C is provided as follows.

Step 1. Generate Y_C from Gam$(r - k, 1)$.
Step 2. Compute

$$X_{L(r)} \mid x_{L(k)} = -\sigma \log \left(\frac{Y_C}{\lambda} + e^{-x_{L(k)}/\sigma} \right).$$

Step 3. Repeat steps 1 and 2, N times.

3. Simulation Study

A simulation study was performed to examine the validity of the proposed pivotal-based approach in terms of the coverage probabilities (CPs) and ALs of the proposed confidence intervals (CIs). The lower record values with sizes $k = 6(2)12$ were generated from the standard EGD distribution with $\lambda = 0.5, 0.8$, and 1.5. To construct the 95% exact CIs described in Section 2.2, $N = 20{,}000$ MCMC samples were generated, and the corresponding CPs and average lengths (ALs) were computed over 10,000 simulations. The results are reported in Table 1 along with those for the classical inference (see Proof) (Appendix A) for comparison. Table 1 shows that the CIs using MCMC samples have nearly same results as those using the classical method, and all considered CIs are well matched to their corresponding nominal levels; however, the CIs based on $W(\sigma)$ have shorter length than those based on $V(\sigma)$. In addition, all ALs decrease with an increase in the size of record values k. For ALs, the CIs with the shortest-lengths have shorter lengths than those with equal-tails, as expected.

Table 1. Coverage probabilities (CPs) (average lengths (ALs)) of CIs for σ and λ.

		σ								λ	
		Equal-Tails				**Shortest**				**Equal-Tails**	**Shortest**
		$V(\sigma)$		$W(\sigma)$		$V(\sigma)$		$W(\sigma)$		ψ	
λ	k	**Classical**	**MCMC**	**Classical**	**MCMC**	**Classical**	**MCMC**	**Classical**	**MCMC**	**MCMC**	
0.5	6	0.948(3.094)	0.948(3.093)	0.950(2.600)	0.949(2.565)	0.951(2.623)	0.950(2.603)	0.952(2.214)	0.950(2.194)	0.950(2.942)	0.942(2.440)
	8	0.949(2.456)	0.948(2.455)	0.950(1.957)	0.949(1.923)	0.951(2.172)	0.949(2.150)	0.951(1.740)	0.949(1.712)	0.954(2.791)	0.946(2.344)
	10	0.948(2.123)	0.948(2.120)	0.950(1.618)	0.949(1.604)	0.950(1.925)	0.947(1.905)	0.952(1.475)	0.950(1.454)	0.952(2.731)	0.947(2.268)
	12	0.948(1.915)	0.948(1.919)	0.948(1.405)	0.948(1.401)	0.951(1.765)	0.948(1.760)	0.949(1.301)	0.948(1.297)	0.950(2.580)	0.944(2.174)
0.8	6	0.948(3.094)	0.948(3.093)	0.950(2.600)	0.949(2.565)	0.951(2.623)	0.950(2.603)	0.952(2.214)	0.950(2.194)	0.950(3.517)	0.942(3.006)
	8	0.949(2.456)	0.948(2.455)	0.951(1.957)	0.949(1.923)	0.951(2.172)	0.949(2.150)	0.951(1.740)	0.949(1.712)	0.953(3.406)	0.944(2.942)
	10	0.948(2.123)	0.948(2.120)	0.950(1.618)	0.949(1.604)	0.950(1.925)	0.947(1.905)	0.952(1.475)	0.950(1.454)	0.954(3.370)	0.948(2.887)
	12	0.948(1.915)	0.948(1.919)	0.948(1.405)	0.948(1.401)	0.951(1.765)	0.948(1.760)	0.949(1.301)	0.948(1.297)	0.950(3.225)	0.940(2.801)
1.5	6	0.948(3.094)	0.948(3.093)	0.950(2.600)	0.949(2.565)	0.951(2.623)	0.950(2.603)	0.952(2.214)	0.950(2.194)	0.950(4.561)	0.940(4.068)
	8	0.949(2.456)	0.948(2.455)	0.951(1.957)	0.949(1.923)	0.951(2.172)	0.949(2.150)	0.951(1.740)	0.949(1.712)	0.953(4.486)	0.945(4.027)
	10	0.948(2.123)	0.948(2.120)	0.950(1.618)	0.949(1.604)	0.950(1.925)	0.947(1.905)	0.952(1.475)	0.950(1.454)	0.957(4.482)	0.943(4.007)
	12	0.948(1.915)	0.948(1.919)	0.948(1.405)	0.948(1.401)	0.951(1.765)	0.948(1.760)	0.949(1.301)	0.948(1.297)	0.952(4.346)	0.939(3.939)

4. Application: Arctic Sea Ice

Sea ice maintains the Earth's average temperature by reflecting solar energy and keeping the polar regions cool. Currently, the Arctic is warming faster than any other region on earth. The warming of the Arctic Circle leads to a decrease in sea ice, which again causes warming of the Arctic Circle. In addition, it causes global weather changes such as summer heat waves, winter cold waves, and heavy snow. These climate changes are leading to disturbances of ecosystems formed around Arctic sea ice and changes in habitats. For this reason, the importance of sea prediction systems to cope with climate change is increasing. The National Aeronautics and Space Administration (NASA) reported that the area covered by Arctic sea ice has decreased by about ten percent in the last 30 years (Figure 2).

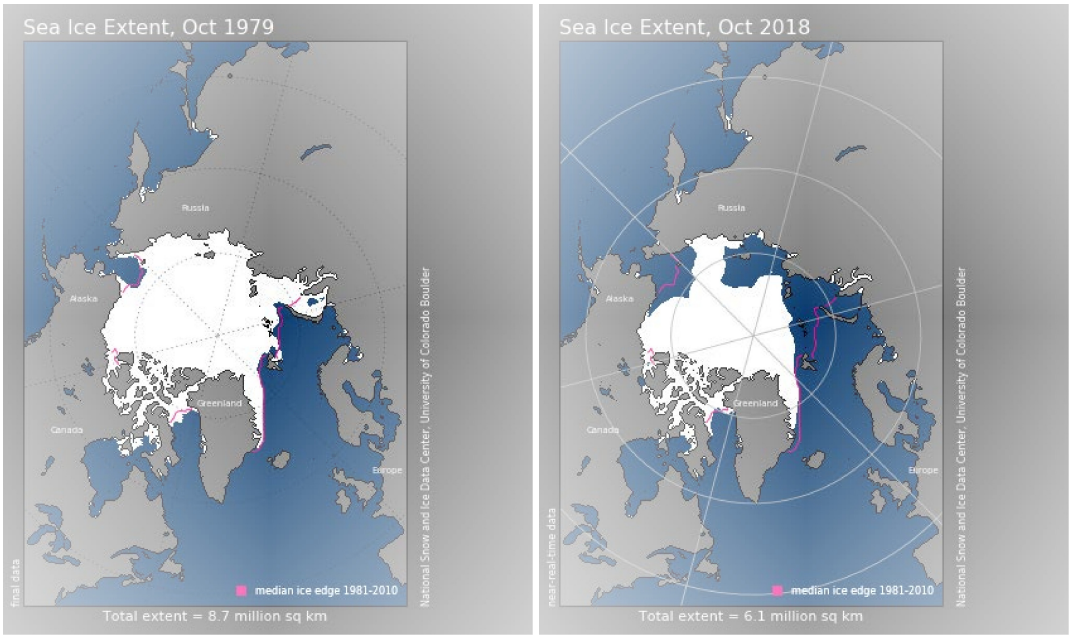

Figure 2. Sea ice extent in October 1979 (**left**) and October 2018 (**right**).

This section analyzes the smallest annual Arctic sea ice extent (see Table 2) from October 1978 to October 2018 extracted from the National Snow & Ice Data Center (NSIDC).

Table 2. Observed record values from Arctic sea ice data

i	1	2	3	4	5	6	7	8	9
$x_{L(i)}$	3.95	2.66	2.47	2.32	2.19	2.17	2.05	1.60	1.29

To measure goodness of fit of the EGD, the replicated data of observed lower record value $x_{L(i)}$ were generated from its marginal density function with σ^* and ψ. All results were obtained by generating $N = 50{,}000$ MCMC samples. In addition, based on the results in the previous simulation study, σ^* from $W(\sigma)$ was only considered in this data analysis.

The 95% confidence region for the replicated data was plotted in Figure 3. It was found that the confidence regions decreased as the record value of the smallest annual Arctic sea ice decreased. The correlation coefficient between the observed and expected lower record values indicates a strong association.

To examine the relationship between observed record values, Figure 4 plots the first trees of the C- and D-vines with the best copula function in terms of the Akaike information criterion (AIC) and corresponding Kendall τ.

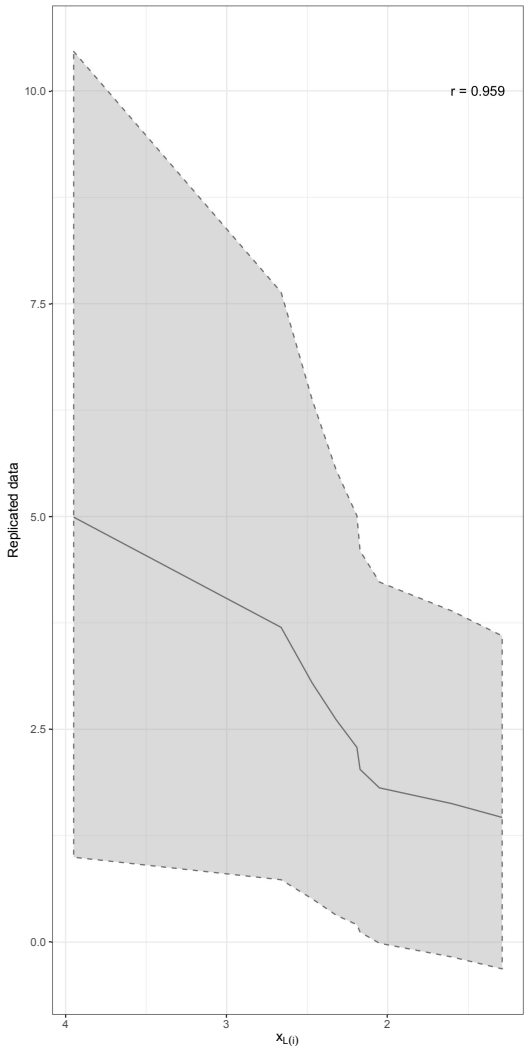

Figure 3. 95% Confidence region of the replicated data. The solid line represents the mean of the replicated data and r is the correlation coefficient of the mean and observed lower record values.

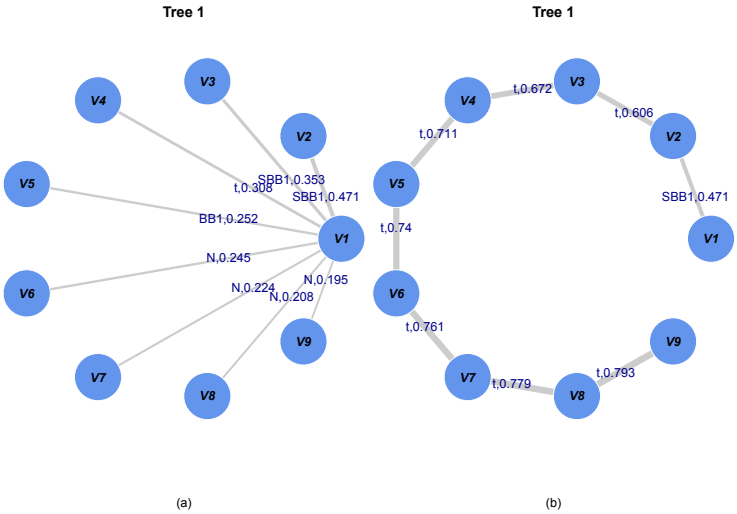

Figure 4. (**a**) First tree of C-vine for the observed record values; (**b**) first tree of D-vine for the observed record values. The labels are the best pair-copula families and corresponding Kendall's τ values. For example, N, t, BB1, and SBB1 represent Gaussian, Student t, Clayton Gumbel, and survival Clayton Gumbel copula, respectively.

Note that the AIC is defined as $-2\ln(L) + 2k$, where L is the likelihood function and k is the number of estimated parameters of the model. Therefore, the smaller the AIC, the better. The entire result for the relationships between observed lower record values is reported in Figure 5. It is shown that the observed lower record values have a positive dependence on each other. In addition, the Kendall's τ values increase as the interval between the lower record times decreases. That indicates that $x_{L(i)}$ and $x_{L(j)}$ such that $j - i = 1$ for $j \neq i$ have the strongest dependency in terms of the Kendall τ. It is worth noting that the strength of dependency between $x_{L(i)}$ and $x_{L(j)}$ such that $j - i = 1$ for $j \neq i$ becomes stronger as the lower record times increase.

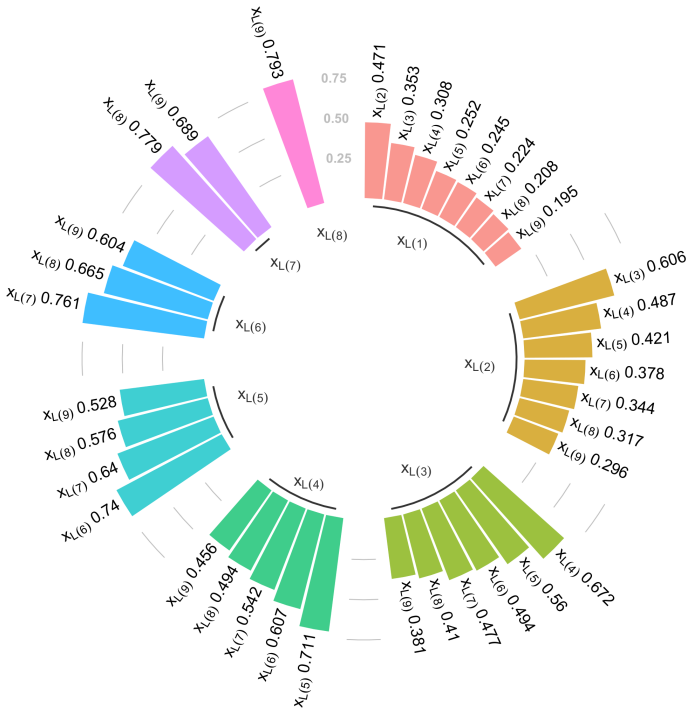

Figure 5. Circular plot for Kendall's τ between two paired record values.

The 95% exact CIs for σ and λ are reported in Table 3, which shows a similar pattern to the simulation results.

Table 3. CIs for σ and λ.

		σ		λ
		$V(\sigma)$	$W(\sigma)$	$W(\sigma)$
Equal-tails	Classical	(0.462, 2.672)	(0.631, 2.635)	-
	MCMC	(0.465, 2.675)	(0.631, 2.619)	(9.747, 77.629)
Shortest	Classical	(0.349, 2.330)	(0.526, 2.333)	-
	MCMC	(0.347, 2.321)	(0.507, 2.302)	(6.130, 65.137)

For prediction, the last lower record value $x_{L(9)}$ was assumed to be unknown, and a time series analysis was conducted, in which it was expected that differences of the observed lower record values could yield a stationary time series because the observed lower record values had a decreasing pattern. In fact, the ARIMA (0, 1, 0) model was chosen as the best model in terms of the AIC from an ARIMA (p, d, q) model, where p is the autoregressive (AR) model order, d is the difference order, and q is the moving average (MA) model order. Table 4 and Figure 6 present the results for future record values of the least annual Arctic sea ice. Table 4 shows that there is little difference in measures of center such as the mean and median for the predictions of the future lower record values based on the pivotal quantity Y_C.

Table 4. Prediction results.

	Mean	Median	Equal-Tails	Shortest
$X^*_{L(9)} \mid x_{L(8)}$	1.457	1.516	(0.987, 1.599)	(1.118, 1.600)
$X^{TS}_{L(9)}$	1.407	-	(1.071, 1.850)	-
$X^*_{L(10)} \mid x_{L(8)}$	1.336	1.401	(0.735, 1.583)	(0.876, 1.600)
$X^{TS}_{L(10)}$	1.237	-	(0.840, 1.821)	-
$X^*_{L(11)} \mid x_{L(8)}$	1.231	1.298	(0.552, 1.560)	(0.693, 1.591)
$X^{TS}_{L(11)}$	1.087	-	(0.677, 1.746)	-

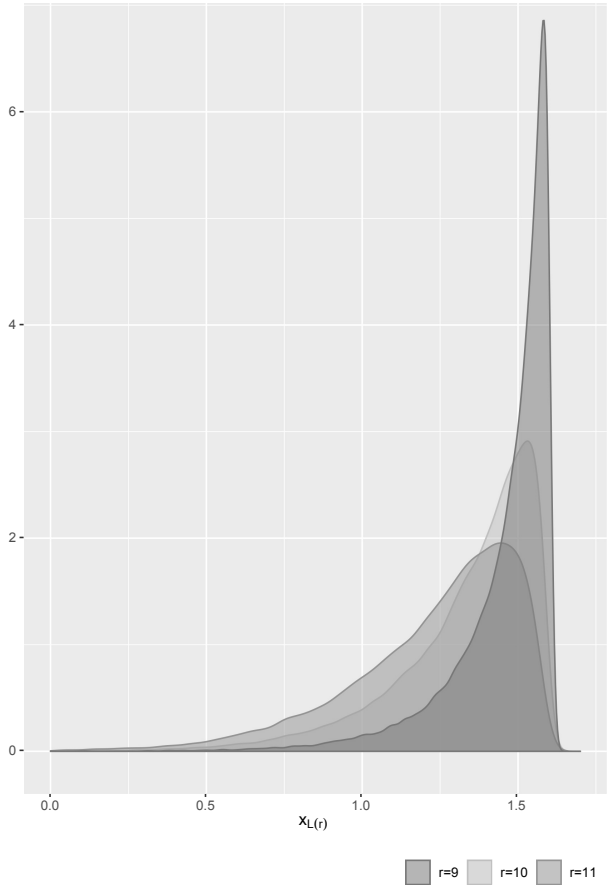

Figure 6. Estimated kernel density functions for $X^*_{L(r)} \mid x_{L(8)}$.

For the last lower record value, the ARIMA (0, 1, 0) model provides a closer predictive value than the mean of $X^*_{L(9)}|x_{L(8)}$ to the actual value of 1.29, while the PI from the ARIMA (0, 1, 0) model has a longer length than that for $X_{L(r)}|x_{L(k)}$ based on the pivotal quantity Y_C. Finally, Figure 6 shows that as the future record time $L(r)$ increases, the variance of the predicted future record value from the conditional density function increases.

5. Conclusions

This article proposed a copula approach with which to model the dependence structure between record values from the EGD and to predict future lower record values using a pivotal-based method. In the pivotal-based method, a new algorithm for dealing with a nuisance parameter has been proposed; it not only is very computationally convenient in constructing exact CIs with the shortest lengths, but also provides very satisfactory results in terms of the CPs and ALs, compared with the classical method. In the approach based on the C- and D-vine copulas, we chose the best copula model in terms of the AIC among 40 paircopula families and it showed very intuitive and reasonable results

in analysis based on real data. An interesting point is that the strength of the dependency between $x_{L(i)}$ and $x_{L(j)}$ such that $j - i = 1$ for $j \neq i$ becomes strong as the lower record times increase in real data analysis. The proposed method is applicable to recording values of other real data that have a probability distribution if the CDF of the probability distribution has a closed form, such as an extreme value distribution. The prediction results of this paper indicate that we should be alert to the decrease in Arctic sea ice extent. In future studies, we envision extending this work to predict the size and decreasing rate of Arctic sea ice extent in real time.

Author Contributions: J.I.S. conceived and designed the research; J.I.S. and J.L. analyzed the data and interpreted the results; J.I.S., J.L., J.J.S., and Y.K. wrote the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (Ministry of Education) (number 2019R1I1A3A01062838).

Acknowledgments: We are grateful to the editor-in-chief, associate editor, and anonymous referees for their helpful comments.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. Proof

Proof. For any $0 < \alpha < 1$, we have

$$\begin{aligned} 1 - \alpha &= P\left[\theta_1(\alpha) < V(\sigma) < \theta_2(\alpha)\right] \\ &= P\left[\frac{X_{L(1)} - X_{L(k)}}{\log\left[1 + (k-1)\theta_2(\alpha)\right]} < \sigma < \frac{X_{L(1)} - X_{L(k)}}{\log\left[1 + (k-1)\theta_1(\alpha)\right]}\right]. \end{aligned} \tag{A1}$$

Then, the interval length in (A1) is given by

$$L = \left\{\frac{1}{\log\left[1 + (k-1)\theta_1(\alpha)\right]} - \frac{1}{\log\left[1 + (k-1)\theta_2(\alpha)\right]}\right\}\left(X_{L(1)} - X_{L(k)}\right)$$

and the corresponding expected interval is given by

$$EL = \left\{\frac{1}{\log\left[1 + (k-1)\theta_1(\alpha)\right]} - \frac{1}{\log\left[1 + (k-1)\theta_2(\alpha)\right]}\right\}\sigma(\psi_{(1)} - \psi_{(k)})$$

because

$$\begin{aligned} E(X_{L(i)}) &= \int_{-\infty}^{\infty} x f_{X_{L(i)}}(x)\mathrm{d}x \\ &= \sigma\left[\log(\lambda) - \psi(k)\right], \end{aligned}$$

where $\psi(\cdot)$ is the digamma function. The equal-tailed CI based on $V(\sigma)$ is obtained setting $\theta_1(\alpha) = F_{1-\alpha/2,(2k-2,2)}$ and $\theta_2(\alpha) = F_{\alpha/2,(2k-2,2)}$ in (A1) because $V(\sigma)$ has the F distribution with $2k-2$ and 2 degrees of freedom. To find $\theta_1(\alpha)$ and $\theta_2(\alpha)$ that minimizes the length such that

$$\int_{\theta_1(\alpha)}^{\theta_2(\alpha)} g(t)dt = 1 - \alpha, \tag{A2}$$

where $g(\cdot)$ is the PDF of the F distribution with $2(k-1)$ and 2 degrees of freedom, we have

$$\frac{dL}{d\theta_1(\alpha)} = \left[-h(\theta_1(\alpha)) + h(\theta_2(\alpha))\frac{d\theta_2(\alpha)}{d\theta_1(\alpha)}\right]\left(X_{L(1)} - X_{L(k)}\right)$$

and

$$\frac{d\theta_2(\alpha)}{d\theta_1(\alpha)} = \frac{g(\theta_1(\alpha))}{g(\theta_2(\alpha))},$$

so that

$$\frac{dL}{d\theta_1(\alpha)} = \left[-h(\theta_1(\alpha)) + h(\theta_2(\alpha))\frac{g(\theta_1(\alpha))}{g(\theta_2(\alpha))}\right](X_{L(1)} - X_{L(k)}),$$

where

$$h(\theta_1(\alpha)) = \frac{k-1}{[1+(k-1)\theta_1(\alpha)]\left[\log\left(1+(k-1)\theta_1(\alpha)\right)\right]^2},$$
$$h(\theta_2(\alpha)) = \frac{k-1}{[1+(k-1)\theta_2(\alpha)]\left[\log\left(1+(k-1)\theta_2(\alpha)\right)\right]^2}.$$

It follows that the minimum occurs at

$$h(\theta_1(\alpha))g(\theta_2(\alpha)) = h_2(\theta_2(\alpha))g(\theta_1(\alpha)). \tag{A3}$$

That is, we choose $\theta_1(\alpha)$ and $\theta_2(\alpha)$ that satisfy the conditions (A2) and (A3) to construct the shortest CI for σ based on $V(\sigma)$.

Similarly, for any $0 < \alpha < 1$, the equal-tailed CI based on $W(\sigma)$ is obtained setting $\theta_1(\alpha) = \chi^2_{1-\alpha/2,2k-2}$ and $\theta_2(\alpha) = \chi^2_{\alpha/2,2k-2}$ in

$$\begin{aligned} 1-\alpha &= P\left[\theta_1(\alpha) < W(\sigma) < \theta_2(\alpha)\right] \\ &= P\left[\frac{2\left(\sum_{j=1}^{k} X_{L(j)} - kX_{L(k)}\right)}{\theta_2(\alpha)} < \sigma < \frac{2\left(\sum_{j=1}^{k} X_{L(j)} - kX_{L(k)}\right)}{\theta_1(\alpha)}\right] \end{aligned}$$

and the shortest CI can be constructed choosing $\theta_1(\alpha)$ and $\theta_2(\alpha)$ that satisfy

$$\int_{\theta_1(\alpha)}^{\theta_2(\alpha)} g(t)dt = 1-\alpha$$

and

$$[\theta_1(\alpha)]^2 g(\theta_1(\alpha)) = [\theta_2(\alpha)]^2 g(\theta_2(\alpha)).$$

□

References

1. Chandler, K.N. The distribution and frequency of record values. *J. R. Stat. Soc. Ser. B* **1952**, *14*, 220–228. [CrossRef]
2. Balakrishnan, N.; Ahsanullah, M.; Chan, P.S. Relations for single and product moments of record values from Gumbel distribution. *Stat. Probab. Lett.* **1992**, *15*, 223–227. [CrossRef]
3. Coles, S.G.; Tawn, J.A.A. Bayesian analysis of extreme rainfall data. *J. R. Stat. Soc. Ser. C* **1996**, *45*, 463–478. [CrossRef]
4. Wang, B.X.; Yu, K.; Coolen, F.P. Interval estimation for proportional reversed hazard family based on lower record values. *Stat. Probab. Lett.* **2015**, *98*, 115–122. [CrossRef]
5. Seo, J.I.; Kim, Y. Statistical inference on Gumbel distribution using record values. *J. Korean Stat. Soc.* **2016**, *45*, 342–357. [CrossRef]

6. Seo, J.I.; Kim, Y. Objective Bayesian analysis based on upper record values from two-parameter Rayleigh distribution with partial information. *J. Appl. Stat.* **2017**, *44*, 2222–2237. [CrossRef]
7. Seo, J.I.; Kim, Y. Objective Bayesian entropy inference for two-parameter logistic distribution using upper record values. *Entropy* **2017**, *19*, 208–219. [CrossRef]
8. Imani, M.; Braga-Neto, U.M. Finite-horizon LQR controller for partially-observed Boolean dynamical systems. *Automatica* **2018**, *95*, 172–179. [CrossRef]
9. Imani, M.; Dougherty, E.R.; Braga-Neto, U. Boolean Kalman filter and smoother under model uncertainty. *Automatica* **2020**, *111*, 108609. [CrossRef]
10. Sklar, M. Fonctions de répartition *á* n dimensions et leurs marges. *Publ. Inst. Stat. Univ. Paris* **1959**, *8*, 229–231.
11. Tsung, F.; Zhang, K.; Cheng, L.; Song, Z. Statistical transfer learning: A review and some extensions to statistical process control. *Qual. Eng.* **2018**, *30*, 115–128. [CrossRef]
12. Rocher, L.; Hendrickx, J.M.; De Montjoye, Y.A. Estimating the success of re-identifications in incomplete datasets using generative models. *Nat. Commun.* **2019**, *10*, 1–9. [CrossRef] [PubMed]
13. Joe, H. Families of $m-$variate distributions with given margins and $m(m-1)/2$ bivariate dependence parameters. *Lect. Notes Monogr. Ser.* **1996**, *28*, 120–141.
14. Aas, K.; Czado, C.; Frigessi, A.; Bakken, H. Pair-copula constructions of multiple dependence. *Insur. Math. Econ.* **2009**, *44*, 182–198. [CrossRef]
15. Berg, D.; Aas, K. Models for construction of multivariate dependence: A comparison study. *Eur. J. Financ.* **2009**, *15*, 639–659.
16. Fischer, M.; Köck, C.; Schlüter, S.; Weigert, F. An empirical analysis of multivariate copula models. *Quant. Financ.* **2009**, *9*, 839–854. [CrossRef]
17. Nelsen, R.B. *An Introduction to Copulas*, 2nd ed.; Springer: New York, NY, USA, 2006.
18. Ahsanullah, M. *Record Statistics*; Nova Science Publishers, Inc.: New York, NY, USA, 1995.
19. Arnold, B.C.; Balakrishnan, N.; Nagaraja, H.N. *Records*; Wiley: New York, NY, USA, 1998.
20. Wang, B.X.; Yu, K.; Jones, M.C. Inference under progressively Type II right-censored sampling for certain lifetime distributions. *Technometrics* **2010**, *52*, 453–460. [CrossRef]