

Research Paper

Extraction of joint probabilistic distribution functions of characteristics of precipitation using a four-dimensional D-vine tree structureMaryam Shafaei¹,Rasoul Mirabbasi Najafabadi^{2*},¹ Ph. D in Water Resources Engineering, Department of Water Engineering, Tabriz University, Tabriz, Iran.² Associate Professor. Department of Water Engineering, Shahrekord University, Shahrekord, Iran.*- Corresponding Author Email: mirabbasi_r@yahoo.com

10.22125/IWE.2021.285684.1507

Received:
May.10.2021
Accepted:
September.22.2021
Available online:
October.05.2022

Keywords:
Copula function,
D-vine,
Multivariate
Frequency analysis,
Pair-copula,
Precipitation
characteristics.

Abstract

Multivariate frequency analysis of hydrological phenomena by considering the dependence between the basic characteristics of these phenomena will lead to more their accurate estimation. Due to the high flexibility provided by vine tree copulas in problems with dimension greater than two. In this study the D-vine function is used to determine the four-dimensional probabilistic distribution function of main characteristics of the precipitation events of Cremona station in Italy (including maximum rainfall intensity, total rainfall depth, duration of wet period and dry period). First, due to the significant dependence between the main characteristics of precipitation events and also using their permutation, D-vine tree structures were obtained. After fitting the various Archimedean and elliptical copula families to the pair-copulas of each D-vine tree structure, the most suitable copula families were determined for fitting the pair-copulas of each D-vine structure by the maximum log-likelihood, Akaike (AIC) and Bayesian Information Criterion (BIC). Then, in order to evaluate the accuracy of the four-dimensional probabilistic distribution functions of the important characteristics of precipitation events, the mentioned functions were compared with the corresponding four-dimensional empirical copulas. Finally, the M-R-D-L four-dimensional D-vine structure according to the evaluation criteria of $R^2 = 0.991$, $RMSE = 0.031$, and $MAE = 0.024$, was selected as the most appropriate structure for constructing of the joint distribution function of main characteristics of precipitation in Cremona station.

* **Corresponding Author:** Rasoul Mirabbasi Najafabadi**Address:** Department of Water Engineering,
Shahrekord University, Shahrekord, Iran**Email:** mirabbasi_r@yahoo.com**Tel:** 0913333275

1. Introduction

The use of copula functions in high-dimensional problems increases the computational load and mathematical complexity and each of the multidimensional copulas has its limitations, but the vine copula structure provides necessary flexibility in high-dimensional problems by using breaking down multivariate distributions into two-dimensional distributions. Also, vine copula structure lacks the weaknesses of these copulas. Joe 1996 for the first time proposed the pair-copula structure for a multivariate structure in the form of distribution functions (Joe, 1996). Bedford and Cook proposed a graphical model called Normal Vine for high-dimensional distributions (Bedford and Cooke, 2001, 2002).

2. Materials and Methods

The method of decomposing the four-dimensional distribution density function by the four-dimensional D-vine structure is shown in Figure (1). In Figure (1), each square represents a node and each line between two squares represents an edge. Each node corresponds to a variable and each edge corresponds to a pair-copula density. If d is the number of variables, all of the decompositions are performed by $(d-1)/2$ edge or the same pair-copulas density. The labels of the edges of the tree T_{i+1} are created by connecting the nodes connected to the tree T_i (i is the number of the trees) and the condition of connecting the two edges in the tree T_{i+1} is sharing the two edges in a node in the tree T_i (Aas et al., 2009). According to Figure (1), the difference between the first tree T_1 and the trees T_2 and T_3 is that the nodes of the first tree are unconditional, while the second and third trees contain the conditional nodes. In the graphical structure of the D-Vine model, the variables are placed linearly (in a row) and each variable will be connected to one variable or two variables. According to Figure (1), the first tree T_1 consists four variables (nodes) s_1, s_2, s_3 and s_4 and the C_{12}, C_{23} and C_{34} unconditional edges. But as shown in Figure (1), the nodes of the tree T_2 contain the $F_{2|1}, F_{3|2}$ and $F_{4|3}$ conditional distribution functions. Furthermore, the nodes of the tree T_3 contain the $F_{13|2}$ and $F_{24|3}$ conditional distribution functions, which are obtained of the partial derivative of the edges of the first tree i.e. C_{12}, C_{23}, C_{34} , and the edges of the second tree i.e. $C_{13|2}$ and $C_{24|3}$.

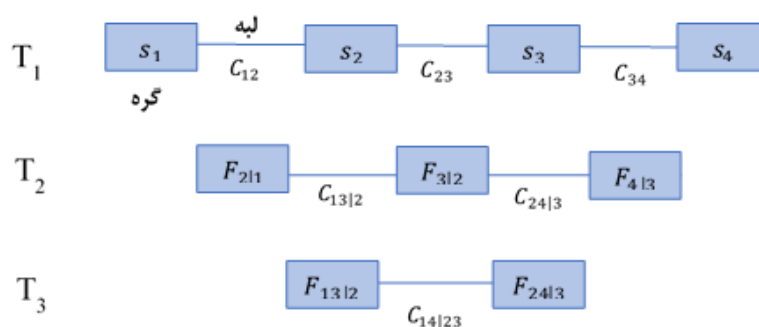


Fig 1-the trees of four-dimensional D-vine structure

3. Results

The aim of this study is to achieve an optimal four-dimensional D-Vine structure to determine the joint four-dimensional distribution of important characteristics of precipitation events at Cremona station in Italy. According to what has been explained, the optimal structures, which include a combination of different copula families for fitting on each pair-copulas of the D-Vine four-dimensional structure, were obtained. Among the D-Vine four-dimensional structures, The M-R-D-L D-Vine structure was identified as the most accurate D-Vine structure according to the log-likelihood criteria and the AIC and BIC criteria.

4. Discussion and Conclusion

Finally, studies show that: 1- Using the main characteristics of precipitation with the highest correlation in the first tree does not lead to achieving the optimal structure of D-Vine and a structure that necessarily has pairs-variables with the strongest correlation in the first tree, it was not recognized as the best structure. 2- Using pairs-variables with the maximum correlation in the first tree does not increase the tau-kendal between the pair-conditional variables of the second and third trees of the D-Vine structure. 3- Considering the physical characteristics of each of the important characteristics of hydrological phenomena, it is necessary to perform all permutations between the variables to achieve the best connection, which ultimately leads to the most accurate D-Vine distribution.

5. Six important references

1. Aas, K., Czado, C., Frigessi, A and Bakken, H. 2009. Pair-copula constructions of multiple dependence. *Insurance: Mathematics and Economics*. 44 (2): 182–198.
2. Ayantobo, O.O., Li, Y. and Song, S. 2019. Multivariate Drought Frequency Analysis using Four-Variate Symmetric and Asymmetric Archimedean Copula Functions. *Water Resources Management*. 33: 103–127.
3. Bedford, T. and Cooke, R. 2001. Probability density decomposition for conditionally dependent random variables modeled by vines. *Annals of Mathematics and Artificial Intelligence*. 32 (1): 245–268.
4. Bedford, T. and Cooke, R. 2002. Vines – A new graphical model for dependent random variables. *Annals of Statistics*. 30 (4): 1031–1068.
5. Brechmann, E. C., Czado, C. and Aas, K. 2012. Truncated regular vines in high dimensions with applications to financial data. *Canadian Journal of Statistics*. 40 (1): 68-85.
6. De Michele, C., Salvadori, G., Passoni, G., and Vezzoli, R. 2007. A multivariate model of sea storms using copulas. *Coastal Engineering*. 54 (10): 734–751.

Conflict of Interest

Authors declared no conflict of interest.

Acknowledgments

Carlo Demichele for supporting data of this research.

استخراج توابع توزیع احتمالاتی توأم مشخصه‌های بارش با استفاده از ساختار درختی دی- واین چهاربعدی

مریم شفاei^۱، رسول میرعباسی نجف آبادی^۲

تاریخ ارسال: ۱۴۰۰/۰۲/۲۰

تاریخ پذیرش: ۱۴۰۰/۰۶/۳۱

مقاله پژوهشی

چکیده

تحلیل فراوانی چند متغیره پدیده‌های هیدرولوژیکی، با در نظر گرفتن وابستگی بین مشخصه‌های اساسی این پدیده‌ها، منجر به تخمین دقیق‌تر آنها خواهد شد. با توجه به انعطاف‌پذیری بالایی که مفصل‌های درختی واین در مسائل با ابعاد بالاتر از دو ارائه داده‌اند، در این مطالعه از تابع مفصل دی- واین جهت تعیین تابع توزیع احتمالاتی توأم چهاربعدی مشخصه‌های مهم رویدادهای بارش ایستگاه سریمونا واقع در کشور ایتالیا شامل بیشینه شدت بارش (M)، عمق کل بارش (R)، مدت زمان دوره مرطوب (L) و مدت دوره خشک (D) استفاده شده است. بدین منظور، در ابتدا با توجه به معنی‌داری وابستگی بین مشخصه‌های مهم رویدادهای بارش و همچنین با استفاده از جایگشت آنها ساختارهای منتخب درختی دی- واین حاصل گردید. پس از برآزش مفصل‌های مختلف خانواده‌های ارشمیدسی و بیضوی بر جفت- مفصل‌های هر یک از ساختارهای درختی دی- واین، مناسب‌ترین خانواده‌های مفصل جهت برآزش بر جفت- مفصل‌های هر یک از ساختارهای دی- واین بواسطه معیارهای حداکثر لوگ درست- نمائی، اطلاعات آکائیکه (AIC) و اطلاعات بیزین (BIC) تعیین گردیدند. سپس به منظور ارزیابی دقت توابع توزیع احتمالاتی توأم چهاربعدی مشخصه‌های مهم رویدادهای بارش، توابع مذکور با مفصل چهاربعدی تجربی مقایسه گردیدند. در نهایت ساختار دی- واین چهار بعدی M-R-D-L با توجه به میزان معیارهای ارزیابی $R^2=0/991$ ، $RMSE=0/031$ و $MAE=0/024$ ، و روش گرافیکی بعنوان مناسب‌ترین ساختار جهت تعیین تابع توزیع توأم مشخصه‌های مهم بارش در ایستگاه سریمونا انتخاب گردید.

واژه‌های کلیدی: تحلیل فراوانی چندمتغیره، جفت-مفصل، مشخصه‌های بارش، ساختار دی- واین، تابع مفصل

^۱ دکتری مهندسی منابع آب، گروه مهندسی آب، دانشگاه تبریز، تبریز، ایران m.shafaei65@gmail.com
^۲ دانشیار گروه مهندسی آب، دانشگاه شهرکرد، شهرکرد، ایران. نویسنده مسئول: mirabbasi_r@yahoo.com

مقدمه

مهندسی منابع آب محدود به چند مطالعه انگشت شمار می‌باشد. در این خصوص می‌توان به مطالعات معدودی نظیر مطالعه دی میچله و همکاران (۲۰۰۷) در ارزیابی رویدادهای توفان دریا با استفاده از یک مفصل چهاربندی (De Michele et al., 2007)، مطالعه آیانتوبو و سانگ (۲۰۱۸) در تحلیل فراوانی چهاربندی مشخصه‌های خشکسالی با استفاده از توابع مفصل متقارن^۱ و نامتقارن^۲ چهاربندی، خانی تملیه و همکاران (۱۳۹۹) در تحلیل فراوانی مشخصات پدیده خشکسالی با بکارگیری مفصل تودرتو چهاربندی اشاره نمود. اما بطور کلی، استفاده از توابع مفصل در مسائل با ابعاد بالا سبب افزایش بار محاسباتی و پیچیدگی ریاضیاتی می‌شود و هر یک از مفصل‌های چند متغیره در مواجه با مسائل با ابعاد بالا با محدودیت‌هایی مواجه می‌شوند. بعنوان مثال در اینگونه موارد، بکارگیری مفصل‌های ارشمیدسی نامتقارن دارای مشکلاتی من جمله وابستگی‌های مثبت بین متغیرهای درونی و کوچکتر بودن وابستگی بیرونی نسبت به وابستگی درونی می‌باشد که این موضوع موجب محدودیت‌هایی در استفاده از این مفصل‌ها در مسائل با ابعاد بالاتر از دو می‌باشد. همچنین در مفصل‌های نیمه بیضوی^۳، با برآورد ریشه مربع ماتریس کوواریانس با تجزیه کولسکی^۴، چندک‌های متغیرهای مهم که در تفکیک بارش حائز اهمیت می‌باشد، حفظ نمی‌شوند. علاوه بر این در مسائلی که ماتریس همبستگی دارای اعداد منفی می‌باشد، تصحیحاتی جهت تضمین مثبت بودن ماتریس مورد نیاز است. ، اما ساختار مفصل واین^۵ با تجزیه توزیع‌های چندمتغیره به توزیع‌های توأم با بعد دو، انعطاف‌پذیری لازم در مسائل با ابعاد بزرگتر را فراهم آورده و فاقد ضعف‌های مفصل‌های مذکور می‌باشد. برای اولین بار جو ۱۹۹۶ ساختمان جفت-مفصل یک مفصل چندمتغیره را در قالب توابع توزیع ارائه داد (Joe, 1996). بیدفورد و کوک مدل گرافیکی با عنوان واین نرمال^۶ را برای توزیع‌های با ابعاد بالا ارائه دادند (Bedford and Cooke, 2001, 2002). واین نرمال دارای یک ساختار

استفاده از توابع مفصل یک روش انعطاف‌پذیر برای ایجاد تابع توزیع توأم مستقل از توزیع‌های حاشیه‌ای را امکان‌پذیر می‌سازد (Sklar, 1959). توابع مفصل برای مدل سازی ساختار وابستگی بین دو یا چند متغیر بکار می‌رود. تاکنون خانواده‌های گوناگونی از توابع مفصل توسعه یافته‌اند که می‌توانند برای توصیف ساختار وابستگی بین متغیرهای مختلف هیدرولوژیکی بکار روند. توابع مفصل، امکان استفاده همزمان از چند توزیع تک متغیره (با توزیع-های مختلف) را برای ساخت یک توزیع چند متغیره (که وابستگی بین متغیرها را نیز در نظر می‌گیرد) فراهم می‌آورند (Shiau, 2006). در سال‌های گذشته، توابع مفصل به‌طور گسترده‌ای در تحلیل فراوانی متغیرهای هیدرولوژیکی مورد استفاده قرار گرفته‌اند و کاربرد آنها در زمینه‌های مختلف هیدرولوژی مانند تحلیل فراوانی بارش، سیل، خشکسالی ... به‌سرعت رو به فزونی می‌باشد. بطوریکه در زمینه تحلیل فراوانی پدیده سیل می‌توان به مطالعات (De Michele et al., 2005; Grimaldi and Serinald, 2006; Zhang and Singh, 2006; Ganguli and Reddy, 2012; Kuchment and Demidov, 2013; Sraj et al., 2015; Dodangeh et al., 2020) تحلیل رویدادهای خشکسالی به مطالعات (Shiau, 2006; Kao and Govindaraju, 2010; Song and Sing, 2010; Mirabbasi et al., 2012; Mirabbasi et al., 2013; Salvadori and De Michele, 2015; Ghafari et al., 2019) و در زمینه تحلیل فراوانی بارش به (De Michele and Salvadori, 2003; Grimaldi and Serinald, 2006; Zhang and Singh 2006; Vandenberghe et al., 2011; Ghosh, 2010; Palynchuk and Guo, 2011; Ariff et al., 2012; Pandey et al., 2018; Buliah and Yie. 2020; Christopher Dzure et al., 2020) اشاره نمود. بررسی‌ها نشان می‌دهد که تاکنون اکثر تحقیقات انجام شده بر تحلیل فراوانی متغیرهای هیدرولوژیکی با استفاده از مفصل‌های دوبعدی و سه‌بعدی متمرکز بوده‌اند، بطوریکه کاربرد توابع مفصل چند بعدی (ابعاد بالاتر از سه) در

⁴ Cholesky

⁵ Vine Copula

⁶ Regular vine

¹ Symmetrical

² Asymmetric

³ Meta-elliptic



ساختار دی-واین چهاربعدی مشخصه‌های مهم رویدادهای بارش (شامل عمق کل بارش، ماکزیمم شدت بارش، مدت زمان دوره مرطوب و دوره خشک) تعیین و سپس توابع توزیع توأم چهاربعدی با توجه به ساختارهای دی-واین منتخب استخراج گردید.

مواد و روش‌ها

منطقه مورد مطالعه

در این مطالعه، آمار بارش ساعتی ایستگاه باران‌سنجی سریمونا^۳ واقع در شهر سریمونا در کشور ایتالیا به مدت ۲۶ سال (۲۰۲۰-۱۹۹۵) بدلیل ضرورت استفاده از یک آمار کامل و طولانی‌مدت بارش مورد استفاده و بررسی قرار گرفته است. شهر سریمونا در شمال ایتالیا و در ساحل چپ رودخانه پرو قرار دارد و در طول جغرافیایی ۱۰°۲' و عرض جغرافیایی ۴۵°۸' و در ارتفاع ۴۹ متری از سطح دریا قرار دارد. آب و هوای شهر سریمونا گرم و معتدل می‌باشد. میانگین دمای آن ۱۳/۹ درجه سانتی‌گراد و میانگین بارش آن ۹۸۹ میلی‌متر در سال می‌باشد. سریمونا در زمره مناطق پربارش قرار دارد و حتی در خشک‌ترین ماه‌ها دارای میانگین بارش بالایی می‌باشد. در این مطالعه، به‌منظور جداسازی رویدادهای بارش از معیار توقف بارش به مدت حداقل ۷ ساعت استفاده شد و بدین ترتیب ۳۶۶ رویداد بارش شناسائی شدند. لازم به ذکر است که دی میچله و سالوادوری (۲۰۰۳) برای استخراج رویدادهای بارش از معیار توقف ۷ ساعته استفاده نمودند. در این مطالعه، دوره مرطوب (L) به دوره‌ای اطلاق می‌شود که رویدادهای بارش دقیقاً در آن ساعات رخ داده‌اند و دوره خشک به دوره‌ای اطلاق می‌گردد که در یک رویداد بارش واقع شده است و مدت آن کمتر از ۷ ساعت باشد با توجه به مطالب بالا می‌توان نتیجه‌گیری کرد در صورتیکه فاصله بین دو رویداد بارش کمتر از ۷ ساعت باشد فاصله بین دو رویداد جزو مدت زمان دوره خشک بوده و دو رویداد متوالی بعنوان یک رویداد محسوب می‌شوند. در شکل (۱) روش بکار رفته در این مطالعه جهت تشخیص مشخصه‌های یک رویداد بارش

کلی و عمومی می‌باشد و شامل تجزیه‌های جفت-مفصل^۱ متعددی می‌باشد. بطور کلی، واین نرمال به دو نوع سی-واین^۲ و دی-واین^۳ تقسیم بندی می‌شوند که هرکدام از واین‌ها روش مشخصی را برای تجزیه تابع چگالی ارائه می‌دهند (Kurowicka and Cooke, 2006) مفصل واین دارای ویژگی سلسله مراتبی می‌باشد؛ عبارت دیگر، این مفصل با استفاده از آبشاری از مفصل‌های دو بعدی مفصل-های چند بعدی را محاسبه می‌نماید. در سالهای اخیر، توابع مفصل واین بعنوان ساختارهای درختی با انعطاف-پذیری بالا در علوم مختلف مورد توجه قرار گرفته‌اند. با این وجود، کاربرد مفصل واین در تعیین توابع توزیع متغیرهای هیدرولوژیکی چندبعدی و شبیه‌سازی آنها به مطالعات کمی محدود می‌باشد (Graler et al., 2013; Vernieuwe et al., 2015; Shafaei et al., 2016; Ni et al., 2020) بطوریکه گرالر و همکاران در تحلیل فراوانی سه بعدی مشخصات مهم سیل از ساختار دی-واین سه بعدی بهره گرفتند و با استفاده از توزیع احتمالاتی بدست آمده، دوره بازگشت‌های مختلف مشخصه‌های سیل را محاسبه نمودند (Graler et al., 2013). شفائی و همکاران (۱۳۹۵) با استفاده از ساختار سی-واین چهار بعدی و انتخاب مناسب-ترین خانواده‌های مفصل ارشمیدسی و بیضوی بر هر یک از جفت-مفصل‌های ساختارها، دقیق‌ترین توابع توزیع توأم چهاربعدی مشخصه‌های مهم سیل (یعنی زمان پایه، زمان تا اوج، حجم سیل و دبی اوج سیل) را محاسبه نمودند. نی و همکاران (۲۰۲۰) توابع توزیع احتمالاتی چهاربعدی مشخصه‌های مهم رویدادهای خشکسالی را با بکارگیری ساختارهای درختی سی-واین و دی-واین محاسبه نمودند (Ni et al., 2020). امینی و همکاران (۱۳۹۹) برای تحلیل فراوانی سیل در حوضه آبخیز بازفت استان چهارمحال و بختیاری از مفصل‌های سی-واین و دی-واین سه‌متغیره و چهارمتغیره استفاده نمودند. در این مطالعه برای اولین بار تابع توزیع توأم چهاربعدی مشخصه‌های مهم بارش ایستگاه باران‌سنجی سریمونا واقع در کشور ایتالیا با استفاده از ساختار دی-واین چهاربعدی استخراج گردید. بدین منظور، ابتدا با استفاده از معیارهای ارزیابی مختلف، دقیق‌ترین

³ Drawable vine (D-vine)

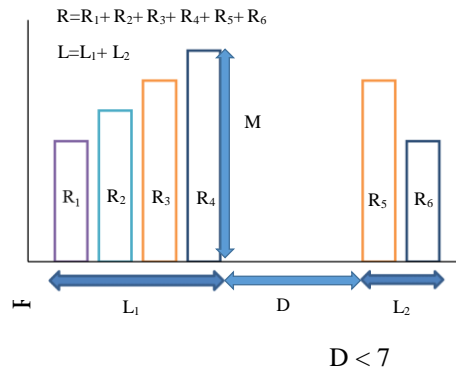
⁴ Cremona

¹ Pair-copula

² Canonical vine (C-vine)

استخراج شدند که مشخصات آماری مشخصه‌های مهم رویدادهای بارش در جدول (۱) ارائه شده اند.

بصورت شماتیک نشان داده شده است. مشخصه‌های مهم این رویدادها یعنی عمق کل بارش (R) ماکزیمم شدت بارش (M) مدت زمان دوره مرطوب (L) و دوره خشک (D)



شکل (۱): نحوه تعریف مشخصه‌های بارش در یک رویداد بارش (hr) L

جدول (۱): مشخصات آماری رویدادهای بارش مورد استفاده در مطالعه حاضر (ایستگاه سریمونا، ایتالیا)

عمق کل بارش R(mm)	بیشینه بارش M (mm/hr)	مدت زمان دوره مرطوب L (hr)	مدت زمان دوره خشک D (hr)	آماره‌های مشخصات رویدادهای بارش
۱۶۱/۴	۳۵/۰	۶۳/۰	۲۸/۰	بیشینه
۰/۸	۰/۲	۳/۰	۰	کمینه
۱۹/۹	۵/۰	۹/۷	۴/۸	انحراف معیار
۲/۷	۲/۹	۱/۴	۱/۵	چولگی
۱۹/۵	۴/۲	۱۴/۹	۴/۹	میانگین

درخت‌ها با حداکثر N متغیر و حداکثر N-1 لبه^۱ می‌باشد. هر لبه نشاندهنده یک مفصل یعنی توزیعی دو بعدی در فضای [۰, ۱]^۲ می‌باشد. در هر درخت، متغیرها با استفاده از مفصل‌هایی که بعنوان لبه شناخته می‌شوند توزیع توأمی را تشکیل می‌دهند که همان جفت-مفصل می‌باشد (Czado, 2010). شایان ذکر است که جفت-مفصل‌ها در ساختار و این همان مفصل‌های دوبعدی می‌باشند. علاوه بر این نوع و ترتیب اتصال متغیرها و جفت-مفصل‌ها در هر درخت و این، مهمترین مسئله در تعیین نوع ساختار و این می‌باشد. به نظر می‌رسد، مدل‌های گرافیکی (درخت‌های

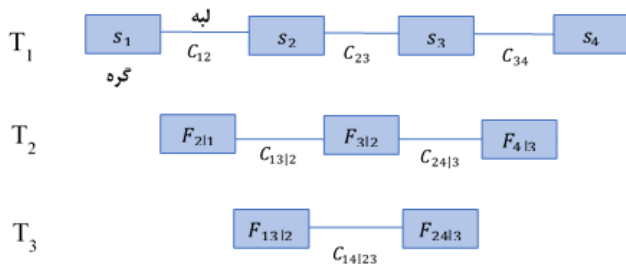
مفصل و این

با توجه به محدودیتهای مفصل‌های چند بعدی (ابعاد بالاتر از دو) در تعیین توابع توزیع توأم چند بعدی، برای اولین بار جو در سال ۱۹۹۶، مفصل‌های چند متغیره‌ای را که ساختار آنها شامل بلوک‌های مفصل دو متغیره می‌شد را معرفی نمود (Joe, 1996). این بلوک‌ها عمدتاً جفت-مفصل نامیده می‌شوند. مجموعه‌ای از درخت‌های متصل به هم به منظور اتصال بین جفت-مفصل‌ها مورد استفاده قرار گرفت که و این نرمال نامیده می‌شوند (Bedford and Cooke, 2001, 2002). مدل‌های و این شامل دنباله‌ای از

¹ Edge



متصل بهم درخت T_i ایجاد می‌شوند (i شماره درخت می‌باشد) و شرط اتصال دو لبه در درخت T_{i+1} اشتراک دو لبه مورد نظر در یک گره در درخت T_i می‌باشد (Aas et al., 2009). همانگونه که در شکل (۲) مشاهده می‌شود، وجه تمایز درخت اول T_1 با درخت‌های T_2 و T_3 در غیر شرطی بودن گره‌های درخت اول می‌باشد، در حالیکه درخت‌های دوم و سوم شامل گره‌های شرطی می‌باشند. در ساختار گرافیکی مدل دی-واین، متغیرها بصورت خطی (پشت سرهم) قرار می‌گیرند و هر متغیر حداکثر به دو متغیر متصل خواهد می‌شود. مطابق با شکل (۲)، درخت اول T_1 از چهار متغیر (گره) s_1, s_2, s_3, s_4 و لبه‌های غیرشرطی C_{12}, C_{23}, C_{34} تشکیل شده است. اما همانطور که در شکل (۲) نمایش داده شده است، گره‌های درخت-های T_2 شامل توابع توزیع شرطی $F_{2|1}, F_{3|2}, F_{4|3}$ و گره‌های درخت T_3 شامل توابع توزیع شرطی $F_{13|2}, F_{24|3}$ می‌باشند که به ترتیب حاصل مشتق جزئی مفصل-های (لبه‌ها) درخت اول یعنی C_{12}, C_{23}, C_{34} و لبه‌های درخت دوم یعنی $C_{13|2}$ و $C_{24|3}$ می‌باشند. لازم بذکر است که روش محاسبه جفت-مفصل‌ها در قسمت بعدی توضیح داده خواهد شد.



شکل (۲): درخت‌های ساختار دی-واین چهار بعدی

توابع مفصل

اگر F توزیع توأم چند متغیره متغیرهای تصادفی s_1, s_2, \dots, s_N با توزیع‌های حاشیه‌ای $F_{S_1}(s_1), F_{S_2}(s_2), \dots, F_{S_N}(s_N)$ که $F_{S_i}(s_i) = u_i$ باشد، بنابراین، مفصل C در دامنه R برای متغیرهای s_1, s_2, \dots, s_N بصورت زیر تعیین می‌شود:

متصل بهم) مدل‌های مناسبی برای ارائه توزیع‌های احتمالاتی توأم با ابعاد بالاتر از دو باشند، بطوریکه این مدل‌ها، بطور تصویری به توصیف وابستگی‌های موجود در توزیع‌های توأم می‌پردازند.

یکی از ویژگی‌های مهم مدل‌های گرافیکی، توانایی آنها در توصیف ساختارهای پیچیده به شکل مدولار^۱ می‌باشد، بطوریکه در مدل‌های مذکور، وابستگی بین عناصر همسایه ترکیب می‌شوند. واین نرمال به دو نوع واین به نام‌های سی-واین (C-vine) و دی-واین (D-vine) تقسیم‌بندی می‌شوند که ساختار دی-واین شامل درخت‌های با ساختار خطی و ساختار سی-واین شامل درخت‌های با ساختار ستاره‌ای می‌باشند، بطوریکه هر یک از درخت‌های ساختار سی-واین دارای یک متغیر مرکزی می‌باشند. بدفورد و کوک تجزیه‌های جفت-مفصل را بطور گرافیکی و بصورت دنباله‌ای از درخت‌های تو در تو (درخت واین) با لبه‌های بدون جهت نمایش دادند (Bedford and Cooke, 2001, 2002). مطابق با مطالعه کورویکا و کوک، درخت واین نرمال d متغیره شامل درخت‌های متصل شده $T_i = \{T_1, \dots, T_{d-1}\}$ که $i=1, \dots, d-1$ با گره‌های N_i و لبه‌های E_i می‌باشد و شرایط زیر را ارضاء می‌نماید (Kurowicka and Cooke, 2006)

۱- T_1 دارای گره‌های $N_1 = \{1, \dots, d\}$ و E_1 می‌باشد

۲- برای $i=2, \dots, d-1$ درخت T_i دارای گره‌های N_i

E_{i-1} می‌باشد.

۳- دو لبه متعلق به درخت T_i در صورت اشتراک آنها در یکی از گره‌های درخت T_i ، در درخت T_{i+1} به یکدیگر متصل می‌گردند

شیوه تجزیه تابع چگالی توزیع چهار بعدی توسط ساختار دی-واین چهار بعدی در شکل (۲) نشان داده شده است. در شکل (۲) هر مربع نشانگر یک گره و هر خط بین دو مربع نمایانگر یک لبه می‌باشد. هر گره مطابق با یک متغیر و هر لبه با یک چگالی جفت-مفصل متناظر می‌باشد. اگر d تعداد متغیرها باشد، کل تجزیه‌ها بواسطه $(d-1)/2$ لبه یا همان چگالی جفت-مفصل انجام می‌گیرد. برچسب‌های لبه‌های درخت T_{i+1} از اتصال گره‌های

² Node

¹ Modular

(Bedford and Cooke, 2001, 2002). همانطور که رابطه فوق نشان می‌دهد، روابط توابع مفصل دومتغیره شامل توابع توزیع حاشیه‌ای شرطی و تک متغیره نیز می‌شوند که توابع توزیع تک متغیره با استفاده از انواع توابع توزیع احتمالاتی قابل محاسبه می‌باشد، اما توابع توزیع حاشیه‌ای شرطی با استفاده از مشتق جزئی تابع مفصل دوبعدی قابل محاسبه می‌باشند. در این مطالعه به منظور تعیین توابع حاشیه‌ای شرطی موجود در رابطه زنجیره‌ای چگالی چهار بعدی دی- و این فوق از روابط زیر استفاده شده است (Joe, 1996).

$$h_{2|1} = F(s_2|s_1) = \frac{\partial C_{12}(F(s_1), F(s_2), \theta_{12})}{\partial F(s_1)} \quad (5)$$

$$h_{3|2} = F(s_3|s_2) = \frac{\partial C_{23}(F(s_3), F(s_2), \theta_{32})}{\partial F(s_2)} \quad (6)$$

$$h_{4|3} = F(s_4|s_3) = \frac{\partial C_{34}(F(s_3), F(s_4), \theta_{34})}{\partial F(s_3)} \quad (7)$$

$$h_{1|23} = F(s_1|s_2, s_3) = \frac{\delta C_{13|2}(F(s_1|s_2), F(s_3|s_2); \theta_{13|2})}{\delta F(s_3|s_2)} \quad (8)$$

$$h_{4|23} = F(s_4|s_2, s_3) = \frac{\delta C_{42|3}(F(s_4|s_3), F(s_2|s_3); \theta_{42|3})}{\delta F(s_2|s_3)} \quad (9)$$

با توجه به روابط بالا، می‌توان نتیجه گرفت که بطور کلی توزیع جمعیتی شرطی بصورت $h(u_i, u_j, \theta_{ij}) = F(s_j|s_i)$ تعریف می‌گردد. بطوریکه در این رابطه برای تعیین تابع h ، از مفصل دوبعدی C نسبت به متغیر دوم مشتق جزئی گرفته میشود. $\theta_{ij|k}$ پارامتر تابع توزیع مفصل توأم بین دو متغیر $F(s_i|s_k)$ و $F(s_j|s_k)$ می‌باشد (Aas et al., 2009).

محاسبه پارامترهای ساختار دی- و این چهاربعدی

در این مطالعه، از کاربرد روش ترتیبی جهت تخمین پارامترهای ساختار دی- و این چهار بعدی استفاده شده است که مراحل آن به تفکیک در ادامه ارائه شده است: ۱- برآزش و انتخاب یک خانواده مفصل مناسب دوبعدی بر u_1

$$F_{S_1, \dots, S_N}(s_1, s_2, \dots, s_N) = C(u_1, u_2, \dots, u_n, \theta) = C(F_{S_1}(s_1), F_{S_2}(s_2), \dots, F_{S_N}(s_N), \theta) \quad (1)$$

$$C(u_1, u_2, \dots, u_n, \theta) = P(S_1 \leq s_1, S_2 \leq s_2, \dots, S_n \leq s_n) \quad (2)$$

تابع C یک تابع مفصل منحصر بفرد n بعدی می‌باشد که توزیع‌های یکنواخت حاشیه‌ای تک متغیره را به توزیع جمعیتی n بعدی توأم نگاشت می‌نماید. بطوریکه C نگاشت $[0, 1]^n \rightarrow [0, 1]$ می‌باشد. θ بر پارامتر تابع مفصل دلالت دارد که مقدار وابستگی بین توزیع‌های حاشیه‌ای را نشان می‌دهد. تئوری معکوس نظریه اسکالر بصورت زیر تعریف می‌شود:

$$C(u_1, u_2, \dots, u_n, \theta) = F_{S_1, \dots, S_N}(F_{S_1}^{-1}(u_1), F_{S_2}^{-1}(u_2), \dots, F_{S_N}^{-1}(u_n)) \quad (3)$$

u_n در رابطه بالا توابع توزیع حاشیه‌ای متغیر n ام و $F_n^{-1}(\cdot)$ تابع چندک تابع F می‌باشد. خانواده‌های توابع مفصلی مختلفی نظیر توابع مفصل ارشمیدسی، بیضوی، مقدر حدی و ... برای تعیین ساختمان همبستگی توابع موجود هستند که با انتخاب مناسب خانواده مفصل می‌توان ساختار وابستگی چندمتغیره را به خوبی مدل‌سازی نمود (Aas et al., 2009).

تابع چگالی توأم چهاربعدی دی- و این

مطابق با قانون زنجیره‌ای، تابع چگالی توأم چهاربعدی دی- و این بصورت رابطه (۴) تعریف می‌گردد.

$$f(s_1, s_2, s_3, s_4) = f(s_1) \cdot f(s_2) \cdot f(s_3) \cdot f(s_4) \cdot c_{12}\{F(s_1), F(s_2), \theta_{12}\} \cdot c_{23}\{F(s_2), F(s_3), \theta_{23}\} \cdot c_{34}\{F(s_3), F(s_4), \theta_{34}\} \cdot c_{13|2}\{F(s_2|s_1), F((s_3|s_2), \theta_{13|2})\} \cdot c_{24|3}\{F(s_3|s_2), F((s_4|s_3), \theta_{24|3})\} \cdot c_{14|23}\{F(s_1|s_2, s_3), F((s_4|s_2, s_3), \theta_{14|23})\} \quad (4)$$

در رابطه بالا، C تابع چگالی مفصل دوبعدی می‌باشد. این نوع تجزیه تابع چگالی چند بعدی به توابع چگالی تک متغیره و مفصل‌های دوبعدی، برای اولین بار توسط بدفورد و کوک شرح داده شد که تجزیه دی- و این نامیده می‌شود



در رابطه (10) ، C_{ij} تابع چگالی مفصل دوبعدی و $h_{j,t}$ و $h_{i,t}$ توزیع‌های شرطی تولید شده با استفاده از روش ترتیبی فوق و تابع h در زمان t و یا توزیع‌های حاشیه‌ای تک متغیره می‌باشند.

مراحل دستیابی به ساختار دی-واین منتخب جهت تعیین تابع توزیع توأم چهاربعدی

در این مطالعه، به منظور انتخاب بهترین ساختار دی-واین مشخصه‌های مهم رویدادهای بارش یعنی (ماکزیمم شدت بارش، عمق کل، مدت زمان دوره خشک و مرطوب رویدادهای بارش) مرحله زیر اجرا گردید: (۱) محاسبه ضرائب تاو-کندال بین همه جفت-متغیرها جهت اتصال متغیرهای وابسته در مفصل دی-واین و همچنین حذف جفت-متغیرهای غیروابسته در صورت لزوم. ضریب تاو-کندال از تفریق احتمال انطباق و احتمال عدم انطباق حاصل می‌شود. اگر فرض کنیم $(x_1, y_1), \dots, (x_n, y_n)$ جفت متغیر تصادفی باشند، (x_i, y_i) و (x_j, y_j) بر هم منطبقند اگر، $[(x_j - x_i)(y_j - y_i)] > 0$ و در غیر اینصورت غیر منطبقند. اگر a و b به ترتیب جفت‌های منطبق و غیرمنطبق باشند، ضریب تاو-کندال بصورت $\tau = (a - b) / \binom{n}{2}$ (Kendall, 1938) تعریف می‌گردد. (۲) تعیین توزیع‌های حاشیه‌ای مربوط به هر یک از مشخصات مهم رویدادهای بارش با آزمون توزیع‌های پارامتری مختلف بر متغیرها و انتخاب دقیق‌ترین آنها با توجه به معیارهای ارزیابی مختلف (۳) جایگشت توزیع‌های حاشیه‌ای متغیرها (مشخصات مهم رویدادهای بارش) در ساختار مفصل دی-واین به منظور بدست آوردن ساختارهای واین مختلف. بطوریکه در هر جایگشت، جابجایی متغیرها در ساختارهای دی-واین منجر به ایجاد ساختارهای مختلف می‌شود. (۴) استفاده از خانواده‌های ارشمیدسی و بیضوی مختلف جهت برآزش بر جفت-مفصل‌های موجود در ساختارهای ایجاد شده دی-واین حاصل از مرحله قبل، بعبارت دیگر، بر هر یک از جفت-مفصل‌های موجود در درخت‌های ساختار واین، همه‌ی خانواده‌های مفصل ارشمیدسی و بیضوی مورد آزمون قرار می‌گیرند تا بهترین خانواده مفصل جهت برآزش بر جفت-مفصل‌های مختلف انتخاب شوند. (۵) تخمین پارامترهای جفت-مفصل‌های خانواده‌های مختلف مطابق با

u_2 جهت تخمین پارامتر مفصل θ_{12} . ۲- برآزش و انتخاب یک خانواده مفصل مناسب دوبعدی بر u_2 و u_3 جهت تخمین پارامتر مفصل θ_{23} . ۳- برآزش و انتخاب یک خانواده مفصل مناسب دوبعدی بر u_3 و u_4 جهت تخمین پارامتر مفصل θ_{34} . ۴- محاسبه سری $h_{2|1}$ با توجه به پارامتر مفصل تخمینی و رابطه (۵) $h_{2|1} = h(u_1, u_2, \theta_{12})$. ۵- محاسبه سری h_{23} با توجه به پارامتر مفصل تخمینی و رابطه (۶) $h_{3|2} = h(u_2, u_3, \theta_{23})$. ۶- محاسبه سری h_{34} با توجه به پارامتر مفصل تخمینی و رابطه (۷) $h_{4|3} = h(u_3, u_4, \theta_{34})$. ۷- برآزش و انتخاب یک خانواده مفصل مناسب دوبعدی بر $h_{2|1}$ و $h_{3|2}$ جهت تخمین پارامتر مفصل $\theta_{31|2}$. ۸- برآزش و انتخاب یک خانواده مفصل مناسب دوبعدی بر $h_{3|2}$ و $h_{4|3}$ جهت تخمین پارامتر مفصل $\theta_{24|3}$. ۹- محاسبه سری $h_{1|23}$ با قرار دادن پارامتر مفصل تخمینی در رابطه (۸) $h_{1|23} = h(h_{2|1}, h_{3|2}, \theta_{31|2})$. ۱۰- محاسبه سری $h_{4|23}$ با قرار دادن پارامتر مفصل تخمینی در رابطه (۹) $h_{4|23} = h(h_{3|2}, h_{4|3}, \theta_{24|3})$. ۱۱- برآزش و انتخاب یک خانواده مفصل مناسب دوبعدی بر $h_{1|23}$ و $h_{4|23}$ جهت تخمین پارامتر مفصل $\theta_{14|23}$ شایان ذکر است که در این مطالعه، برای انتخاب و تخمین پارامترهای مفصل دوبعدی θ_{ij} از روش حداکثر لوگ درست‌نمایی بهره گرفته شده است. بطوریکه ابتدا هریک از مفصل‌های خانواده‌های ارشمیدسی (کلایتون، گامبل، جو و فرانک) و بیضوی (تی-استیودنت و گوسین) بر هر یک از جفت-مفصل‌های C_{ij} و G_{ijkl} ساختار دی-واین برآزش داده شدند. سپس پارامترهای تخمین زده شده θ_{ij} و θ_{ijkl} که حاصل از برآزش هر کدام از خانواده‌های مفصل‌های مختلف بر جفت-مفصل‌ها هستند، در رابطه (۱۰) قرار داده می‌شوند. سپس هر خانواده مفصلی که موجب حداکثر شدن میزان لوگ درست‌نمایی در رابطه ذیل گردد، بعنوان خانواده منتخب جهت برآزش بر هر جفت-مفصل موجود در ساختار دی-واین تعیین می‌گردد. بدین ترتیب در نهایت برای هر یک ساختار دی-واین ترکیبی از مناسب‌ترین خانواده‌های مفصل حاصل می‌گردد که موجب حداکثر شدن لوگ درست‌نمایی ساختار دی-واین می‌گردد.

$$L(\theta) = \sum_{t=1}^T \{\log c_{i,j}(h_{i,t}, h_{j,t}; \theta)\} \quad (10)$$

صفر است. n ، تعداد داده‌ها، R_i, M_i, L_i و D_i رتبه هریک از نمونه‌ها در مشخصات مهم رویدادهای بارش می‌باشند (Genest and Rivest, 1993). در نهایت سه معیار ارزیابی ریشه میانگین مربعات خطا^۵ (RMSE)، میانگین خطای مطلق^۶ (MAE) و ضریب تبیین^۷ (R^2) جهت ارزیابی میزان تطابق تابع توزیع توأم چهاربعدی دی-واین تعیین شده بامفصل تجربی چهاربعدی استفاده شدند. معیارهای ارزیابی مذکور مطابق با روابط (۱۵) تا (۱۷) محاسبه می‌شوند:

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (C_{ti} - C_{ei})^2}{n}} \quad (15)$$

$$MAE = \frac{\sum_{i=1}^n |C_{ti} - C_{ei}|}{n} \quad (16)$$

$$R^2 = \left(1 - \frac{\sum_{i=1}^n (C_{ti} - C_{ei})^2}{\sum_{i=1}^n C_{ti}^2 - \frac{\sum_{i=1}^n C_{ei}^2}{n}}\right) \quad (17)$$

در روابط بالا، C_t و C_e به ترتیب توزیع توأم دی-واین چهاربعدی محاسبه شده و مفصل تجربی چهاربعدی می‌باشند. همچنین n تعداد داده‌ها و i شماره داده‌ها می‌باشند. شایان ذکر است که در این مطالعه کلیه محاسبات مفصل واین با استفاده از کد نویسی در محیط برنامه‌نویسی R انجام شده است.

نتایج و بحث

تعیین ساختارهای دی-واین مشخصات مهم بارش
در این مطالعه، به منظور تعیین توزیع توأم چهاربعدی دی-واین مشخصات رویدادهای بارش، داده‌های بارش ساعتی ایستگاه سریمونا واقع در کشور ایتالیا از سال‌های (۱۹۹۵-۲۰۲۰) با توجه به داده‌های کامل و طولانی‌مدت آن مورد استفاده قرار گرفت و همانگونه که پیش‌تر نیز گفته شد با استفاده از معیار توقف بارش ۷ ساعته، در نهایت ۳۶۶ رویداد بارش استخراج و سپس مشخصات مهم رویدادهای

روشی که در قسمت قبل شرح داده شد. (۶) انتخاب مناسب‌ترین خانواده مفصل برای هر جفت-مفصل که در نهایت، بیشترین مقدار تابع لوگ درست نمائی و کمترین مقدار معیارهای اطلاعات آکائیکه^۱ (AIC) و اطلاعات بیزین^۲ (BIC) را ارائه می‌دهند. نحوه محاسبه معیارهای مذکور در روابط ۱۱-۱۳ برای ساختار دی-واین چهار بعدی شرح داده شده است (Brechmann et al., 2012).

$$L(u, h, \theta) = \sum_{t=1}^T \{ \log c_{12}(u_{1,t}, u_{2,t}; \theta_{12}) + \log c_{23}(u_{2,t}, u_{3,t}; \theta_{23}) + \log c_{34}(u_{3,t}, u_{4,t}; \theta_{34}) + \log c_{31|2}(h_{2|1,t}, h_{3|2,t}; \theta_{13|2}) + \log c_{24|3}(h_{3|2,t}, h_{4|3,t}; \theta_{24|3}) + \log c_{14|23}(h_{13|2,t}, h_{24|3,t}; \theta_{14|23}) \} \quad (11)$$

$$AIC = -2L(u, h, \theta) + 2 \quad (12)$$

$$BIC = -2L(u, h, \theta) + 2k * \ln(n) \quad (13)$$

در روابط فوق، $L(u, h, \theta)$ ، لوگ درست نمائی و k و n به ترتیب تعداد پارامترها و تعداد داده‌های می‌باشند. c_{ij} نیز بر چگالی مفصل دویبعدی بین i و j دلالت دارد.

ارزیابی دقت ساختار توزیع توأم چهار بعدی دی-واین

در این مطالعه به منظور سنجش میزان دقت تابع توزیع توأم چهاربعدی تعیین شده با مفصل تجربی^۳ چهاربعدی مشخصات مهم رویدادهای بارش مقایسه گردید. مفصل تجربی مشخصات مهم رویدادهای بارش (شامل ماکزیمم شدت بارش (M)، عمق کل بارش (R)، مدت زمان دوره مرطوب (L) و دوره خشک (D)) توسط رابطه (۱۴) تعیین می‌گردد:

$$C_e(r, m, l, d) = \frac{1}{n} \sum_{i=1}^n I\left(\frac{R_i}{n+1} \leq r, \frac{M_i}{n+1} \leq m, \frac{L_i}{n+1} \leq l, \frac{D_i}{n+1} \leq d\right) \quad (14)$$

در رابطه بالا I تابع شاخص^۴ می‌باشد، که اگر عبارت داخل پرانتز درست باشد مقدار آن برابر ۱ و در غیر اینصورت برابر

⁵ Root Mean Square Error

⁶ Mean Absolute Error

⁷ Coefficient of Determination

¹ Akaike information criterion

² Bayesian information criterion

³ Empirical Copula

⁴ Indicator function

اسمیرنوف و کای اسکور ارائه شده‌اند. مطابق با جدول ۲ و با توجه به آزمون‌های کولموگروف-اسمیرنوف و کای اسکور، توزیع های بور^۱ و لوگ لوجستیک^۲ به ترتیب جهت مدلسازی متغیرهای عمق کل بارش و حداکثر بارش و توزیع پارتو تعمیم یافته^۳ جهت برآزش بر متغیرهای مدت زمان دوره‌های مرطوب و دوره خشک در سطح معنی‌داری ۵ درصد انتخاب شدند. پس از جایگذاری پارامترهای محاسبه شده در توابع توزیع منتخب، توزیع‌های حاشیه‌ای مربوط به هر یک از مشخصه‌های مهم رویدادهای بارش حاصل گردید

بارش شامل عمق کل بارش (R)، ماکزیمم شدت بارش (M)، مدت زمان دوره مرطوب (L) و مدت زمان دوره خشک (D) تعیین گردیدند. پس از استخراج مشخصه‌های مهم رویدادهای بارش، گام بعدی تعیین توزیع‌های احتمالاتی مناسب جهت برآزش بر هر یک از چهار مشخصه مهم رویدادهای بارش و سپس تخمین پارامترهای توزیع‌های منتخب می‌باشد. بنابراین در این مطالعه توزیع‌های پارامتری مختلفی جهت مدل‌سازی مشخصه‌های مهم بارش مورد آزمون قرار گرفتند. در جدول ۲ فرمول‌های توزیع‌های منتخب به همراه میزان پارامترهای تخمین زده شده مربوط به توابع توزیع منتخب و P-value آزمون‌های کولموگروف-

جدول (۲): پارامترهای توابع توزیع حاشیه‌ای به همراه P-value مقدار آزمون‌های کولموگروف-اسمیرنوف و کای اسکور

توابع توزیع-مشخصه رویداد بارش	فرمول	α	β	γ	k	μ	σ	P_{KS}	P_{χ^2}
بور-عمق کل بارش	$F(x) = 1 - (1 + (\frac{x-\gamma}{\beta})^\alpha)^{-k}$	۱/۲	۵۶/۴	۰/۸	۴/۳	-	-	۰/۶	۰/۹
لوگ لوجستیک-حداکثر شدت بارش	$F(x) = (1 + (\frac{\beta}{x-\gamma})^\alpha)^{-1}$	۱/۶	۲/۴	۰/۲	-	-	-	۰/۵	۰/۲
پارتو تعمیم یافته-زمان دوره مرطوب	$F(x) = 1 - (1 + k(\frac{x-\mu}{\sigma})^{-1/k})^{-1/k}$	-	-	-	-۰/۲	۳/۶	۱۳/۰	۰/۳	۰/۷
پارتو تعمیم یافته-زمان دوره خشک	$F(x) = 1 - (1 + k(\frac{x-\mu}{\sigma})^{-1/k})^{-1/k}$	-	-	-	-۰/۱	-۰/۵	۶/۳	۰/۳	۰/۶

در جدول ۳ ارائه شده است. نتایج جدول ۳ نشان می‌دهد که میزان P-value کلیه جفت-متغیرها به استثناء جفت متغیر M-D کمتر از ۰/۰۵ می‌باشد. بنابراین در بین جفت-متغیرهای مشخصات مهم رویدادهای بارش، همبستگی بین متغیرهای M و D در سطح ۵ درصد معنی‌دار نمی‌باشد، در حالیکه در سایر جفت-متغیرها همبستگی معنی‌داری بین دو متغیر برقرار می‌باشد

پس از تعیین توزیع‌های حاشیه‌ای مربوط به هر یک از مشخصه‌های بارش، میزان وابستگی یا عدم وابستگی بین هر جفت از مشخصه‌های مهم بارش مورد ارزیابی قرار گرفت. در این مطالعه برای سنجش میزان وابستگی بین هر یک از جفت-متغیرها از ضریب وابستگی تاو-کندال استفاده شد. میزان ضریب تاو-کندال بین جفت-متغیرهای مشخصه‌های بارش به همراه میزان P-value آنها

جدول (۳): ضریب تاو-کندال بین جفت-متغیرهای مشخصه‌های بارش به همراه P-value آنها

L-D	R-D	R-L	M-D	M-L	M-R	جفت-متغیرها ضریب تاو-کندال (P-value)
۰/۲۶ (۰)	۰/۱۳ (۰)	۰/۴۳ (۰)	۰/۰۲ (۰/۵)	۰/۱۱ (۰)	۰/۶۰ (۰)	

³ Generalized Preto

¹ Burr

² Log-Logistic

خانواده‌های مفصل بهینه جهت برازش بر جفت-مفصل‌های یک ساختار دی-واین انتخاب می‌شوند. در جدول ۴ پارامترهای محاسبه شده خانواده‌های مفصل منتخب جهت برازش بر جفت-مفصل‌های هر یک از ۶ ساختار دی-واین به‌مراه معیارهای لوگ درست‌نمائی، آکائیکه و بی‌زین و همچنین میزان ضرائب تاو-کندال بین همه توزیع‌های حاشیه‌ای (u_1, u_2) ، (u_2, u_3) و (u_3, u_4) در درخت اول و همچنین بین سری‌های شرطی $(h_{2|1}, h_{3|2})$ و $(h_{3|2}, h_{4|3})$ در درخت دوم و بین سری‌های شرطی $(h_{4|23}, h_{1|23})$ در درخت سوم ساختار دی-واین ارائه شده‌اند. بعنوان مثال در ساختار R-M-L-D، θ_{23} ، θ_{12} و θ_{34} به ترتیب پارامترهای مربوط به جفت-مفصل‌های C_{RM} ، C_{ML} و C_{LD} در درخت اول، $\theta_{13|2}$ ، $\theta_{24|3}$ به ترتیب پارامترهای جفت-مفصل‌های $C_{MD|L}$ و $C_{RL|M}$ در درخت دوم و پارامتر $\theta_{14|23}$ مربوط به جفت-مفصل $C_{RD|LM}$ در درخت سوم می‌باشند. همانطور که جدول ۴ نشان می‌دهد، به ترتیب خانواده‌های مفصل گامبل، فرانک و گوسین جهت برازش بر جفت-مفصل‌های C_{RM} ، C_{ML} و C_{LD} و خانواده‌های گامبل و فرانک جهت برازش بر جفت-مفصل‌های $C_{MD|L}$ و $C_{RL|M}$ و در نهایت مفصل فرانک جهت برازش بر جفت-مفصل $C_{RD|LM}$ انتخاب شده است. در این مطالعه برای اولین بار میزان همبستگی بین کلیه جفت-متغیرهای توزیع‌های حاشیه‌ای در درخت اول و همچنین توزیع‌های شرطی تولید شده در درخت‌های دوم و سوم محاسبه شده است تا تأثیر کاربرد جفت-متغیرهای با میزان همبستگی‌های متفاوت را در دستیابی به ساختار بهینه دی-واین بررسی شود. علاوه بر این در این مطالعه سعی شده است تا اثر استفاده از جفت-متغیرهای با قوی‌ترین میزان همبستگی در درخت اول بر میزان همبستگی جفت-متغیرهای مربوط به درخت‌های دوم و سوم و همچنین تأثیر آن در دقت ساختار دی-واین بررسی گردد. لازم بذکر است که برای آزمون نکوئی برازش پارامترهای محاسبه شده از تست

یکی از اهداف مهم این مطالعه ایجاد ساختارهای مختلف دی-واین چهار متغیره با استفاده از جایگشت‌های مختلف مشخصه‌های مهم رویدادهای بارش (شامل عمق کل بارش (R)، ماکزیمم شدت بارش (M)، مدت زمان دوره مرطوب (L) و مدت زمان دوره خشک (D)) می‌باشد. مطابق با رابطه قانون جایگشت‌ها، n متغیر می‌تواند n! تعداد ترکیب با جایگشت‌های متفاوت ایجاد نمایند. از آنجا که از تعداد n! ترکیب ساختارهای دی-واین، نصف تعداد ساختارهای آنها دارای جفت-مفصل‌های یکسان در درخت اول ساختار خود هستند و تنها در ترتیب قرار گرفتن متفاوت می‌باشند، لذا ساختار خطی دی-واین با n متغیر می‌تواند $\frac{n!}{2}$ ساختار دی-واین متفاوت ایجاد نمایند بطوریکه مطابق با رابطه مزبور، ۴ متغیر می‌تواند $\frac{4!}{2}$ ساختار تولید یا عبارتی ۱۲ ساختار دی-واین متفاوت تشکیل دهند. اما با توجه به اینکه جفت-متغیر M-D دارای همبستگی معنی‌داری نمی‌باشد، لذا ساختارهایی که شامل این جفت متغیر می‌باشند از بین ۱۲ ساختار دی-واین حذف می‌گردند و به این ترتیب در این مطالعه ۶ ساختار دی-واین حذف و ادامه مطالعه با ۶ ساختار دی-واین باقی مانده انجام گرفت. در گام بعدی هدف تعیین مناسب‌ترین خانواده مفصل برای هر کدام از جفت-مفصل‌های موجود در درخت‌های واین ساختارهای دی-واین ایجاد شده، می‌باشد. لذا، مطابق با آنچه در قسمت قبل شرح داده شد (تخمین پارامترهای مفصل دی-واین با استفاده از روش ترتیبی)، خانواده‌های مختلف مفصل ارشمیدسی (کلایتون، فرانک، جو و گامبل) و بیضوی (تی-استیودنت و گوسین) بر هر یک از جفت-مفصل‌های موجود در درخت‌های ساختارهای دی-واین برازش داده می‌شوند. پس از تخمین پارامترهای مفصل با استفاده از روش ترتیبی و جایگذاری پارامترها در روابط (۱۱)، (۱۲) و (۱۳)، به ترتیب معیارهای لوگ درست‌نمائی، معیارهای آکائیکه و بی‌زین محاسبه می‌گردند و در نهایت ترکیبی از خانواده‌های مفصل مختلف که منجر به ماکزیمم شدن لوگ درست‌نمائی و مینیمم شدن معیارهای آکائیکه و بی‌زین در ساختار دی-واین می‌گردند، بعنوان ترکیبی از



مجموع مقادیر ضرائب تاو-کندال در درخت اول ساختار M-R-L-D برابر با ۱/۲۹ می‌باشد بطوریکه دارای بیشترین میزان در بین مجموع مقادیر تاو-کندال جفت-متغیرهای درخت‌های اول سایر ساختارها می‌باشد اما لزوماً دارای بیشترین مقادیر لوگ درست‌نمایی و یا کمترین مقادیر AIC و BIC در مقایسه با سایر با سایر ساختارهای دی-واین نمی‌باشد.

شفائی و همکاران (۲۰۱۷) نیز در مطالعه‌ای با استفاده از جایگشت متغیرها در ساختار دی-واین، ساختارهای مختلفی را به منظور رسیدن به یک ساختار بهینه برای مشخصه‌های اصلی سیل ایستگاه هیدرومتری ریپتا واقع در ایتالیا مورد آزمون قرار دادند و برخلاف مطالعه حاضر تنها به محاسبه میزان همبستگی در جفت-متغیرهای درخت اول اکتفا نمودند و در نهایت

نتیجه‌گیری کردند که استفاده از مشخصه‌های اصلی سیل با بیشترین همبستگی در درخت اول موجب رسیدن به یک ساختار بهینه نمی‌گردد و ساختاری که لزوماً دارای جفت-متغیرهای با قوی‌ترین همبستگی نبود بعنوان بهترین ساختار شناسایی گردید (Shafaei et al., 2017). علاوه بر این پس از جمع بستن میزان ضرائب تاو-کندال در کلیه جفت-متغیرهای درخت‌های هر ساختار دی-واین می‌توان نتیجه‌گیری کرد که جمع مقادیر تاو-کندال‌های درخت‌های ساختار M-L-R-D معادل ۱/۷۱ می‌باشد که این میزان نسبت به مجموع میزان تاو-کندال‌های درخت-های سایر ساختارها بیشتر می‌باشد.

تبدیل انتگرال احتمالی (PIT) ^۱ استفاده شده است (Rosenblatt, 1952). نتایج مقادیر P-value آزمون مذکور در جدول ۵ آورده شده است. مطابق با جدول ۵ مشاهده می‌شود که مقادیر P-value همه ساختارها بزرگتر از ۵ درصد می‌باشند. بنابراین می‌توان نتیجه گرفت که پارامترهای تخمینی در سطح ۵ درصد معنی‌دار می‌باشند. مطابق با جدول ۴ ساختار M-R-D-L دارای ماکزیمم میزان لوگ درست‌نمایی ۴۶۱/۸ و کمترین مقادیر ۹۱۱/۵- AIC و BIC = -۸۸۱/۱ و بعنوان دقیق‌ترین ساختار دی-واین در بین ساختارهای دی-واین چهاربعدی مشخصه‌های مهم بارش شناخته شده است. بررسی مطالعات اخیر (Aas et al., 2009; Vernieuwe et al., 2015) نشان می‌دهد که محققان برای دستیابی به ساختار بهینه واین از جفت-متغیرهای با بیشترین میزان همبستگی در درخت اول استفاده نمودند و ساختارهای دیگر دی-واین را که از جایگشت‌های مختلف متغیرها ایجاد شده است را مورد ارزیابی قرار ندادند. اما مقایسه مقادیر تاو-کندال در جفت-متغیرهای مربوط به درخت‌های ساختارهای مختلف دی-واین (جدول ۴) نشان می‌دهد که در مجموع میزان ضریب تاو-کندال در جفت-متغیرهای درخت اول ساختار M-R-D-L برابر با مقدار ۰/۹۹ می‌باشد و نسبت به سایر ساختارهای دی-واین لزوماً بیشتر نمی‌باشد و با این حال دارای حداکثر میزان لوگ درست‌نمایی و کمترین مقادیر معیارهای AIC و BIC می‌باشد. همچنین با بررسی دقیق‌تر جدول ۴ می‌توان دریافت که برخلاف آنچه در مطالعه (Aas et al., 2009; Vernieuwe et al., 2015) عنوان شده است، استفاده از ساختاری که دارای بیشترین میزان ضرائب تاو-کندال در جفت-متغیرهای درخت اول خود می‌باشد منجر به ارائه دقیق‌ترین ساختار نمی‌گردد، چرا که

¹ Probability integral transform (PIT)

جدول (۴): پارامترهای مفصل ساختار دی-واین، میزان معیارهای ارزیابی و تاو-کندال

BIC	AIC	Log-Lik	$\theta_{14 23}$	$\theta_{24 3}$	$\theta_{13 2}$	θ_{34}	θ_{23}	θ_{12}	۱-۲-۳-۴	
			درخت سوم	درخت دوم			درخت اول			
-۸۴۱/۰	-۸۶۴/۵	۴۳۸/۲	F (-۱/۲) -۰/۱۳	F (-۰/۱۵) ۰/۰۲	G (۲/۵) ۰/۶۰	Ga (۰/۴) -۰/۲۶	F (۱/۰) ۰/۱۱	G (۲/۴) -۰/۶۰	R-M-L-D Tau	
-۸۴۸/۱	-۸۷۱/۵	۴۴۱/۸	G (۳/۴) ۰/۷۱	F (-۰/۱۵) -۰/۰۲	G (۱/۷) ۰/۴۲	F (۱/۰) -۰/۱۱	Ga (۰/۴) ۰/۲۶	C (۰/۳) -۰/۱۳	R-D-L-M Tau	
-۸۸۱/۴	-۹۰۴/۸	۴۵۸/۴	F (۱/۱) ۰/۱۲	F (-۰/۷) -۰/۰۸	F (-۶/۰) -۰/۵۲	Ga (۰/۴) -۰/۲۶	G (۱/۸) ۰/۴۳	G (۲/۴) -۰/۶۰	M-R-L-D Tau	
-۸۸۱/۱	-۹۱۱/۵	۴۶۱/۸	F (-۶/۱) -۰/۵۲	G (۱/۷) ۰/۴۲	Ga (-۰/۲) -۰/۱۰	Ga (۰/۴) -۰/۲۶	C (۰/۳) ۰/۱۳	G (۲/۴) -۰/۶۰	M-R-D-L Tau	
-۸۵۳/۰	-۸۷۶/۵	۴۴۴/۲	F (۱/۰) ۰/۱۱	F (۲/۲) ۰/۲۳	G (۳/۴) ۰/۷۰	C (۰/۳) -۰/۱۳	G (۱/۸) ۰/۴۳	F (۱/۰) -۰/۱۱	M-L-R-D tau	
-۸۴۹/۵	-۸۷۲/۹	۴۴۴/۴	F (۲/۳) ۰/۲۴	Ga (-۰/۲) -۰/۱۰	G (۲/۵) ۰/۶۰	C (۰/۳) ۰/۱۳	G (۲/۴) ۰/۶۰	F (۱/۰) ۰/۱۱	L-M-R-D Tau	

G (گامبل)، F (فرانک)، Ga (گوسین)، C (کلایتون)

از ساختار دی-واین برای شبیه‌سازی مشخصه‌های بارش ایستگاه بن در استان چهارمحال و بختیاری استفاده نمودند. آنها نتیجه گرفتند که با توجه به معیارهای ارزیابی، ساختار دی-واین M-R-D-L دقیق‌ترین ساختار می‌باشد. لذا می‌توان نتیجه گرفت که ساختار دی-واین M-R-D-L بارش در مطالعه آنها و مطالعه حاضر بعنوان ساختار منتخب شناسایی شده است. ارزیابی میزان دقت و تطابق تابع توزیع توأم چهاربعدی دی-واین مشخصه‌های مهم بارش بر مفصل تجربی چهاربعدی با استفاده از معیارهای ریشه میانگین مربعات خطاها، میانگین خطای مطلق و ضریب تبیین انجام گرفت. جدول ۵ مقادیر معیارهای ارزیابی حاصل از مقایسه مفصل تجربی چهاربعدی را با توزیع توأم دی-واین چهاربعدی مشخصه‌های مهم بارش نشان می‌دهد. مطابق با معیارهای جدول ۵ ساختار دی-واین M-R-D-L دارای بیشترین تطابق با مفصل چهاربعدی دی-واین می‌باشد، چرا که در مقایسه با سایر ساختارهای دی-واین دارای کمترین $RMSE=0/031$ و $MAE=0/024$ و بیشترین $R^2=0/991$ و $P\text{-value}=0/35$ می‌باشد. مقادیر معیارهای ارزیابی در ساختارهای M-R-L-D و M-L-R-D موید این

یکی دیگر از بررسی‌هایی که در این مطالعه مورد توجه قرار گرفته است، بررسی تأثیر استفاده از جفت-متغیرهای با بیشترین همبستگی در درخت اول بر میزان همبستگی جفت-متغیرهای سایر درخت‌ها می‌باشد. همانطور که گفته شد، مقادیر ضرائب تاو-کندال در ساختار M-R-L-D دارای بیشترین میزان نسبت به مقادیر ضرائب تاو-کندال جفت-متغیرهای درخت‌های اول سایر ساختارهای دی-واین می‌باشد، اما با این حال این موضوع سبب بالا رفتن میزان تاو-کندال جفت-متغیرهای درخت‌های دوم و سوم آن نگردیده است بطوریکه حتی دو ضریب تاو-کندال آن رقم-های منفی می‌باشند. در مرحله نهایی با اعمال پارامترهای تخمین زده شده‌ی جفت-مفصل‌های مناسب هر یک از ساختارهای منتخب دی-واین (پارامترهای تخمین زده شده در جدول ۴) و همچنین سری‌های شرطی تولید شده $(h_{2|1}, h_{3|2}, h_{4|3}, h_{4|23}, h_{1|23})$ و توزیع‌های حاشیه‌ای تک متغیره (u_1, u_2, u_3, u_4) در رابطه چگالی زنجیره‌ای^۱ (رابطه ۴) می‌توان توابع توزیع توأم چهاربعدی مشخصات مهم رویدادهای بارش را برای هر یک از ساختارهای منتخب دی-واین محاسبه نمود. شفائی و همکاران (۱۳۹۵) در مطالعه‌ای

¹ Chain density

درخت‌ها همانند ساختار M-L-R-D، موجب دستیابی به دقیق‌ترین ساختار که کاملاً منطبق بر مفصل تجربی باشند، نمی‌گردد

است که لزوماً دارا بودن حداکثر همبستگی در جفت-متغیرهای درخت اول همانند ساختار M-R-L-D و یا حداکثر شدن مجموع همبستگی‌های جفت-متغیرها در کل

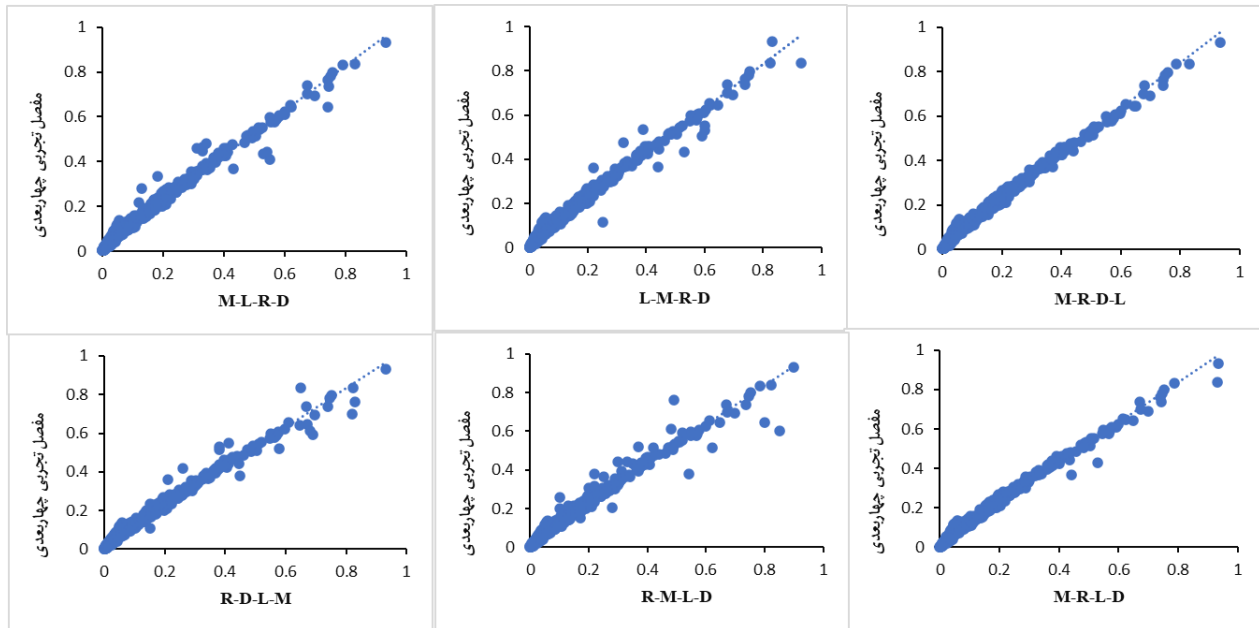
جدول (۵): مقادیر معیارهای ارزیابی حاصل از مقایسه مفصل تجربی و توزیع توأم دی- واین چهاربعدی و مقادیر تبدیل انتگرال

احتمال PIT

PIT	P-value	R ²	MAE	RMSE	ساختارها
۰/۶۵	۰/۱۲	۰/۹۶۳	۰/۰۳۳	۰/۰۴۶	R-M-L-D
۰/۶۸	۰/۲۵	۰/۹۷۷	۰/۰۳۰	۰/۰۳۹	R-D-L-M
۰/۷۰	۰/۲۸	۰/۹۸۷	۰/۰۲۹	۰/۰۳۵	M-R-L-D
۰/۷۵	۰/۳۵	۰/۹۹۱	۰/۰۲۴	۰/۰۳۱	M-R-D-L
۰/۶۷	۰/۲۷	۰/۹۷۹	۰/۰۳۰	۰/۰۳۸	M-L-R-D
۰/۶۹	۰/۲۷	۰/۹۸۰	۰/۰۳۰	۰/۰۳۸	L-M-R-D

نیمساز ۴۵ درجه نمی‌باشد)، ترسیم شده است. با توجه به شکل ۳ نقاط آبی در ساختار M-R-D-L دارای کمترین میزان پراکندگی در اطراف خط نقطه چین نسبت به سایر ساختارها می‌باشد که این موضوع دلالت بر همپوشانی بالای توزیع توأم چهاربعدی دی- واین M-R-D-L بر مفصل تجربی چهاربعدی دارد. علاوه بر این شکل ۳ نشان می‌دهد که پراکندگی نقاط آبی رنگ ساختار R-M-L-D در اطراف خط نقطه چین نسبت به ساختارهای دیگر بالاتر می‌باشد که نشانگر دقت پایین مدل مذکور می‌باشد. همانطور که مشاهده می‌شود نتایج تست گرافیکی و معیارهای ارزیابی هر دو دقت بالای ساختار M-R-D-L نسبت به سایر ساختارهای دی- واین را تایید می‌نمایند

بلکه آنچه مورد اهمیت می‌باشد، ایجاد کلیه ترکیب‌های ممکن با استفاده از جایگشت متغیرها برای ایجاد ساختارهای دی- واین با ترکیب‌های مختلف و آزمون آنها جهت رسیدن به ساختارهای بهینه می‌باشد. زیرا با توجه به مشخصات فیزیکی هر یک از مشخصات پدیده‌های هیدرولوژیکی نوع اتصال ساختار بهینه متفاوت می‌باشد و ارتباطی به میزان همبستگی آنها ندارد. به منظور مقایسه مفصل تجربی چهاربعدی با مفصل دی- واین چهار بعدی مشخصه‌های مهم بارش از روش گرافیکی نیز بهره گرفته شده است. در شکل ۳ مفصل تجربی چهاربعدی در مقابل مفصل دی- واین چهار بعدی مشخصه‌های مهم بارش (نقاط آبی) به همراه خط نقطه چین که نشان‌دهنده رابطه خطی بین دو مفصل مذکور می‌باشد (توجه: خط نقطه چین



شکل (۳): مقایسه مفصل‌های تجربی چهاربعدی با ساختارهای مختلف توزیع توأم چهاربعدی دی-واین

نتیجه‌گیری

هدف این مطالعه دستیابی به یک ساختار بهینه دی-واین چهاربعدی جهت تعیین توزیع توأم چهاربعدی مشخصات مهم رویدادهای بارش ایستگاه سریمونا در ایتالیا می‌باشد. در بین ساختارهای دی-واین چهاربعدی بدست آمده، ساختار دی-واین M-R-D-L بعنوان دقیق‌ترین ساختار دی-واین شناسایی گردید. در مرحله بعد ۶ تابع توزیع توأم چهاربعدی با استفاده از ۶ ساختار درختی دی-واین که با استفاده از جایگشت متغیرها ایجاد شده بودند، حاصل گردیدند. مطابق با معیارهای مختلف توزیع توأم چهاربعدی که با استفاده از درخت‌های ساختار M-R-D-L ساخته شده بود، بعنوان دقیق‌ترین توزیع توأم چهاربعدی مشخصه‌های مهم بارش شناخته شد. در این مطالعه برای اولین بار علاوه بر محاسبه میزان همبستگی بین کلیه جفت-متغیرهای توزیع‌های حاشیه‌ای در درخت اول، میزان همبستگی توزیع‌های شرطی تولید شده در درخت‌های دوم و سوم نیز محاسبه شده است تا تأثیر کاربرد جفت-متغیرهای با میزان همبستگی متفاوت در درخت‌های مختلف یک ساختار در دستیابی به ساختار بهینه دی-واین بررسی شود. در نهایت بررسی‌ها نشان می‌دهد که ۱-

استفاده از مشخصه‌های اصلی بارش با بیشترین همبستگی در درخت اول منجر به دستیابی به ساختار بهینه دی-واین نمی‌گردد و ساختاری که لزوماً دارای جفت-متغیرهای با قوی‌ترین همبستگی در درخت اول نبود بعنوان بهترین ساختار شناخته شد. ۲- بکارگیری جفت-متغیرهای با حداکثر میزان همبستگی در درخت اول به هیچ وجه موجب بالارفتن میزان تاو-کندال بین جفت-متغیرهای شرطی درخت‌های دوم و سوم ساختار دی-واین نمی‌گردد. ۳- حداکثر شدن جمع مقادیر تاو-کندال جفت-متغیرهای درخت‌های یک ساختار دی-واین نسبت به مجموع میزان تاو-کندال‌ها در درخت‌های سایر ساختارها موجب بالارفتن مقادیر معیار لوگ درست‌نمایی و کاهش معیارهای AIC و BIC و در نتیجه رسیدن به دقیق‌ترین ساختار نمی‌گردد. ۴- با توجه به مشخصات فیزیکی هر یک از مشخصه‌های مهم پدیده‌های هیدرولوژیکی، اجرای کلیه جایگشت‌های بین متغیرها برای دستیابی به بهترین اتصال که در نهایت منجر به ایجاد دقیق‌ترین توزیع توأم دی-واین گردد، امری ضروری می‌باشد. ۵- ساختار دی-واین، با تجزیه درختی خود منجر به تعیین دقیق توابع توزیع احتمالاتی توأم با ابعاد بالاتر از ۳ می‌گردد. ۶- مطالعات نشان دادند که ساختار منتخب دی-واین مشخصه‌های بارش برای

مفصل‌های متداول چندبعدی را حل نموده و موجب مدلسازی دقیق توابع توزیع توأم چند بعدی مشخصه‌های مهم پدیده‌های هیدرولوژیکی می‌گردد. بطوریکه افزایش دقت تخمین توابع توزیع توأم در تحلیل هرچه دقیق‌تر ریسک پدیده‌های هیدرولوژیکی نظیر سیل، بارش، خشکسالی و ... و در نتیجه طراحی مناسب سازه‌های هیدرولیکی تأثیر بسزایی خواهد داشت

ایستگاه‌های بارش بن و همچنین سریمونا یکسان می‌باشند. لیکن با این وجود برای اثبات یکسان بودن ساختارهای منتخب دی-واین برای مشخصه‌های بارش در مناطق مختلف، به نظر می‌رسد که نیاز به مطالعات بیشتری بر روی ایستگاه‌های بارش مختلف می‌باشد. در نهایت می‌توان نتیجه‌گیری نمود که مفصل دی-واین با ایجاد ارتباط بین مفصل‌های دوبعدی در ساختار درختی خود، پیچیدگی‌های

منابع

- امینی، س.، ر. زارع بیدکی، ر. میرعباسی، و م. شفائی. ۱۳۹۹. تحلیل چندمتغیره سیل با استفاده از مفصل‌های Vine در حوزه آبریز بازفت استان چهارمحال و بختیاری. نشریه مرتع و آبخیزداری، مجله منابع طبیعی ایران. دوره ۷۳، شماره ۴، جلد ۱۰، ص ۶۹۰-۶۷۴
- خانی تملیه، ذ.، ح. رضائی و ر. میرعباسی. ۱۳۹۹. کاربرد توابع مفصل تودرتو برای تحلیل فراوانی چهار متغیره خشکسالی های هواشناسی (مطالعه موردی: غرب ایران). نشریه حفاظت منابع آب و خاک. سال دهم، شماره ۱، ص ۱۱۲-۹۳.
- شفائی، م.، ا. فاخری فرد، ی. دین پژوه، و ر. میرعباسی. ۱۳۹۵. مدلسازی تابع توزیع توأم چهار بعدی ویژگی های مهم سیل با استفاده از ساختار سی-واین. نشریه آبیاری و زهکشی ایران. شماره ۳، جلد ۱۰، ص ۳۳۷-۳۲۷.
- شفائی، م.، ا. فاخری فرد، ی. دین پژوه، و ر. میرعباسی. ۱۳۹۵. مدلسازی مشخصات رویدادهای بارش با استفاده از مفصل دی-نشریه حفاظت منابع آب و خاک. سال ششم، شماره ۲، ص ۴۵-۵۸.
- Aas, K., C. Czado, A. Frigessi and H. Bakken. 2009. Pair-copula constructions of multiple dependence. *Insurance: Mathematics and Economics*, 44 (2): 182-198.
- Ariff, N.M., A.A. Jemain, K. Ibrahim and W.Z. Wan Zin. 2012. IDF relationships using bivariate copula for storm events in Peninsular Malaysia. *Journal of Hydrology*, 470-471:158-171.
- Ayantobo, O.O., Y. Li and S. Song. 2019. Multivariate Drought Frequency Analysis using Four-Variate Symmetric and Asymmetric Archimedean Copula Functions. *Water Resources Management*, 33: 103-127.
- Bedford, T. and R. Cooke. 2001. Probability density decomposition for conditionally dependent random variables modeled by vines. *Annals of Mathematics and Artificial Intelligence*, 32 (1): 245-268.
- Bedford, T. and R. Cooke. 2002. Vines – A new graphical model for dependent random variables, *Annals of Statistics*, 30 (4): 1031-1068.
- Brechmann, E. C., C. Czado and K. Aas. 2012. Truncated regular vines in high dimensions with applications to financial data. *Canadian Journal of Statistics*, 40 (1): 68-85.
- Buliah, N.M. and W. Yie. 2020. Modelling of extreme rainfall using copula. *AIP Conference Proceedings* 2266, 090007.



- Christopher Dzapire, N., P. Ngarea L. Odongoac, 2020. A copula based bi-variate model for temperature and rainfall processes. *Scientific African*, 8: e00365.
- Czado C. .2010. Pair-Copula Constructions of Multivariate Copulas. In: Jaworski P., Durante F., Härdle W., Rychlik T. (eds) *Copula Theory and Its Applications*. Lecture Notes in Statistics, vol 198. Springer, Berlin, Heidelberg.
- De Michele, C and G. Salvadori. 2003. A generalized Pareto intensity-duration model of storm rainfall exploiting 2-copulas. *Journal of Geophysical Research*, 108(D2): 4067.
- De Michele, C., G. Salvadori, M. Canossi, A. Petaccia and R. Rosso. 2005. Bivariate statistical approach to check adequacy of dam spillway. *Journal of Hydrologic Engineering*, 10 (1): 50–57.
- De Michele, C., G. Salvadori, G. Passoni and R. Vezzoli. 2007. A multivariate model of sea storms using copulas. *Coastal Engineering*, 54 (10): 734–751.
- Dodangeh, E., V.P. Singh, and B.T. Pham. 2020. Flood Frequency Analysis of Interconnected Rivers by Copulas. *Water Resources Management*, 34: 3533–3549.
- Genest, C. and L.P. Rivest. 1993. Statistical inference procedures for bivariate Archimedean copulas. *Journal of the American Statistical Association*, 88 (423): 1034–1043.
- Ghafori, V., H. Sedghi and R.A. Sharifan. 2020. Regional Frequency Analysis of Droughts Using Copula Functions (Case Study: Part of Semiarid Climate of Fars Province, Iran). *Iranian Journal of Science and Technology* 44, 1223–1235.
- Ghosh, S. 2010. Modelling bivariate rainfall distribution and generating bivariate correlated rainfall data in neighboring meteorological subdivisions using copula. *Hydrological Processes*. 24, 3558–3567.
- Gräler, B., M. J. Van den Berg, S.Vandenberghe, A. Petroselli, S. Grimaldi, B. De Baets, and Verhoest, N. E. C. 2013. Multivariate return periods in hydrology: a critical and practical review focusing on synthetic design hydrograph estimation. *Hydrology and Earth System Science*, 17 (4): 1281–1296.
- Joe, H. 1996. Families of m-variate distributions with given margins and $m(m-1)/2$ bivariate dependence parameters. In L. Rueschendorf, B. Schweizer, and M. D. Taylor (Eds.), *Distributions with fixed marginal and related topics*. Hayward: Institute of Mathematical Statistics. 28: 120-141.
- Kao, SC. and RS. Govindaraju. 2010. A copula-based joint deficit index for droughts. *Journal of Hydrology*, 380(1–2):121–134.
- Kendall, M. G. .1938. A new measure of rank correlation, *Biometrika*, 30(1-2): 81–93
- Kurowicka, D. and R. Cooke. 2006. *Uncertainty Analysis with High Dimensional Dependence Modelling*. Wiley, Chichester, 308 p.
- Kuchment, L. S. and V. N. Demidov. 2013. On the Application of Copula Theory for Determination of Probabilistic Characteristics of Springflood., *Russian Meteorology and Hydrology*, 38(4): 263–271.
- Mirabbasi, R., A. Fakheri-Fard and Y. Dinpashoh. 2012. Bivariate drought frequency analysis using the copula method. *Theoretical and Applied Climatology*, 108 (1): 191–206.
- Mirabbasi, R., E.N. Anagnostou, A. Fakheri-Fard, Y. Dinpashoh and S. Eslamian. 2013. Analysis of meteorological drought in northwest Iran using the Joint Deficit Index. *Journal of Hydrology*, 492: 35–48.



- Ni, L., D. Wang, J. Wu, Y. Wang, Y. Tao, J. Zhang, J. Liu and F. Xie. 2020. Vine copula selection using mutual information for hydrological dependence modeling. *Environmental Research*, 186:1-12.
- Palynchuk, B.A. and Y. Guo. 2011. A probabilistic description of rain storms incorporating peak intensities. *Journal of Hydrology*, 409: 71–80.
- Pandey, P.K., L. Das. and D. Jhajharia. 2018. Modelling of interdependence between rainfall and temperature using copula. *Modeling Earth Systems and Environment*, 4: 867–879.
- Rosenblatt, M. 1952. Remarks on a multivariate transformation. *Annals of Mathematical Statistics*, 27: 832–837.
- Salvadori, G. and C. De Michele. 2015. Multivariate real-time assessment of droughts via copula-based multi-site Hazard Trajectories and Fans. *Journal of Hydrology*, 526: 101-115.
- Shafaei, M., A. Fakheri-Fard, Y. Dinpashoh, R. Mirabbasi and C. De Michele. 2017. Modeling flood event characteristics using D-vine structures, *Theoretical and Applied Climatology*, 130: 713–724.
- Shiau, J. T. 2006. Fitting drought duration and severity with two-dimensional copulas. *Water Resources Management*, 20: 795–815.
- Song, S and V. Singh. 2010. Meta-elliptical copulas for drought frequency analysis of periodic hydrologic data. *Stochastic Environmental and Research Risk A*, 24 (3): 425–444.
- Sraj, M., N. Bezak and M. Brilly. 2015. Bivariate flood frequency analysis using the copula function: a case study of the Litija station on the Sava River. *Hydrologic Processes*, 29 (2): 225–238.
- Vandenberghe, S., N. E. C. Verhoest, C. Onof and B. De.Baets. 2011. A comparative copula-based bivariate frequency analysis of observed and simulated storm events: A case study on Bartlett-Lewis modeled rainfall. *Water Resources Researches*, 47, W07529: 1-16.
- Vernieuwe, H., S. Vandenberghe, B. De Baets and NEC. Verhoest. 2015. A continuous rainfall model based on vine copulas. *Hydrology and Earth System Science*, 19(6): 2685–2699.
- Zhang, L., and V. P. Singh. 2006. Bivariate rainfall frequency distributions using Archimedean copulas. *Journal of Hydrology*, 332 (1–2): 109-193.
- Zhang, L., V. P. Singh. 2006. Bivariate flood frequency analysis using copula method. *Journal of Hydrologic Engineering*, 11 (2): 150–164.



Extraction of joint probabilistic distribution functions of characteristics of precipitation using a four-dimensional D-vine tree structure

Maryam Shafaei¹. Rasoul Mirabbasi Najafabadi^{2*}

Multivariate frequency analysis of hydrological phenomena by considering the dependence between the basic characteristics of these phenomena will lead to more their accurate estimation. Due to the high flexibility provided by vine tree copulas in problems with dimension greater than two. In this study the D-vine function is used to determine the four-dimensional probabilistic distribution function of main characteristics of the precipitation events of Cremona station in Italy (including maximum rainfall intensity, total rainfall depth, duration of wet period and dry period). First, due to the significant dependence between the main characteristics of precipitation events and also using their permutation, D-vine tree structures were obtained. After fitting the various Archimedean and elliptical copula families to the pair-copulas of each D-vine tree structure, the most suitable copula families were determined for fitting the pair-copulas of each D-vine structure by the maximum log-likelihood, Akaike (AIC) and Bayesian Information Criterion (BIC). Then, in order to evaluate the accuracy of the four-dimensional probabilistic distribution functions of the important characteristics of precipitation events, the mentioned functions were compared with the corresponding four-dimensional empirical copulas. Finally, the M-R-D-L four-dimensional D-vine structure according to the evaluation criteria of $R^2 = 0.991$, $RMSE = 0.031$, and $MAE = 0.024$, was selected as the most appropriate structure for constructing of the joint distribution function of main characteristics of precipitation in Cremona station.

Key words: Copula function, D-vine, Multivariate Frequency analysis, Pair-copula, Precipitation characteristics.

¹Ph. D in Water Resources Engineering, Department of Water Engineering, Tabriz University, Tabriz. Iran.

² Associate Professor. Department of Water Engineering, Shahrekord University, Shahrekord, Iran.

*- Corresponding Author Email: mirabbasi_r@yahoo.com