

## مقایسه عملکرد مفصل‌های درختی سی-واین و دی-واین در تحلیل چندمتغیره مشخصه‌های بارش

مریم شفاعی<sup>۱\*</sup> و رسول میرعباسی نجف آبادی<sup>۲</sup>

\* دانش آموخته دکتری مهندسی منابع آب، گروه مهندسی آب دانشگاه تبریز، تبریز، ایران.

<sup>۲</sup> دانشیار گروه مهندسی آب، دانشگاه شهرکرد، شهرکرد، ایران.

\* ایمیل نویسنده مسئول: m.shafaei65@gmail.com

تاریخ پذیرش: ۱۴۰۰/۰۶/۲۹

تاریخ دریافت: ۱۴۰۰/۰۳/۰۴

### چکیده:

در این مطالعه، قابلیت‌های مفصل‌های درختی واین، نظیر توانایی در تجزیه توزیع‌های چند متغیره به توزیع‌های دوبعدی، انعطاف آن در مسائل با ابعاد بالا و استفاده از وابستگی شرطی بین متغیرها مورد بررسی قرار گرفته است. هدف استفاده از ساختارهای درختی سی-واین و دی-واین جهت تعیین تابع توزیع احتمالاتی توأم چهاربندی مشخصه‌های مهم رویدادهای بارش ۲۶ ساله ایستگاه بارش سریمونا واقع در کشور ایتالیا، شامل ماکزیمم شدت بارش، عمق کل بارش، مدت زمان دوره مرطوب و دوره خشک می‌باشد. بدین منظور، ابتدا ترکیبی از مناسب‌ترین خانواده‌های مفصل ارشمیدسی و بیضوی جهت برازش بر جفت-مفصل‌های هر یک از ساختارهای سی-واین و دی-واین مشخص شدند. توابع توزیع توأم بهینه ساختارهای سی-واین و دی-واین نیز با استفاده از توابع چگالی زنجیره‌ای محاسبه و میزان انطباق آنها با مفصل تجربی چهاربندی مشخصه‌های بارش متناظر مورد بررسی قرار گرفتند. در نهایت میزان دقت ساختارهای درختی سی-واین و دی-واین در تعیین توابع توزیع توأم مشخصه‌های مهم بارش، مقایسه گردید. نتایج نشان داد که ساختار سی-واین R-D-L-M دارای مینیمم مقدار معیارهای ارزیابی  $RMSE=0/029$  و  $MAE=0/022$  و همچنین ماکزیمم  $P\text{-value}=0/35$  و  $R^2=0/998$  در بین کلیه ساختارهای سی-واین و دی-واین می‌باشد و در نتیجه برای تحلیل فراوانی مشخصات بارش ایستگاه سریمونا ایتالیا دارای بیشترین دقت می‌باشد.

**کلید واژه‌ها:** بارش؛ توزیع توأم؛ مفصل؛ واین؛ جفت-مفصل

### مقدمه

بین دو یا چند متغیر بکار می‌رود. توابع مفصل، امکان استفاده همزمان از چند توزیع تک متغیره (با توزیع‌های مختلف) را برای ساخت یک توزیع چند متغیره (که وابستگی بین متغیرها را نیز در نظر می‌گیرد) فراهم می‌آورند (Shiau, 2006). گرچه مفصل‌ها در مدل‌سازی دوبعدی وابستگی‌ها دارای توانایی بالایی می‌باشند، اما اغلب ایجاد مفصل‌های با ابعاد بالاتر از دو سبب افزایش بار محاسباتی می‌گردد. در سال‌های اخیر بررسی‌هایی در

پدیده‌های هیدرولوژیکی، اغلب دارای مشخصه‌های مختلفی هستند که برای تحلیل فراوانی آنها، بایستی مؤلفه‌های مهم پدیده مورد نظر به صورت توأم مورد بررسی قرار گیرد، چرا که مطالعه تک‌متغیره پدیده‌های مذکور اغلب موجب ایجاد خطا در مدل‌سازی آنها می‌گردد. توابع مفصل<sup>۱</sup> برای مدل سازی ساختار وابستگی

<sup>۱</sup> Copula Function

مواجهه با مسائل با ابعاد بالا با محدودیت‌هایی مواجه می‌شوند. مفصل درختی و این با تجزیه توابع توزیع چندمتغیره به مفصل‌های دوبعدی، انعطاف لازم را در مسائل با ابعاد بالاتر از دو ایجاد نموده و فاقد ضعف‌های مفصل‌های متداول می‌باشد. در سال‌های اخیر توابع مفصل و این بعنوان ساختارهای درختی با انعطاف‌پذیری بالا در علوم مختلف مورد توجه قرار گرفته‌اند. با این وجود از مفصل و این در تعیین توابع توزیع رویدادهای هیدرولوژیکی چندمتغیره و شبیه‌سازی آنها در مطالعات معدودی استفاده شده است. Graler و همکاران (2013) از ساختار دی-واین سه بعدی جهت تحلیل فراوانی مشخصه‌های سیل استفاده نمودند. شفائی و همکاران (۱۳۹۵) ابتدا با استفاده از روش ترتیبی و معیارهای ارزیابی مختلف ترکیبی از خانواده‌های مفصل بهینه را انتخاب و سپس با استفاده از توابع چگالی زنجیره‌ای ساختار سی-واین، توابع توزیع توأم چهاربعدی مشخصه-های مهم سیل (زمان پایه، زمان تا اوج، حجم کل و دبی اوج سیل) را محاسبه نمودند. Ni و همکاران (2020) توابع توزیع احتمالاتی چهاربعدی مشخصه‌های مهم رویدادهای خشکسالی را با کاربرد ساختارهای درختی سی-واین و دی-واین محاسبه نمودند. امینی و همکاران (۱۳۹۹) برای تجزیه و تحلیل سه بعدی و چهاربعدی مشخصات سیل ایستگاه لندی واقع در حوضه آبخیز بازفت استان چهارمحال و بختیاری از توابع مفصل سی-واین و دی-واین استفاده نمودند. در این مطالعه، تجزیه درختی مفصل‌های دی-واین و سی-واین جهت محاسبه توابع توزیع توأم چهاربعدی مشخصه‌های مهم بارش ایستگاه باران‌سنجی سریمونا واقع در کشور ایتالیا (شامل ماکزیمم شدت بارش، عمق کل بارش، مدت زمان دوره مرطوب و دوره خشک) بکار برده شده است. علاوه بر این، برای اولین بار میزان دقت مفصل‌های سی-واین و دی-واین در تعیین توابع توزیع توأم چهاربعدی مشخصه‌های بارش مورد مقایسه قرار گرفت.

خصوص ساخت مفصل چندمتغیره از مفصل‌های دوبعدی ارشمیدسی با استفاده از ساختمان ارشمیدسی تودرتو<sup>۱</sup> انجام شده است. عبداللهی اسدآبادی و همکاران (۱۳۹۷) مدلی احتمالاتی و چندمتغیره را برای شبیه‌سازی فرایند پیچیده بارش-رواناب با بهره‌گیری از مفهوم توابع مفصل نامتقارن ارشمیدسی یا تودرتو توسعه دادند. علاوه بر این خانی تملیه و همکاران (۱۳۹۹)، برای تحلیل فراوانی چهارمتغیره مشخصات پدیده خشکسالی غرب کشور از مفصل تودرتو چهاربعدی استفاده نمودند. مفصل‌های بیضوی، نظیر گوسین یا تی-استیوندت نیز از جمله مفصل‌های محبوب در حل مسائل با ابعاد بالاتر از دو می‌باشند. 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 ... به سرعت رو به فزونی می‌باشد. بطوریکه در زمینه تحلیل فراوانی پدیده سیل می‌توان به مطالعات

(Karmakar and Simonovic, 2009; Kuchment and Demidov, 2013; Sraj et al., 2015; Dodangeh et al., 2020) در زمینه تحلیل رویدادهای خشکسالی به مطالعات (Shiau, 2006; Kao and Govindaraju, 2010; Mirabbasi et al., 2013; Salvadori and De Michele, 2015; Ghafari et al., 2019) و در زمینه تحلیل فراوانی بارش به بررسی‌های (De Michele and Salvadori, 2003; Ghosh, 2010; Ariff et al., 2012; Yee et al., 2014; ; Pandey et al., 2018; Buliah and Yie. 2020; Christopher Dzipire et al., 2020) اشاره نمود. تاکنون اکثر مطالعات کاربرد توابع مفصل، بر کاربرد مفصل‌های دوبعدی و سه‌بعدی در تحلیل فراوانی پدیده‌های هیدرولوژیکی متمرکز بوده‌اند و به ندرت مطالعه‌ای در زمینه تحلیل فراوانی با مفصل چهاربعدی یا بالاتر انجام شده است. اما بطور کلی، استفاده از توابع مفصل در مسائل با ابعاد بالا سبب افزایش پیچیدگی ریاضیاتی می‌شود و هر یک از مفصل‌های چندمتغیره در

<sup>۱</sup> Nested Copula

## مواد و روش‌ها

### منطقه مورد مطالعه و داده‌های مورد استفاده

در این مطالعه، آمار بارش ساعتی ایستگاه باران‌سنجی سریمونا واقع در شهر سریمونا در کشور ایتالیا به مدت ۲۶ سال (۱۹۹۵-۲۰۲۰) بدلیل لزوم بکارگیری آمار کامل و طولانی‌مدت بارش ساعتی، مورد استفاده قرار گرفته است. شهر سریمونا در شمال ایتالیا و در ساحل چپ رودخانه پرو و در طول جغرافیایی  $10^{\circ}2'$  و عرض جغرافیایی  $45^{\circ}8'$  و در ارتفاع ۴۹ متری از سطح دریا قرار دارد. میانگین دمای آن  $13/9$  درجه سانتی‌گراد و میانگین بارش آن ۹۸۹ میلی‌متر در سال می‌باشد موقعیت ایستگاه باران سنجی سریمونا واقع در شهر سریمونا در کشور ایتالیا در شکل ۱ نشان داده شده‌اند.



شکل ۱. موقعیت ایستگاه باران‌سنجی سریمونا واقع در شهر سریمونا در کشور ایتالیا

لازم به ذکر است که Salvadori و De Michele (۲۰۰۳) برای استخراج رویدادهای بارش از معیار توقف ۷ ساعته استفاده نمودند. بنابراین در این بررسی، برای جداسازی رویدادهای بارش، از روش معیار توقف ۷ ساعته بارش استفاده شد و بدین ترتیب ۳۶۶ رویداد بارش استخراج گردیدند. در این مطالعه، عمق کل بارش (R) معادل مجموع عمق بارش‌های ساعتی در هر رویداد بارش و ماکزیمم شدت بارش (M) معادل ماکزیمم میزان بارش ساعتی در هر رویداد بارش می‌باشند. علاوه بر این، دوره مرطوب (L) به دوره‌ای اطلاق می‌شود که رویدادهای بارش دقیقاً در آن ساعات رخ داده‌اند و دوره خشک (D) به مدت زمان دوره خشکی گفته می‌شود که در یک رویداد بارش واقع شده است و مدت زمان آن کمتر از ۷ ساعت باشد. در نهایت، مشخصه‌های مهم رویدادهای بارش استخراج شدند که مشخصات آماری آنها در جدول (۱) ارائه شده است.

### مفصل واین

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 مواجهه با مسائل با ابعاد بالا با محدودیت‌هایی مواجه می‌شوند. بعنوان مثال در اینگونه موارد، بکارگیری مفصل‌های ارشمیدسی نامتقارن تودرتو دارای مشکلاتی من جمله وابستگی‌های مثبت بین متغیرهای درونی و کوچکتر بودن وابستگی بیرونی نسبت به وابستگی درونی می‌باشد که این موضوع

جدول ۱- مشخصات آماری مشخصه‌های رویدادهای بارش مورد استفاده در مطالعه (ایستگاه سریمونا طی سالهای سال (۱۹۹۵-۲۰۲۰))

آماره‌های مشخصه‌های بارش	مدت زمان دوره خشک (hr)	مدت زمان دوره مرطوب (hr)	ماکزیمم بارش (mm)	عمق کل بارش (mm)
ماکزیمم	۲۸/۰	۶۳/۰	۳۵/۰	۱۶۱/۴
مینیمم	۰	۳/۰	۰/۲	۰/۸
انحراف معیار	۴/۸	۹/۷	۵/۰	۱۹/۹
میانگین	۴/۹	۱۴/۹	۴/۲	۱۹/۵

موجب محدودیت‌هایی در استفاده از این مفصل‌ها در مسائل با ابعاد بالاتر از دو می‌باشد. اما ساختار مفصل و این با تجزیه توزیع‌های چندمتغیره به توزیع‌های توأم با بعد دو، انعطاف‌پذیری لازم در مسائل با ابعاد بزرگتر را فراهم آورده و فاقد ضعف‌های مفصل‌های تودرتو می‌گردد. با توجه به محدودیت‌های مفصل‌های چندبعدی (ابعاد بالاتر از دو) در تعیین توابع توزیع توأم چندبعدی، برای اولین بار جو در سال ۱۹۹۶، مفصل‌های چندمتغیره‌ای را که ساختار آنها شامل بلوک‌های مفصل دومتغیره می‌شد را معرفی نمود (Joe, 1996). در علم هیدرولوژی متغیرهای پدیده‌های هیدرولوژیکی نقش متغیرهای مفصل‌های و این را بازی می‌کنند. بعنوان مثال در مطالعه حاضر مشخصه‌های وقوع عمق کل بارش (R)، ماکزیمم شدت بارش (M) مدت زمان دوره مرطوب (L) و دوره خشک (D) نقش متغیرهای و این را بازی می‌کنند. بهمین صورت در هر قسمتی از مقاله که به متغیرهای مفصل اشاره شده است در اصل منظور همان مشخصه‌های هیدرولوژیکی می‌باشند.

مجموعه‌ای از درخت‌های متصل به هم به منظور اتصال بین جفت-مفصل‌ها مورد استفاده قرار گرفت که و این نرمال<sup>۱</sup> نامیده می‌شوند (Bedford and Cooke, 2001, 2002). مدل‌های و این شامل دنباله‌ای از درخت‌ها با حداکثر N متغیر و حداکثر N-1 لبه می‌باشند. هر لبه نشان‌دهنده یک مفصل دوبعدی در فضای  $[0, 1]^2$  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 مفصل-های دوبعدی می‌باشند. به نظر می‌رسد مدل‌های گرافیکی (درخت‌های متصل بهم) مدل‌های مناسبی برای ارائه توزیع‌های احتمالاتی توأم با ابعاد بالاتر از دو باشند، بطوریکه این مدل‌ها، بطور تصویری به توصیف وابستگی-

های موجود در توزیع‌های توأم می‌پردازند. یکی از ویژگی‌های مهم مدل‌های گرافیکی، توانایی آنها در توصیف ساختارهای پیچیده به شکل مدولار می‌باشد، بطوریکه در مدل‌های مذکور، وابستگی بین عناصر همسایه ترکیب می‌شوند. و این نرمال به دو نوع سی- و این (C-vine) و دی- و این (D-vine) تقسیم‌بندی می‌شوند. ساختار دی- و این شامل درخت‌های با ساختار خطی و ساختار سی- و این شامل درخت‌های با ساختار ستاره‌ای می‌باشند، بطوریکه هر یک از درخت‌های ساختار سی- و این دارای یک متغیر مرکزی است. بدفورد و کوک تجزیه‌های جفت-مفصل را بطور گرافیکی و بصورت دنباله‌ای از درخت‌های تودرتو (درخت و این) با لبه‌های بدون جهت نمایش دادند (Bedford and Cooke, 2001, 2002). شیوه تجزیه تابع چگالی توزیع چهاربعدی توسط ساختارهای سی- و این و دی- و این چهاربعدی به ترتیب در شکل‌های ۲ و ۳ نشان داده شده است. همانطور که مشاهده می‌شود ساختارهای و این از گره‌ها و لبه‌ها تشکیل شده‌اند که به ترتیب برای نمایش دادن متغیرها و جفت-مفصل بکار برده می‌شوند. مطابق با شکل‌های ۲ و ۳، 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. علاوه بر این در ساختار گرافیکی درخت‌های مدل سی- و این، متغیر مرکزی حداقل به دو متغیر دیگر متصل می‌باشد و سایر متغیرها حداکثر به یک متغیر مرکزی اتصال دارند. اما در مدل دی- و این، متغیرها بصورت خطی (پشت سرهم) قرار می‌گیرند و هر متغیر حداکثر به دو متغیر متصل می‌شود. بعنوان مثال، با توجه به شکل ۲ در ساختار سی- و این چهاربعدی، درخت اول  $T_1$  از چهار متغیر (گره)  $s_1, s_2, s_3$  و  $s_4$  و لبه‌های غیرشرطی  $C_{12}, C_{13}, C_{14}$  تشکیل شده است. لیکن همانطور که در شکل ۲ نمایش داده شده است، گره‌های درخت  $T_2$  شامل توابع توزیع شرطی  $F_{3|1}, F_{2|1}F_{4|1}$  می‌باشند که از مشتق جزئی

<sup>۱</sup> Regular Vine

(۲)

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

تابع  $C$  یک تابع مفصل منحصر بفرد  $n$  بعدی می‌باشد که توزیع‌های یکنواخت حاشیه‌ای تک متغیره را به توزیع تجمعی  $n$  بعدی توأم نگاشت می‌نماید. علاوه بر این  $\theta$  پارامتر مفصل و  $u_i$  توزیع حاشیه‌ای می‌باشند.

### تابع چگالی توأم چهاربعده سی-واین و دی-واین

به منظور تعیین توابع چگالی توأم سی-واین، با توجه به تجزیه درختی سی-واین، از قانون زنجیره‌ای تابع چگالی توأم چهاربعده سی-واین استفاده می‌شود. رابطه تابع چگالی زنجیره‌ای با روش سی-واین در رابطه (۳) ارائه شده است (Aas et al., 2009):

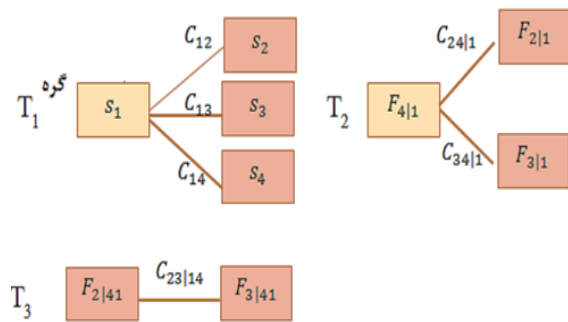
(۳)

$$\begin{aligned} f(s_1, s_2, s_3, s_4) = & f(s_1) \cdot f(s_2) \cdot f(s_3) \cdot f(s_4) \cdot c_{12}\{F(s_1), F(s_2), \theta_{12}\} \cdot \\ & c_{13}\{F(s_1), F(s_3), \theta_{13}\} \cdot c_{14}\{F(s_1), F(s_4), \theta_{14}\} \cdot \\ & c_{24|1}\{F(s_2|s_1), F(s_4|s_1), \theta_{24|1}\} \cdot \\ & c_{34|1}\{F(s_3|s_1), F(s_4|s_1), \theta_{34|1}\} \cdot \\ & c_{23|14}\{F(s_2|s_4, s_1), F(s_3|s_4, s_1), \theta_{23|14}\} \end{aligned}$$

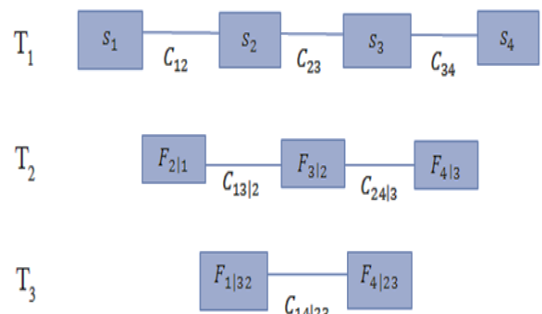
در رابطه بالا،  $C$  تابع چگالی جفت-مفصل می‌باشد. تجزیه توابع چگالی چندبعده (با ابعاد بالاتر از دو) به توابع چگالی جفت-مفصل و توابع چگالی تک بعدی، برای اولین بار توسط بدفورد و کوک شرح داده شد (Bedford and Cooke, 2001, 2002). مطابق با رابطه (۳)،

توابع چگالی جفت-مفصل  $C$  غیرشرطی بصورت تابعی از توابع توزیع غیرشرطی تعریف می‌شوند. در حالیکه توابع چگالی جفت-مفصل  $C$  شرطی بصورت تابعی از توابع توزیع شرطی می‌باشند. لازم بذکر است که توابع توزیع حاشیه‌ای شرطی با استفاده از مشتق جزئی تابع جفت-مفصل دوبعدی قابل محاسبه می‌باشند. محاسبه توابع توزیع حاشیه‌ای شرطی موجود در رابطه تابع چگالی زنجیره‌ای سی-واین با استفاده از روابط (۴) تا (۸) انجام می‌گیرد (Joe, 1996):

جفت-مفصل‌های (لبه‌ها) درخت اول، یعنی  $C_{12}, C_{13}, C_{14}$  حاصل می‌شوند. علاوه بر این، گره‌های درخت  $T_3$  شامل توابع توزیع شرطی  $F_{2|41}$  و  $F_{3|41}$  می‌باشند که از طریق مشتق جزئی لبه‌های درخت دوم یعنی  $C_{24|1}$  و  $C_{34|1}$  بدست می‌آیند. روش محاسبه توابع توزیع شرطی با استفاده از مشتق جزئی مفصل‌ها در بخش بعدی ارائه خواهد شد.



شکل ۲. نمایش شماتیک درخت‌های ساختار سی-واین چهار بعدی



شکل ۳. نمایش شماتیک درخت‌های ساختار دی-واین چهار بعدی

### توابع مفصل

اگر  $F$  توزیع توأم چندمتغیره متغیره‌های تصادفی  $s_1, s_2, \dots, s_n$  بترتیب با توزیع‌های حاشیه‌ای  $F_{S_1}(s_1), F_{S_2}(s_2), \dots, F_{S_N}(s_N)$  باشد، بنابراین، مفصل  $C$  در دامنه  $R$  برای متغیره‌های  $s_1, s_2, s_3, \dots, s_n$  بصورت زیر تعیین می‌شود (Aas et al., 2009):

(۱)

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

محاسبه پارامترهای ساختارهای سی-واین و دی-واین

چهاربعدی

تخمین پارامترهای ساختارهای سی-واین چهار بعدی به روش ترتیبی بشرح زیر انجام می‌گیرد (Aas et al., 2009):

۱- انتخاب یک خانواده جفت-مفصل مناسب بر توزیع-های حاشیه‌ای  $u_1$  و  $u_2$  جهت تخمین پارامتر مفصل  $\theta_{12}$ .

۲- انتخاب یک خانواده جفت-مفصل مناسب بر توزیع-های حاشیه‌ای  $u_1$  و  $u_3$  جهت تخمین پارامتر مفصل  $\theta_{13}$ .

۳- انتخاب یک خانواده جفت-مفصل مناسب بر توزیع-های حاشیه‌ای  $u_1$  و  $u_4$  جهت تخمین پارامتر مفصل  $\theta_{14}$ .

۴- محاسبه سری  $h_{2|1}$  با توجه به پارامتر مفصل  $\theta_{12}$  رابطه (۴)  $h_{2|1} = h(u_1, u_2, \theta_{12})$ .

۵- محاسبه سری  $h_{3|1}$  با توجه به پارامتر مفصل  $\theta_{13}$  و رابطه (۵)  $h_{3|1} = h(u_1, u_3, \theta_{13})$ .

۶- محاسبه سری  $h_{4|1}$  با توجه به پارامتر مفصل  $\theta_{14}$  رابطه (۶)  $h_{4|1} = h(u_1, u_4, \theta_{14})$ .

۷- انتخاب یک خانواده جفت-مفصل مناسب بر توزیع-های حاشیه‌ای  $h_{2|1}$  و  $h_{4|1}$  جهت تخمین پارامتر مفصل  $\theta_{24|1}$ .

۸- برآزش و انتخاب یک خانواده جفت-مفصل مناسب بر  $h_{3|1}$  و  $h_{4|1}$  جهت تخمین پارامتر مفصل  $\theta_{34|1}$ .

۹- محاسبه سری  $h_{2|41}$  با قرار دادن پارامتر مفصل  $\theta_{24|1}$  در رابطه (۷)  $h_{2|41} = h(h_{2|1}, h_{4|1}, \theta_{24|1})$ .

۱۰- محاسبه سری  $h_{3|41}$  با قرار دادن پارامتر مفصل  $\theta_{34|1}$  در رابطه (۸)  $h_{3|41} = h(h_{3|1}, h_{4|1}, \theta_{34|1})$ .

مفصل مناسب دویبعدی بر  $h_{3|41}$  و  $h_{2|41}$  جهت تخمین پارامتر مفصل  $\theta_{23|41}$ .

تخمین پارامترهای ساختار درختی دی-واین نیز با استفاده از روش ترتیبی و مشابه با تخمین پارامتر در روش سی-واین می‌باشد، که در این مقاله به جهت جلوگیری از افزایش حجم مطالب تنها به تشریح روش تخمین پارامترهای ساختار سی-واین اکتفا شده است (Aas et al., 2009).

$$h_{2|1} = F(s_2|s_1) = \frac{\partial C_{12}(F(s_1), F(s_2), \theta_{12})}{\partial F(s_1)} \quad (۴)$$

$$h_{3|1} = F(s_3|s_1) = \frac{\partial C_{13}(F(s_3), F(s_1), \theta_{13})}{\partial F(s_1)} \quad (۵)$$

$$h_{4|1} = F(s_4|s_1) = \frac{\partial C_{14}(F(s_4), F(s_1), \theta_{14})}{\partial F(s_1)} \quad (۶)$$

$$h_{2|41} = F(s_2|s_4, s_1) = \frac{\partial C_{24|1}(F(s_2|s_1), F(s_4|s_1); \theta_{24|1})}{\partial F(s_4|s_1)} \quad (۷)$$

$$h_{3|41} = F(s_3|s_4, s_1) = \frac{\partial C_{34|1}(F(s_3|s_1), F(s_4|s_1); \theta_{34|1})}{\partial F(s_4|s_1)} \quad (۸)$$

مطابق با قانون زنجیره‌ای، تابع توزیع چگالی توأم چهاربعدی دی-واین بصورت رابطه (۹) تعریف می‌گردد:

$$(۹)$$

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

با توجه به روابط بالا، می‌توان نتیجه گرفت که بطور کلی توزیع تجمعی 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 دویبعدی  $C$  نسبت به متغیر دوم مشتق جزئی گرفته می‌شود. پارامتر  $\theta_{ij|k}$  تابع توزیع مفصل توأم بین دو متغیر  $F(s_i|s_k)$  و  $F(s_j|s_k)$  می‌باشد (Aas et al., 2009). در این مقاله با توجه به مشابهت روش محاسبه توابع توزیع شرطی مربوط به تابع چگالی ساختار D-Vine با C-vine و همچنین به منظور جلوگیری از افزایش حجم مقاله تنها به تشریح روش محاسبه توابع توزیع شرطی موجود برای ساختار سی-واین بسنده شده است.

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

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

$n$  و  $k$  لوگ درست نمائی و  $L(u, h, \theta)$  در روابط فوق، نیز  $c_{ij}$  به ترتیب تعداد پارامترها و تعداد داده‌ها می‌باشند. دلالت دارد. زو ابر چگالی مفصل دویبعی بین دو متغیر بررسی دقت ساختار توزیع‌های توأم چهاربعی سی-واین و دی-واین

در این مطالعه، برای ارزیابی میزان دقت تابع توزیع توأم چهاربعی منتخب با مفصل تجربی<sup>۳</sup> چهاربعی مشخصه‌های مهم رویدادهای بارش مقایسه گردید. مفصل تجربی مشخصه‌های 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:

$$(14)$$

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

در رابطه بالا، I تابع شاخص، n تعداد داده‌ها،  $R_i, M_i, L_i, D_i$  رتبه هریک از نمونه‌ها در مشخصه‌های رویدادهای بارش می‌باشند (Genest and Rivest, 1993). همچنین متغیرهای  $r, m, l, d$  نشاندهنده رقم‌هایی هستند که به ترتیب برای محاسبه احتمال عمق کل بارش (R)، مدت زمان دوره مرطوب (L) و دوره خشک (D) در رقم‌های پایین تر از  $r, m, l, d$  بکار می‌روند. در نهایت سه معیار ارزیابی ریشه میانگین مربعات خطا (RMSE)، میانگین خطای مطلق (MAE) و ضریب تبیین ( $R^2$ ) جهت ارزیابی میزان تطابق توابع توزیع‌های توأم چهاربعی سی-واین و دی-واین تعیین شده بر مفصل تجربی چهاربعی استفاده شدند. معیارهای ارزیابی مذکور مطابق با روابط (۱۵) تا (۱۷) محاسبه می‌شوند:

$$(17)$$

## انتخاب مناسب‌ترین خانواده مفصل برای هر جفت-مفصل ساختارهای سی-واین و دی-واین

ابتدا هریک از مفصل‌های خانواده‌های ارشمیدسی (کلاتون، گامبل، جو و فرانک) و بیضوی (تی-استیودنت و گوسی) بر هر یک از جفت-مفصل‌های ساختارهای سی-واین و دی-واین برازش داده می‌شوند. سپس پارامترهای محاسبه شده که حاصل از برازش هر کدام از خانواده‌های مفصل‌های مختلف بر جفت-مفصل‌ها هستند، در رابطه‌های (۱۰) تا (۱۳) قرار داده می‌شوند. سپس هر خانواده مفصلی که موجب حداکثر شدن مقدار لوگ درست‌نمایی یا مینیمم شدن مقدار معیارهای اطلاعات آکائیکه<sup>۱</sup> (AIC) و اطلاعات بیزین<sup>۲</sup> (BIC) شود در رابطه‌های زیر گردد، بعنوان خانواده منتخب جهت برازش بر هر جفت-مفصل موجود در ساختارهای سی-واین و دی-واین تعیین می‌گردد. نحوه محاسبه لوگ درست-نمایی برای ساختار سی-واین چهار بعدی به شرح زیر می‌باشد (Brechmann et al., 2012):

$$(10)$$

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

محاسبه معیار لوگ درست‌نمایی در ساختار سی-واین مطابق با فرمول (۱۱) انجام می‌گیرد. همچنین معیارهای (AIC) و (BIC) هر دو ساختار سی-واین و دی-واین بترتیب با روابط (۱۲) و (۱۳) محاسبه می‌گردند:

$$(11)$$

$$L(u, h, \theta) = \sum_{t=1}^T \{ \log c_{12}(u_{1,t}, u_{2,t}; \theta_{12}) + \log c_{13}(u_{1,t}, u_{3,t}; \theta_{13}) + \log c_{14}(u_{1,t}, u_{4,t}; \theta_{14}) + \log c_{42|1}(h_{2|1,t}, h_{4|1,t}; \theta_{24|1}) + \log c_{34|1}(h_{3|1,t}, h_{4|1,t}; \theta_{34|1}) + \log c_{23|14}(h_{34|1,t}, h_{24|1,t}; \theta_{23|14}) \}$$

<sup>1</sup> Akaike information criterion

<sup>2</sup> Bayesian information criterion

<sup>3</sup> Empirical Copula

مرطوب و دوره خشک در سطح معنی داری ۵ درصد تشخیص داده شدند.

روابط توزیع‌های منتخب، مقدار پارامترهای تخمین زده شده آنها و P-value تست‌های کولموگروف-اسمیرنوف و کای اسکوتر در جدول ۲ ارائه شده‌اند. در نهایت پارامترهای محاسبه شده در توابع توزیع منتخب جایگذاری و توابع توزیع‌های حاشیه‌ای سری‌های زمانی چهار مشخصه مهم رویدادهای بارش بدست آمدند.

در این مطالعه، اتصال بین متغیرهای ساختارهای سی-واین و دی-واین بواسطه میزان معنی‌داری همبستگی بین متغیرها صورت گرفت. بطوریکه دو متغیر در درخت اول ساختارهای واین، در صورتی بهم متصل می‌گردند که همبستگی معنی‌داری بین آنها برقرار باشد. لازم بذکر است که برای محاسبه میزان وابستگی بین جفت-متغیرها از ضریب وابستگی تاو-کندال استفاده شد (Kendall, 1938). میزان ضریب تاو-کندال بین جفت-متغیرهای مشخصه‌های مهم بارش و میزان P-value تاو-کندال آنها در جدول ۳ ارائه شده است. مقادیر جدول ۳ نشان می‌دهد که میزان P-value تاو-کندال، بین همه متغیرها به جز متغیرهای M و D کمتر از ۰/۰۵ می‌باشد. بنابراین در بین جفت-متغیرهای مشخصه‌های مهم رویدادهای بارش، همبستگی بین متغیرهای M و D در سطح ۵ درصد معنی‌دار نمی‌باشد، در حالیکه همبستگی بین سایر جفت-متغیرها معنی‌دار می‌باشد.

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

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

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

در روابط فوق،  $C_t$  و  $C_e$  به ترتیب توزیع توأم و این (سی- واین و دی- واین) چهاربعدی محاسبه شده و مفصل تجربی چهاربعدی می‌باشند. همچنین n تعداد داده‌ها و i شماره داده‌ها می‌باشند.

## نتایج و بحث

در این مطالعه، پس از استخراج مشخصه‌های مهم رویدادهای بارش ایستگاه سریمونا واقع در ایتالیا از سال-های (۲۰۲۰-۱۹۹۵)، توزیع‌های احتمالاتی مختلف بر سری زمانی هر یک از چهار مشخصه رویدادهای بارش برآزش داده شدند. سپس توزیع‌های حاشیه‌ای مناسب انتخاب و پارامترهای توزیع‌های منتخب محاسبه گردیدند. بنابراین در این مطالعه ابتدا توزیع‌های پارامتری مختلفی جهت مدل‌سازی مشخصه‌های مهم بارش مورد آزمون قرار گرفتند، سپس با توجه به مقدار P-Value آزمون‌های کولموگروف-اسمیرنوف و کای اسکوتر، توزیع‌های بور و لوگ لوجستیک به ترتیب بیشترین تطابق را با سری زمانی متغیرهای عمق کل بارش و حداکثر شدت بارش داشتند و توزیع پارتو تعمیم یافته بعنوان مناسب‌ترین توزیع برای برآزش بر متغیرهای مدت زمان دوره‌های

جدول ۲. پارامترهای توابع توزیع حاشیه‌ای مشخصه‌های بارش به همراه مقدار P-value آزمون‌های کولموگروف-اسمیرنوف و کای اسکوتر

توابع توزیع-مشخصه رویداد بارش	فرمول	$\alpha$	$\beta$	$\gamma$	k	$\mu$	$\sigma$	$P_{KS}$	$P_{\chi^2}$
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



جدول ۳. ضریب تاو-کندال بین جفت-متغیرهای مشخصه‌های بارش ایستگاه سریمونا واقع در ایتالیا به‌مراه میزان P-value آنها

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

ساختار سی-واین حذف و تخمین پارامترها و در نهایت تعیین توابع توزیع توأم با ۶ ساختار باقیمانده صورت گرفت. در گام بعدی هدف تعیین ترکیبی از مناسب‌ترین خانواده‌های مفصل برای هریک از جفت-مفصل‌های موجود در درخت‌های اول، دوم و سوم ساختارهای سی-واین و دی-واین می‌باشد. لذا مطابق با روش ترتیبی که در قسمت قبل برای تخمین پارامترهای ساختارهای سی-واین و دی-واین شرح داده شد، خانواده‌های مختلف مفصل‌های ارشمیدسی (کلاتون، فرانک، جو و گامبل) و بیضوی (تی-استیودنت و گوسی) بر هر یک از جفت-مفصل‌های موجود در درخت‌های ساختارهای سی-واین و دی-واین برازش داده شدند. پس از تخمین پارامترهای خانواده‌های مفصل‌های مختلف ساختارهای واین و جایگذاری پارامترها در روابط لوگ درستنمایی، معیارهای اطلاعات آکائیکه و بیزین، بهترین خانواده‌های مفصل که

ایجاد ساختارهای مختلف سی-واین و دی-واین چهاربعدی با بکارگیری جایگشت‌های مختلف مشخصه-های مهم رویدادهای بارش (شامل عمق کل بارش (R)، ماکزیمم شدت بارش (M)، مدت زمان دوره مرطوب (L) و مدت زمان دوره خشک (D)) یکی از گام‌های اساسی در رسیدن به توزیع توأم چهاربعدی می‌باشد. مطابق با رابطه قانون جایگشت‌ها، ۴ متغیر می‌توانند ۱۲ ساختار دی-واین و سی-واین با جفت-مفصل‌های متفاوت تشکیل دهند. بطوریکه در هریک از ۱۲ ساختار سی-واین و یا دی-واین، درخت‌های اول، دوم و سوم شامل جفت-مفصل‌های متفاوتی می‌شوند که با ساختار دیگر یکسان نمی‌باشد. لیکن با توجه به اینکه در این مطالعه، جفت-متغیر M-D دارای همبستگی معنی‌داری نمی‌باشد، لذا ساختارهایی که شامل این جفت متغیر می‌باشند، از بین ۱۲ ساختار دی-واین و سی-واین حذف می‌گردند و به این ترتیب در این مطالعه ۶ ساختار دی-واین و ۶

جدول ۴. پارامترهای ساختارسی-واین و خانواده مفصل برازش شده با معیارهای لوگ درستنمایی، آکائیکه و بیزین و تاو-کندال بین

سری‌های تولید شده

BIC	AIC	Log-Lik	$\theta_{23 14}$	$\theta_{34 1}$	$\theta_{24 1}$	$\theta_{14}$	$\theta_{13}$	$\theta_{12}$	۱-۲-۳-۴ ساختار
-۸۹۵/۳	-۹۱۸/۷	۴۶۵/۳	F (۲/۵)	F (-۶/۰)	Ga (-۰/۲)	G (۲/۴)	G (۱/۸)	C (۰/۳)	R-D-L-M $\tau$
-۸۸۴/۴	-۹۰۷/۸	۴۵۹/۹	F (۱/۰)	F (-۶/۰)	F (۲/۲)	G (۱/۸)	G (۲/۴)	C (۰/۳)	R-D-M-L $\tau$
-۸۹۲/۳	-۹۱۵/۷	۴۶۳/۹	F (-۶/۲)	F (۲/۲)	Ga (-۰/۲)	C (۰/۳)	G (۱/۸)	G (۲/۴)	R-M-L-D $\tau$
-۸۵۳/۹	-۸۷۷/۳	۴۴۴/۶	G (۳/۵)	F (-۰/۷)	F (-۰/۱)	Ga (۰/۴)	G (۱/۸)	F (۱/۰)	L-M-R-D $\tau$
-۸۴۷/۶	-۸۷۱/۰	۴۴۱/۵	F (-۱/۲)	F (-۰/۱)	G (۳/۴)	F (۱/۰)	Ga (۰/۴)	G (۱/۸)	L-R-D-M $\tau$
-۸۴۹/۸	-۸۷۳/۲	۴۴۲/۶	F (۱/۱)	F (-۰/۷)	G (۳/۴)	G (۱/۸)	Ga (۰/۴)	F (۱/۰)	L-M-D-R $\tau$

جدول ۵. پارامترهای ساختار دی- واین و خانواده مفصل برازش داده شده به همراه معیارهای لوگ درستمنائی، آکائیکه و بیزین و میزان تاو-

کندال بین سری‌های تولید شده

BIC	AIC	Log-Lik	$\theta_{14 23}$	$\theta_{24 3}$	$\theta_{13 2}$	$\theta_{34}$	$\theta_{23}$	$\theta_{12}$	۱-۲-۳-۴
			درخت سوم	درخت دوم		درخت اول			ساختار
-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 $\tau$

ترتیب پارامترهای جفت-مفصل‌های  $C_{ML|R}$  و  $C_{MD|R}$  در درخت دوم و پارامتر  $\theta_{23|14}$  مربوط به جفت-مفصل  $C_{LD|RM}$  در درخت سوم می‌باشند. همانطور که جدول ۴ نشان می‌دهد، به ترتیب خانواده‌های مفصل کلایتون، گامبل و گامبل جهت برازش بر جفت-مفصل‌های  $C_{RD}$ ،  $C_{RM}$  و  $C_{RL}$  و خانواده‌های گوسی و فرانک جهت برازش بر جفت-مفصل‌های  $C_{ML|R}$  و  $C_{MD|R}$  و در نهایت مفصل فرانک جهت برازش بر جفت-مفصل  $C_{LD|RM}$  انتخاب شده است. لازم بذکر است که برای آزمون نکوئی برازش پارامترهای محاسبه شده از آزمون انتگرال احتمالی (PIT)<sup>۱</sup> استفاده شده است که نتایج مقادیر P-value آزمون مذکور در جدول ۶ آورده شده است. مطابق با جدول ۶ مشاهده می‌شود که مقادیر P-value کلیه ساختارها بزرگتر از ۵ درصد می‌باشند. بنابراین می‌توان نتیجه گرفت که پارامترهای تخمینی ساختارهای سی-واین و دی-واین در سطح ۵ درصد معنی‌دار می‌باشند. در این مطالعه برای اولین بار، ضرائب تاو-کندال بین تمامی توزیع‌های حاشیه‌ای در درخت‌های اول و همچنین توزیع‌های شرطی

منجر به ماکزیم شدن لوگ درست‌نمایی و مینیم شدن معیارهای آکائیکه و بیزین مفصل واین می‌گردند، انتخاب شدند. بدین ترتیب ترکیبی از خانواده‌های مفصل بهینه برای هر یک از ساختارهای واین حاصل گردیدند. در جدول‌های ۴ و ۵ ساختارهای منتخب و خانواده‌های مفصل بهینه آنها به ترتیب برای ساختارهای سی-واین و دی-واین نمایش داده شده‌اند.

علاوه بر این در جدول‌های ۴ و ۵ پارامترهای محاسبه شده خانواده‌های مفصل منتخب جهت برازش بر جفت-مفصل‌های هر یک از ساختارهای واین به همراه معیارهای لوگ درستمنائی، آکائیکه و بیزین و همچنین میزان ضرائب تاو-کندال بین همه توزیع‌های حاشیه‌ای و شرطی ارائه شده‌اند. در ساختارهای سی-واین اولین متغیر  $(S_1)$  بعنوان متغیر مرکزی در نظر گرفته شده است. بعنوان مثال مطابق با جدول ۴، در ساختار سی-واین R-D-L-M، پارامترهای  $\theta_{12}$ ،  $\theta_{13}$  و  $\theta_{14}$  به ترتیب پارامترهای مربوط به جفت-مفصل‌های  $C_{RD}$ ،  $C_{RL}$  و  $C_{RM}$  در درخت اول هستند. همچنین با توجه به قرارگیری متغیر  $h_{4|1}$  بعنوان متغیر مرکزی درخت دوم پارامترهای  $\theta_{24|1}$  و  $\theta_{34|1}$  به

<sup>۱</sup> Probability integral transform (PIT)

ساختارهای R-D-L-M, R-D-M-L و R-M-L-D برابر یا ۱/۱۶ و نسبت به سایر ساختارهای سی-واین ماکزیمم می‌باشند، اما تفاوت آنها در ترتیب قرارگیری متغیرهای M, D و L در ساختارها موجب شد تا تنها ساختار R-D-L-M دقت بیشتری را با توجه به معیارهای ارزیابی ارائه نماید که علت آن می‌تواند مرکزیت متغیر  $h_{M|R}$  در درخت دوم ساختار مذکور باشد. علاوه بر این، پس از جمع کردن میزان مقادیر تاو-کندال در درخت‌های اول، دوم و سوم کلیه ساختارهای سی-واین، مشاهده می‌شود که ساختار L-M-D-R با مجموع ضریب تاو-کندال ۱/۵۴ دارای بیشترین حاصل جمع مقادیر تاو-کندال در درخت‌های خود می‌باشد، اما با این وجود مقدار معیارهای ارزیابی، حاکی از دقت بالای آن نسبت به سایر ساختارها نمی‌باشند.

مطابق با جدول ۵ ساختار دی-واین M-R-D-L دارای ماکزیمم میزان لوگ درست‌نمایی معادل ۴۶۱/۸ و کمترین مقادیر  $AIC = -۹۱۱/۵$  و  $BIC = -۸۸۱/۱$  می‌باشد و بعنوان دقیق‌ترین ساختار دی-واین شناخته شده است. بررسی مطالعات اخیر (Aas et al., 2009; Vernieuwe

تولید شده در درخت‌های دوم و سوم ساختارهای واین مورد محاسبه قرار گرفته‌اند تا تأثیر کاربرد جفت-متغیرهای با میزان همبستگی‌های متفاوت در دستیابی به ساختار بهینه واین بررسی شود. بعنوان مثال در جدول ۴، با توجه به انتخاب متغیر  $s_1$  بعنوان متغیر مرکزی در درخت اول ساختار سی-واین، ضرائب تاو-کندال بین توزیع‌های حاشیه‌ای  $(u_1, u_2)$ ،  $(u_1, u_3)$  و  $(u_1, u_4)$  در درخت اول ساختار محاسبه و ارائه شده‌اند. همچنین ضرائب تاو-کندال بین سری‌های شرطی  $(h_{2|1}, h_{4|1})$  و  $(h_{3|1}, h_{4|1})$  در درخت دوم و در نهایت ضریب تاو-کندال بین سری‌های شرطی  $(h_{3|41}, h_{2|41})$  در درخت سوم ساختار سی-واین ارائه شده‌اند. مطابق با نتایج جدول ۴، ساختار سی-واین R-D-L-M دارای حداکثر میزان لوگ درست‌نمایی معادل ۴۶۵/۳ و کمترین میزان معیارهای اطلاعات آکائیکه و بیزین به ترتیب با مقادیر ۹۱۸/۷- و ۸۹۵/۳- می‌باشند. با بررسی مقادیر تاو-کندال در جدول ۴ می‌توان نتیجه گرفت که ترتیب قرارگیری متغیرها در درخت اول ساختار واین، تأثیر بسزایی در دستیابی به ساختاری دقیق دارد. بطوریکه علی‌رغم اینکه مجموع میزان همبستگی در جفت-متغیرهای درخت اول

جدول ۶. مقادیر معیارهای ارزیابی حاصل از مقایسه مفصل تجربی و توزیع توأم سی-واین و دی-واین چهاربعدی با مقادیر PIT ساختارها

PIT	P-value	R <sup>2</sup>	MAE	RMSE	ساختارها	نوع ساختار
۰/۸۰	۰/۳۵	۰/۹۹۸	۰/۰۲۲	۰/۰۲۹	R-D-L-M	C-vine
۰/۷۸	۰/۳۱	۰/۹۸۸	۰/۰۲۷	۰/۰۳۲	R-D-M-L	
۰/۷۷	۰/۳۳	۰/۹۹۳	۰/۰۲۳	۰/۰۳۰	R-M-L-D	
۰/۷۳	۰/۲۹	۰/۹۸۲	۰/۰۲۸	۰/۰۳۳	L-M-R-D	
۰/۶۸	۰/۲۸	۰/۹۸۰	۰/۰۳۰	۰/۰۳۹	L-R-D-M	
۰/۷۰	۰/۲۶	۰/۹۸۱	۰/۰۳۱	۰/۰۳۸	L-M-D-R	
۰/۶۵	۰/۱۲	۰/۹۶۳	۰/۰۳۳	۰/۰۴۶	R-M-L-D	D-vine
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	۰/۳۰	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	

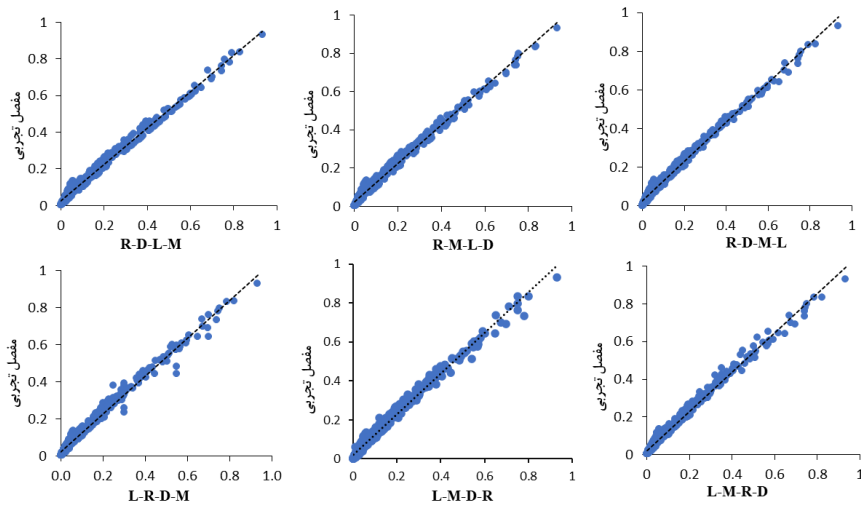
در ایتالیا مورد آزمون قرار دادند و برخلاف مطالعه حاضر تنها به محاسبه میزان همبستگی در جفت-متغیرهای درخت اول اکتفا نمودند و در نهایت نتیجه‌گیری کردند که استفاده از مشخصه‌های اصلی سیل با بیشترین همبستگی در درخت اول موجب رسیدن به یک ساختار بهینه نمی‌گردد و ساختاری که دارای جفت-متغیرهای با بیشترین همبستگی نبود، بعنوان دقیق‌ترین ساختار شناسایی گردید. در مطالعه حاضر، پس از جمع نمودن میزان ضرائب وابستگی تاو-کندال در کلیه جفت-متغیرهای درخت‌های هر ساختار دی-واین می‌توان دریافت که جمع مقادیر تاو-کندال‌های درخت‌های ساختار M-L-R-D نسبت به مجموع میزان تاو-کندال‌های درخت‌های سایر ساختارها بیشتر می‌باشد و همچنین تنها ساختاری است که تمامی ضرائب تاو-کندال درخت‌های آن دارای مقادیر مثبت می‌باشند، اما با این وجود مقدار لوگ درستنمایی آن از لوگ درستنمایی ساختار M-R-D-L کمتر می‌باشد. در ساختار سی-واین نیز علیرغم اینکه مجموع ضرائب وابستگی تاو-کندال ساختار L-M-D-R نسبت به سایر ساختارها بالاتر می‌باشد، اما نسبت به سایر ساختارهای سی-واین دارای دقت بیشتری نمی‌باشد. همچنین با توجه به اینکه هدف اصلی این مطالعه مقایسه دقت مفصل‌های درختی سی-واین و دی-واین می‌باشد، لذا مقایسه مقادیر معیارهای ارزیابی ساختارهای سی-واین و دی-واین در جداول ۴ و ۵ نشان می‌دهد که ساختار سی-واین R-D-L-M با مرکزیت متغیر R بعنوان ساختار منتخب سی-واین، دارای لوگ درستنمایی بیشتری نسبت به ساختار منتخب دی-واین یعنی ساختار M-R-D-L می‌باشد. در مرحله نهائی با اعمال پارامترهای تخمین زده شده جفت-مفصل‌های مناسب هر یک از ساختارهای منتخب سی-واین و دی-واین (پارامترهای تخمین زده شده در جداول ۴ و ۵) و همچنین سری‌های شرطی تولید شده توزیع‌های حاشیه‌ای تک متغیره در روابط چگالی زنجیره‌ای می‌توان توابع توزیع توأم

(et al., 2015) نشان می‌دهد که محققان برای دستیابی به ساختار بهینه واین از جفت-متغیرهای با بیشترین میزان همبستگی در درخت اول استفاده نموده‌اند و ساختارهای دیگر دی-واین که از جایگشت‌های مختلف متغیرها ایجاد شده است را مورد ارزیابی قرار ندادند. اما مقایسه مقادیر تاو-کندال در جفت-متغیرهای مربوط به درخت-های ساختارهای مختلف دی-واین (جدول ۵) نشان می‌دهد که در مجموع میزان ضرائب تاو-کندال در جفت-متغیرهای درخت اول ساختار M-R-D-L 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 دقیق-تر جدول ۵ می‌توان دریافت که برخلاف آنچه در مطالعه (Aas et al., 2009; Vernieuwe et al., 2015) عنوان شده است، لزوماً 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 مقادیر ضرائب تاو-کندال در ساختار M-R-L-D دارای بیشترین میزان در بین 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 ساختارهای دی-واین نمی‌باشد.

شفائی و همکاران (۱۳۹۵) از ساختار سی-واین جهت تعیین توابع توزیع توأم مشخصات مهم سیل رودخانه آجی چای استفاده کردند. پس از استخراج ساختارهای بهینه سی-واین، آنها نتیجه گرفتند که لزوماً ساختاری با بیشتری میزان همبستگی در جفت-متغیرهای درخت اول، دارای بیشترین دقت نمی‌باشد. علاوه بر این Shafaei و همکاران (۲۰۱۶) در مطالعه دیگری 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

چهاربعدی مشخصه‌های مهم رویدادهای بارش را برای هر یک از ساختارهای منتخب سی-واین و دی-واین محاسبه نمود. بررسی میزان دقت و تطابق تابع توزیع توأم چهاربعدی دی-واین مشخصه‌های مهم بارش بر مفصل تجربی چهاربعدی با کاربرد معیارهای ریشه میانگین مربعات خطاها، میانگین خطای مطلق و ضریب تبیین صورت گرفت. جدول ۶ مقادیر معیارهای ارزیابی حاصل از مقایسه مفصل تجربی چهاربعدی را با توزیع توأم دی-واین چهاربعدی مشخصه‌های مهم بارش نشان می‌دهد. نتایج ارائه شده در جدول ۶ نشان می‌دهد که ساختار سی-واین R-D-L-M دارای مینیمم مقدار معیارهای ارزیابی  $RMSE=0/029$  و  $MAE=0/022$  و همچنین ماکزیمم  $value = 0/35P$  و  $R^2=0/998$  در بین کلیه ساختارهای سی-واین و دی-واین می‌باشد که این نتیجه می‌تواند تایید کننده نتایج حاصل از جداول ۴ و ۵ باشد. دقت بالاتر ساختار سی-واین در تعیین توابع توزیع توأم چهاربعدی را می‌توان بعلت استفاده از یک متغیر مرکزی در ساختار سی-واین دانست. بطوریکه ارتباط یک متغیر مرکزی و اتصال آن به سایر متغیرها می‌تواند موجب برآزش بهتر و افزایش دقت ساختار می‌گردد. با توجه به مقادیر معیارهای جدول ۶، ساختار دی-واین M-R-D-L در مقایسه با سایر ساختارهای دی-واین، دارای بیشترین تطابق با مفصل چهاربعدی تجربی می‌باشد، چرا که در مقایسه با سایر ساختارهای دی-واین دارای کمترین  $RMSE=0/031$  و  $MAE=0/024$  و بیشترین  $P-value=0/30$  و  $R^2=0/991$  می‌باشد. مقادیر معیارهای ارزیابی در ساختارهای دی-واین M-R-L-D و M-L-R-D موید این مطلب است که لزوماً دارا بودن حداکثر همبستگی در جفت-متغیرهای درخت اول همانند ساختار M-R-L-D و یا حداکثر شدن مجموع همبستگی-های جفت-متغیرها در کل درخت‌ها همانند ساختار M-

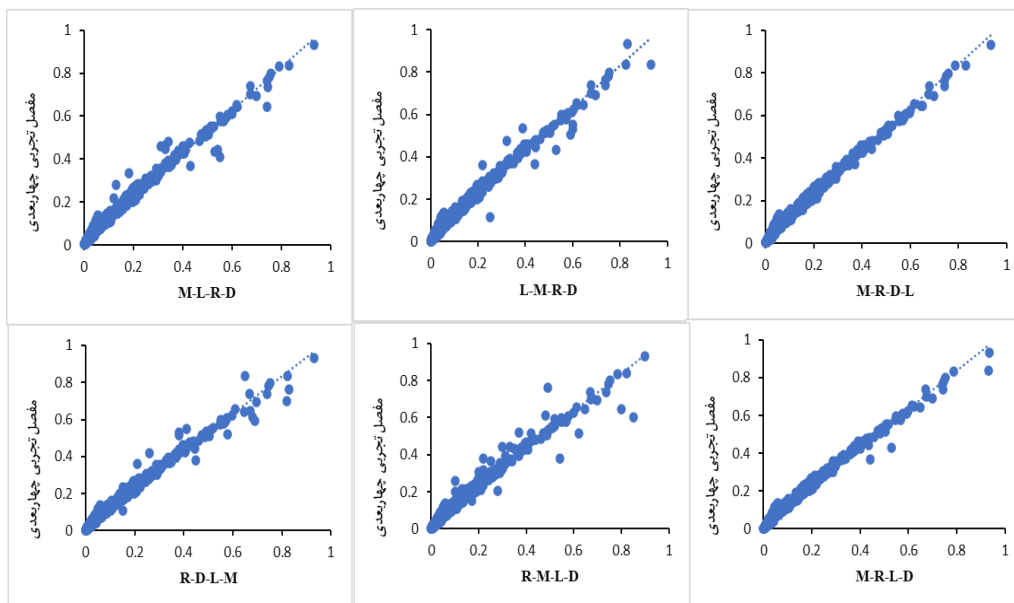
L-R-D، موجب دستیابی به دقیق‌ترین ساختار که کاملاً منطبق بر مفصل تجربی باشند، نمی‌گردد. بلکه ایجاد کلیه ترکیب‌های ممکن از متغیرها با استفاده از جایگشت آنها در ساختار واین مهمترین گام در دستیابی به ساختار بهینه واین می‌باشد. چرا که با توجه به نتایج این مطالعه و همچنین مطالعات قبلی در تحلیل فراوانی سیل (شفائی و همکاران، ۱۳۹۵) و (Shafaei et al., 2017) و تحلیل فراوانی بارش (Vernieuwe et al., 2015) که در خصوص تحلیل فراوانی چهاربعدی پدیده‌های هیدرولوژیکی گرفته شده است، در هیچ یک از مطالعات، ساختار بهینه واین مشخصات بارش و یا ساختار بهینه واین سیل یکسان نمی‌باشد و به نظر می‌رسد که ساختار بهینه سی-واین و دی-واین هر یک از پدیده‌های مذکور مرتبط با مشخصات فیزیکی پدیده هیدرولوژیکی هر منطقه می‌باشد، بطوریکه نوع اتصال متغیرهای پدیده‌های هیدرولوژیکی در ساختارهای بهینه واین مربوط به مناطق مختلف متفاوت می‌باشد. در شکل ۴ مفصل سی-واین چهار بعدی مشخصه‌های مهم بارش در مقابل مفصل تجربی چهاربعدی (نقاط آبی رنگ) به همراه خطی که نشان‌دهنده رابطه خطی بین دو مفصل مذکور می‌باشد (خط نقطه‌چین)، نمایش داده شده است. با توجه به شکل ۴ نمودار ساختارهای سی-واین R-D-L-M، R-M-L-D و R-D-M-L نسبت به سایر ساختارهای سی-واین دارای همپوشانی بیشتری با مفصل تجربی می‌باشند. چرا که پراکندگی نقاط آبی در ساختارهای مذکور نسبت به سایر ساختارها بسیار کمتر می‌باشد. لازم به ذکر است که دقت ساختارهای سی-واین R-D-L-M، R-M-L-D و R-D- نیز با توجه به معیارهای مختلف جداول ۴ و ۶ نیز تایید گردید. بررسی بیشتر شکل ۴ نشان می‌دهد که نقاط آبی ساختار L-R-D-M دارای انسجام کمتری در اطراف خط نقطه چین می‌باشند که این موضوع با توجه به معیارهای جداول ۴ و ۶ موید دقت پایین‌تر این ساختار در تعیین توابع توزیع توأم چهاربعدی می‌باشد.



شکل ۴. مقایسه مفصل‌های تجربی چهاربندی و توزیع توأم چهاربندی سی-واین

ارزیابی هر دو دقت بالای ساختار M-R-D-L نسبت به سایر ساختارهای دی-واین را تایید می‌نمایند. همچنین مشاهده می‌شود که پراکندگی نقاط آبی رنگ ساختار دی-واین R-M-L-D در اطراف خط نقطه‌چین نسبت به ساختارهای دیگر بالاتر می‌باشد که نشانگر دقت پایین مدل مذکور می‌باشد.

شکل ۵ مقایسه توابع توزیع توأم چهاربندی حاصل از مفصل دی-واین را با مفصل تجربی نشان می‌دهد. شکل ۵ حاکی از دقت بالای ساختار دی-واین M-R-D-L در مقایسه با سایر ساختارهای دی-واین، با توجه به پراکندگی کم آن در اطراف خط نقطه‌چین می‌باشد. همانطور که مشاهده می‌شود نتایج نمودار و معیارهای



شکل ۵. مقایسه مفصل‌های تجربی چهاربندی و توزیع توأم چهاربندی دی-واین

در نهایت با مقایسه نمودارهای ساختارهای سی-واین و دی-واین و همچنین مقایسه جداول ۴، ۵ و ۶ می‌توان نتیجه‌گیری کرد که با توجه به آزمون گرافیکی و معیارهای ارزیابی لوگ درست‌نمایی، آکائیکه، بیزین، ضریب تبیین، ریشه مربعات خطاها و میانگین قدرمطلق خطاها، بطور کلی ساختارهای سی-واین دارای دقت بالاتری نسبت به ساختارهای دی-واین می‌باشند. دقت بالاتر ساختار سی-واین در تعیین توابع توزیع توأم چهاربعدی می‌توان بعلت استفاده از یک متغیر مرکزی در ساختار سی-واین دانست. بطوریکه ارتباط یک متغیر مرکزی و اتصال آن به سایر متغیرها می‌تواند موجب افزایش دقت ساختار گردد.

### نتیجه‌گیری

هدف این مطالعه ایجاد توابع توزیع توأم چهاربعدی مشخصه‌های مهم رویدادهای بارش ایستگاه سریمونا در ایتالیا (شامل ماکزیمم شدت بارش، عمق کل بارش، مدت زمان دوره مرطوب و دوره خشک) با استفاده از ساختارهای سی-واین و دی-واین چهاربعدی و در نهایت مقایسه دقت ساختارهای سی-واین و دی-واین در تعیین توابع توزیع توأم می‌باشد. نتایج حاصل از این مطالعه نشان داد که استفاده از مشخصه‌های اصلی بارش با قوی‌ترین همبستگی در درخت اول لزوماً موجب ایجاد ساختار بهینه سی-واین و یا دی-واین نمی‌گردد. بطوریکه علی‌رغم اینکه در تجزیه درختی سی-واین، ساختار R-D-L-M با بیشترین همبستگی بعنوان ساختار بهینه انتخاب گردید، لیکن سایر ساختارهای سی-واین و یا دی-واین

با بیشترین همبستگی در درخت اول بعنوان ساختار بهینه شناخته نشدند. بر اساس معیارهای ارزیابی مختلف، استفاده از متغیر R (مقدار کل بارش) بعنوان متغیر مرکزی در ساختار سی-واین، موجب افزایش دقت توابع توزیع توأم چهاربعدی حاصل از تجزیه درختی ساختارهای سی-واین R-D-L-M و R-M-L-D نسبت به سایر ساختارهای سی-واین و همچنین ساختارهای دی-واین گردید. بررسی وابستگی بین متغیرها نشان داد که حداکثر شدن مجموع مقادیر تاو-کندال جفت-متغیرها در کل درخت‌های یک ساختار دی-واین و یا سی-واین نسبت به مجموع میزان تاو-کندال‌ها در درخت‌های سایر ساختارها موجب افزایش دقت ساختار نمی‌شود. با توجه به ماهیت متفاوت مشخصه‌های مهم پدیده‌های هیدرولوژیکی در مناطق مختلف، استفاده از روش جایگشت متغیرها در ساختارهای سی-واین و دی-واین به منظور دستیابی به بهترین نوع اتصال که موجب تشکیل دقیق‌ترین توزیع توأم واین گردد، حیاتی می‌باشد. در نهایت می‌توان نتیجه‌گیری نمود که استفاده از توزیع‌های توأم با ابعاد بالاتر از دو منجر به استفاده از مشخصه‌های هیدرولوژیکی بیشتر و در نتیجه تحلیل دقیق‌تر پدیده‌های هیدرولوژیکی نسبت به تحلیل‌های دویبعدی می‌گردد. بطور کلی تحلیل فراوانی چند متغیره بارش با توجه به مشخصه‌های کلیدی آن، کمک شایانی به پیش‌بینی بارش-های حدی که در نهایت منجر به سیل‌های با دوره بازگشت‌های بالا می‌شوند، می‌گردد.

### منابع مورد استفاده

- 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. (۱۰): 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، ۱: ۹۳-۱۱۲.

- شفائی، م.، ا. فاخری فرد، ی. دین پژوه، و ر. میرعباسی. ۱۳۹۵. مدلسازی تابع توزیع توأم چهاربعدی ویژگی های مهم سیل با استفاده از ساختار سی-واین. نشریه آبیاری و زهکشی ایران. ۳(۱۰): ۳۲۷-۳۳۷.
- عبداللهی، س.، ع.م. آخوندعلی و ر. میرعباسی. مدل سازی احتمالاتی و داده مبنای بارش- رواناب با بهره گیری از توابع چندمتغیره مفصل. هفدهمین کنفرانس هیدرولیک ایران ۱۳۹۷.
- 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.
- 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 Dzipire, 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.
- Dodangeh, E., Singh, V.P. 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. <https://doi.org/10.1007/s40996-019-00297-5>
- Ghosh, S. 2010. Modelling bivariate rainfall distribution and generating bivariate correlated rainfall data in neighbouring 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 de Paris*: Paris, 229-231
- 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., R. Cooke. 2006. *Uncertainty Analysis with High Dimensional Dependence Modelling*. Wiley, Chichester.
- Kuchment, L. S. and V. N. Demidov. 2013. On the Application of Copula Theory for Determination of Probabilistic Characteristics of Spring flood., *Russian Meteorology and Hydrology*, 38(4): 263-271.
- 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 and J. Zhang. 2020. Vine copula selection using mutual information for hydrological dependence modeling. *Environmental Research*. 186: 10-25.
- 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.
- 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.



- Sklar, A. 1959. Fonction de re'partition a' n dimensions et leurs marges, vol. 8. Publications de L'Institute de Statistique, Universite' de Paris: Paris, 229–231.
- 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.
- 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.
- 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.
- Yee, K.C., J. Suhaila, Y. Yusof, and F.H. Mean. 2014. Bivariate copula in fitting rainfall data. *AIP Conference Proceedings* 1605, 986.



ISSN 2251-7480

## Comparison of Performance of C-Vine and D-Vine Tree Copulas in Multivariate Analysis of Precipitation Characteristics

Maryam Shafaei<sup>1\*</sup> and Rasoul Mirabbasi Najafabadi<sup>2</sup>

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

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

\*Corresponding author email: m.shafaei65@gmail.com

Received: 25-05-2021

Accepted: 20-09-2021

### Abstract

In this study, the basic features of a tree vine copula such as the ability to decompose multivariate distributions into two-dimensional distributions, its flexibility in high-dimensional problems, and the use of conditional dependencies between variables have been considered. The purpose is to use C-Vine and D-Vine structures to determine the four-dimensional probabilistic distribution function of important characteristics of precipitation events of Cremona rain station located in Italy (26 years) including maximum precipitation intensity total precipitation depth, wet period duration and dry period. So that, a combination of the most suitable Archimedean and elliptical copulas families was identified to fit the pair-copulas of each of the C-Vine and D-Vine structures. The optimal combined distribution functions of C-Vine and D-Vine structures were also calculated using chain density functions and compared with the four-dimensional experimental copula of important precipitation characteristics. Finally, the accuracy of C-Vine and D-Vine tree structures in determining the combined distribution functions of important precipitation characteristics were compared. The results showed that the RDLM C-Vine structure has a minimum value of evaluation criteria  $RMSE = 0.029$  and  $MAE = 0.022$ , as well as a maximum of  $P\text{-value} = 0.35$  and  $R^2 = 0.998$  among all C-Vine and D-Vine structures. As a result, it has the highest accuracy for frequency analyzing the precipitation characteristics of Cremona station in Italy.

**Keywords:** Copulas, Joint distribution, Pair-copula, Precipitation, Vine