

## مدل سازی تابع توزیع توام چهار بعدی ویژگی های مهم سیل با استفاده از ساختار سی-واین

مریم شفاei<sup>1\*</sup>، احمد فاخری فرد<sup>2</sup>، یعقوب دین پژوه<sup>3</sup>، رسول میر عباسی نجف آبادی<sup>4</sup>

تاریخ دریافت: 1395/12/16 تاریخ پذیرش: 1395/5/27

### چکیده

تشخیص یک مفصل چند بعدی مناسب به منظور تعیین ساختار وابستگی داده های چند بعدی امر پیچیده ای می باشد چرا که مفصل های چند بعدی نظیر مفصل های بیضوی و ارشمیدسی فاقد انعطاف پذیری لازم جهت مدل سازی ساختار وابستگی بوده و دارای محدودیت های از قبیل محدودیت های پارامتر می باشد. برای غلبه بر این محدودیت ها، مفصل های واین جهت مدل سازی و شبیه سازی داده های هیدرولوژیکی چند بعدی ارائه شده اند. ساختمان سی-واین شامل آشناری از جفت-مفصل ها می باشد که بواسطه آن ها دارای چندین پارامتر بوده و بدین وسیله تشخیص و درک کامل الگوهای وابستگی پنهان و مختلط بین متغیرهای هیدرولوژیکی را فراهم می آورد. در این مقاله مفصل های سی-واین به منظور توصیف ساختار وابستگی بین ویژگی های سیل (حجم سیل (V)، دبی اوج سیل (P)، زمان پایه (B) و زمان اوج سیل (Tp)) پیشنهاد گردید و از مناسب ترین ساختارهای سی-واین جهت مدل سازی توزیع توام چهار متغیره بهره گرفته شد. در این روش ابتدا با استفاده از جایگشت متغیرها ساختارهای سی-واین مختلف حاصل گردید و سپس مناسب ترین جفت-مفصل ها از خانواده های مفصل بیضوی و ارشمیدسی مطابق با معیارهای اطلاعات بیزین (BIC) و آکائیکه (AIC) انتخاب گردید. در نهایت مفصل های سی-واین چهار بعدی حاصل یا مفصل های تجربی چهار بعدی مقایسه گردیدند و با توجه به معیارهای ضریب تبیین، ریشه میانگین مربعات خطا و میانگین قدر مطلق خطاها و روش گرافیکی، ساختار (B-P-V-Tp) دقیق ترین تابع توزیع توام چهار بعدی تشخیص داده شد. بنابراین در نهایت نتیجه گیری شد که انتخاب متغیر زمان پایه B بعنوان متغیر مرکزی انتخاب مناسبی بوده و این متغیر نقش کنترل کننده متغیرهای حجم کل، زمان اوج و دبی اوج سیل را دارا می باشد.

واژه های کلیدی: حجم سیل، جفت-مفصل، مفصل ارشمیدسی مفصل بیضوی، معیار اطلاعات آکائیکه

### مقدمه<sup>1</sup>

اولین مزیت می توان به مدل سازی جداگانه توزیع های حاشیه ای و توام در مفصل ها اشاره نمود و در ثانی مفصل ها لزوماً نیازی به استفاده از توزیع های حاشیه ای بیضوی نمی باشد و ثالثاً مفصل ها قادر به توصیف ساختار وابستگی دمی توزیع های توام می باشند. در زمینه مدل سازی ویژگی های رویدادهای سیل مانند حجم کل، دبی اوج، زمان پایه و زمان اوج سیل مفصل ها توانایی خوبی را در مدل سازی ساختار وابستگی چند بعدی با ابعاد بالاتر از دو ارائه می دهند. کاربرد توابع مفصل چند بعدی (ابعاد بالاتر از دو) در مهندسی منابع آب محدود به چند مطالعه انگشت شمار می باشد که در زمینه مدل سازی چند متغیره سری های زمانی هیدرولوژی می توان تنها به مدل سازی سه متغیره حجم کل، زمان پایه و دبی پیک سیل (Serinaldi and Grimaldi., 2007; Genest et al., 2007) اشاره کرد. کارگیری متغیرهای شدت متوسط بارش، مدت زمان بارش و مدت زمان دوره خشک رویدادهای بارش به (Salvadori and De Michele., 2006)، تحلیل فرکانس سه متغیره عمق، مدت و شدت بارش به (Kao and Govindaraju., 2008) و برای مدل سازی

اخیراً مفصل ها به ابزاری مفید جهت توصیف ساختار وابستگی بین متغیرهای تصادفی تبدیل شده است و تاکنون مدل سازی بر پایه ی مفصل ها به طور گسترده ای در زمینه تحلیل فرکانس خشک-سالی (Kao and Govindaraju, 2010; Mirabbasi et al., 2012) سیل (Grimaldi and Serinaldi, 2006; Zhang and Singh., 2006; Zhang and Singh., 2007a; Karmakar and Simonovic., 2009) و بارش (De Michele and Salvadori., 2003; Singh and Zhang., 2007; Zhang and Singh., 2007b) به کار گرفته شده است. مفصل ها بر پایه نظریه اسکالر (Sklar., 1959) بوده و شامل چندین ویژگی مهم می باشد. به عنوان

1- دانشجوی دکتری مهندسی منابع آب، گروه مهندسی آب، دانشگاه تبریز

2- استاد گروه مهندسی آب، دانشگاه تبریز

3- دانشیار گروه مهندسی آب، دانشگاه تبریز

4- استادیار گروه مهندسی آب، دانشگاه شهر کرد

\* - نویسنده مسئول: (Email: m.shafaei@tabrizu.ac.ir)

شامل تعداد زیادی از تجزیه‌های جفت-مفصل<sup>6</sup> می‌باشد. به‌طور کلی دو نوع مشخص از واین نرمال یعنی واین متعارفی<sup>7</sup> (C-vine) و دی-واین (D-vine) وجود دارد، که هر کدام از واین‌ها روش مشخصی را برای تجزیه تابع چگالی ارائه می‌دهند (Kurowicka and Cooke., 2007). در بخش‌های بعدی مشاهده خواهد شد که جهت تعیین تابع چگالی توزیع توام این تابع به حاصل ضربی از چگالی‌های منفرد و شرطی از متغیرها تجزیه خواهد شد (تجزیه تابع چگالی) که در نهایت با استفاده از جایگذاری متغیرهای مختلف در مکان‌های مختلف ساختار سی-واین تجزیه‌های مختلفی ارائه خواهد شد. مفصل واین دارای ویژگی سلسله مراتبی می‌باشد. به عبارت دیگر مفصل واین چند بعدی (با تعداد متغیرهای بالاتر از دو) را به تعدادی از مفصل‌های دو متغیره تجزیه نموده و بدین ترتیب مفصل‌های چند بعدی را محاسبه می‌نماید. در سال‌های اخیر مفصل‌های واین به ابزاری محبوب جهت مدل‌سازی ساختار چند بعدی هیدرولوژی تبدیل شده‌اند. به طوری که جی یاسی (Gyasi-Agyei Melching., 2012) ساختار سی-واین را به همراه سه متغیر حداکثر عمق، مدت زمان دوره مرطوب و عمق کل رویدادهای بارش جهت شبیه‌سازی ویژگی‌های مذکور به کار گرفتند. گرالر و همکاران با استفاده از ساختار دی-واین و سه متغیر دبی اوج، حجم سیل و زمان پایه سیل مفصل سه بعدی را مدل‌سازی نموده و سپس مفصل سه بعدی حاصل جهت تعیین دوره بازگشت‌های توام متغیرهای سیل به کار گرفته شدند (Gräler et al., 2013). در این مطالعه برای اولین بار، ساختار چهار بعدی مهم‌ترین ویژگی‌های سیل ایستگاه هیدرومتری و نیار در رودخانه آجی چای بر اساس ساختار سی-واین مدل‌سازی گردید و سپس از دقیق‌ترین ساختارهای سی-واین جهت تعیین توزیع توام چهار بعدی حجم کل، دبی اوج، زمان اوج و زمان پایه سیل استفاده شد.

## مواد و روش‌ها

### منطقه مورد مطالعه

رودخانه آجی چای در آذربایجان شرقی از مهم‌ترین رودخانه‌های حوضه آبریز دریاچه ارومیه می‌باشد که در جهت عمومی شرقی-غربی جریان دارد و چهار ایستگاه هیدرومتری بر روی رودخانه آجی چای نصب و تجهیز شده که توسط سازمان آب منطقه‌ای آذربایجان شرقی مورد بهره‌برداری قرار می‌گیرد. در این مطالعه ایستگاه هیدرومتری و نیار به دلیل داشتن آمار کامل، طولانی مدت و دقیق‌تر به عنوان ایستگاه مبنا تعیین و آمار جریان روزانه رودخانه آجی چای در آن ایستگاه مورد استفاده قرار گرفته است. ایستگاه و نیار دارای

خشک‌سالی سه بعدی با استفاده از متغیرهای عمق خشکی، مدت خشکی و فاصله بین خشکی‌های متوالی به (Song and Singh., 2010) اشاره نمود. کاربرد مفصل چهار بعدی در هیدرولوژی بسیار محدود می‌باشد به‌طوری‌که در این زمینه تنها دی میچل و همکاران (De Michele et al., 2007) مفصل چهار بعدی پدیده توفان دریا را تخمین زدند. سرینالدی و گرمی‌مالدی از یک مفصل ارشمیدسی نامتقارن تو در تو<sup>1</sup> برای تحلیل فراوانی سه متغیره استفاده نمودند (Serinaldi and Grimaldi., 2007). با این حال تعمیم مستقیم مفصل‌های ارشمیدسی به ابعاد بالاتر از دو، دارای پیچیدگی‌هایی می‌باشد چرا که محدودیت‌های شدیدی را به وابستگی‌های متقابل دو به دو که در بر می‌گیرد، تحمیل می‌نماید (Genest et al., 2007). برای مثال در سطح سه متغیره، زمانی که یکی از توزیع‌های حاشیه‌ای برابر واحد باشد، مفصل سه متغیره به دو متغیره شامل دو حاشیه باقیمانده تبدیل می‌شود. در عمل فراهم نمودن این شرط بسیار مشکل می‌باشد و خود علت عدم انعطاف مفصل‌های متقارن می‌باشد. همچنین استفاده از مفصل‌های ارشمیدسی نامتقارن نیز خود دارای شرایطی مانند وابستگی‌های مثبت بین متغیرهای درونی و کوچک‌تر بودن وابستگی بیرونی نسبت به وابستگی درونی می‌باشد که خود موجب محدودیت استفاده از این مفصل‌ها در ابعاد بالا می‌باشد. هم‌چنین جنست و همکاران مفصل‌های متا-الپتیک<sup>2</sup> (فوق بیضوی) را برای مدل‌سازی سه متغیره حجم کل، زمان پایه و دبی پیک سیل معرفی نمودند (Genest et al., 2007). با وجود انعطاف‌پذیری این مفصل‌ها، محاسبه ریشه مربع ماتریس کوواریانس با تجزیه کولسکی<sup>3</sup> اجازه حفظ چندک‌های متغیرهای مهم که در تفکیک بارش<sup>4</sup> مهم می‌باشد را نمی‌دهد و به‌علاوه در مواردی که ماتریس همبستگی دارای اعداد منفی می‌باشد تصحیحاتی جهت تضمین مثبت بودن ماتریس مورد نیاز است. به‌طور 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 این مورد محدودیتی نشان می‌دهند، اما ساختار مفصل واین (Kurowicka and Cooke., 2007) با تجزیه توزیع‌های چند متغیره به توزیع‌های توام با بعد دو، انعطاف‌پذیری لازم در مسائل با ابعاد بالا را فراهم آورده و فاقد ضعف‌های مفصل-های مذکور می‌باشد. (Bedford and Cooke, 2001., 2002) یک مدل گرافیکی با عنوان واین نرمال<sup>5</sup> را برای توزیع‌های با ابعاد بالا ارائه دادند. واین نرمال دارای یک ساختار کلی و عمومی می‌باشد و

- 1- Nested Asymmetric Archimedean Copula
- 2- Meta-elliptic
- 3- Cholesky
- 4- Rainfall disaggregation
- 5- Regular vine

6- Pair-copula

7- Canonical vine

کنند (Aas et al., 2009). با استفاده از دو روش اتصال متفاوت جفت-مفصل‌ها دو نوع مشخص از واین نرمال یعنی واین متعارفی<sup>4</sup> (C-vine) و دی-واین (D-vine) معرفی شده‌اند، که هر کدام از واین‌ها روش مشخصی را برای تجزیه تابع چگالی ارائه می‌دهند. تشخیص انواع واین‌ها از شکل گرافیکی آن‌ها قابل تشخیص است، چرا که واین‌ها مجموعه‌ای از درخت‌های تو در تو می‌باشند که از لحاظ ترکیب با هم تفاوت دارند. هریک از ساختارهای واین از مجموعه‌ای از گره‌ها<sup>5</sup> (دایره) و لبه‌ها<sup>6</sup> (خط) تشکیل شده‌اند. در شکل 1 در هر یک از درخت‌ها هر گره توسط یک دایره و هر لبه توسط یک خط نمایش داده شده هر گره مطابق با یک متغیر و هر لبه مطابق با یک چگالی جفت-مفصل می‌باشد. اگر  $n$  تعداد متغیرها باشد کل تجزیه‌ها توسط  $n(n-1)/2$  لبه (چگالی حاشیه‌ای) هر متغیر تعیین می‌گردد. گره‌های درخت  $T_j$  (در اینجا نشان‌دهنده شماره درخت می‌باشد). برای مشخص نمودن برچسب لبه‌های درخت  $T_{j+1}$  ضروری هستند بنابراین لبه‌های درخت  $T_j$  گره‌های درخت  $T_{j+1}$  را تشکیل داده و شرط اتصال دو لبه در درخت  $T_{j+1}$  اشتراک دو لبه مورد نظر در یک گره در درخت  $T_j$  می‌باشد. باید توجه داشت که ساختمان درختی در تجزیه جفت-مفصل ضروری نمی‌باشد اما در تشخیص تجزیه‌های جفت-مفصل‌های مختلف مؤثر هستند (Aas et al., 2009). در این مطالعه از ساختار سی-واین استفاده شد. در شکل 1 روش تجزیه تابع چگالی توزیع چهار بعدی توسط سی-واین همراه با درخت‌های مربوطه نمایش داده شده‌اند. در شکل 1 وجه تمایز درخت اول  $T_1$  و درخت‌های  $T_2$  و  $T_3$  در غیر شرطی بودن متغیرها و یا همان گره‌ها در درخت اول می‌باشد، در صورتی که گره‌های درخت‌های دیگر شرطی و یا به عبارتی جفت-مفصل‌ها می‌باشند. در ساختارهای سی-واین یک متغیر که نقش کنترل‌کننده را بازی می‌نماید و به عنوان متغیر مرکزی در مرکز ساختار قرار می‌گیرد به طوری که دیگر متغیرها به این متغیر مستقیم متصل می‌گردند. ساختار سی-واین در شکل 1 از چهار متغیر (گره)  $x_1, x_2, x_3, x_4$  تشکیل شده است که به ترتیب در مکان‌های 1، 2، 3 و 4 سی-واین در درخت اول قرار گرفته‌اند. در شکل 1 بعنوان مثال از متغیر  $x_1$  بعنوان متغیر مرکزی استفاده شده است. اما در این مطالعه همان‌طور که ذکر شد سه متغیر حجم کل، دبی پایه و دبی اوج بعنوان متغیر مرکزی بکار برده شده‌اند. در درخت  $T_2$  و  $T_3$  از گره‌های شرطی  $y$  که به ترتیب نتیجه محاسبه جفت-مفصل‌های متغیرهای درخت  $T_1$  و  $T_2$  می‌باشد استفاده شد. رویکرد محاسبه جفت-مفصل‌ها در بخش بعدی توضیح داده خواهد شد.

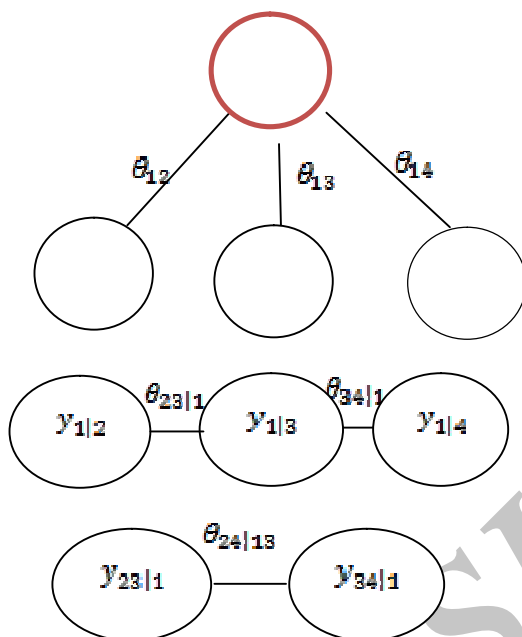
موقعیت جغرافیایی 46 درجه و 24 دقیقه طول شرقی و 38 درجه و 7 دقیقه عرض شمالی بوده و مساحت حوضه آبریز آجی‌چای در محل ایستگاه هیدرومتری ونیار برابر 5/528 کیلومترمربع است. ارتفاع ایستگاه هیدرومتری ونیار از سطح دریا برابر 3882 متر است. طول و شیب متوسط آبراهه اصلی حوضه آبریز آجی‌چای تا ایستگاه ونیار به ترتیب برابر 2/227 کیلومتر و 3/13 درصد می‌باشد. برای سال‌های آماری 1340 الی 1390 رودخانه آجی‌چای در ایستگاه ونیار دارای میانگین جریان سالانه 72/13 مترمکعب بر ثانیه بوده و کم‌ترین و بیش‌ترین جریان سالانه به ترتیب 23/2 و 9/42 مترمکعب بر ثانیه است. در ضمن ضریب تغییرات جریان سالانه رودخانه 51/0 است. در این مطالعه از جریان‌های روزانه رودخانه آجی‌چای در ایستگاه ونیار به مدت 34 سال، از 11 دی سال 1354 تا 29 اسفند 1387 استفاده شده است. حداکثر، متوسط و حداقل جریان روزانه در ایستگاه مذکور به ترتیب 342، 37/10 و مترمکعب بر ثانیه می‌باشد. در این مطالعه هدف تعیین توزیع چهار بعدی ویژگی‌های سیل (حجم کل، زمان اوج، دبی اوج و زمان پایه) می‌باشد. در این تحقیق از روش حداقلی موضعی<sup>1</sup> توسط نرم افزار هیدرو آفیس<sup>2</sup> - به منظور جداسازی دبی پایه از رویدادهای سیل مورد استفاده قرار گرفت. در روش حداقلی موضعی دبی هر روزه را کنترل کرده تا کم‌ترین دبی را در نصف منهای یک فاصله زمانی  $[0.5 * (2n-1)]$  روز قبل و بعد از هر روز را در نظر بگیرد. اگر نتایج رضایت بخش بود مقدار دبی برای آن روز یک حداقلی موضعی در نظر گرفته شده و به مقادیر حداقلی موضعی‌های دیگر با استفاده از درون‌یابی خطی مرتبط می‌شود (Sloto and Crouse, 1996). حداکثر جریان روزانه در طول رویداد سیل به عنوان دبی اوج سیل، تعداد روزهای رویداد سیل به عنوان زمان پایه و حجم تجمعی جریان در طول سیل بعنوان حجم کل سیل در نظر گرفته شدند.

### مفصل‌های واین

بد فورد و کوک یک مدل گرافیکی با عنوان واین نرمال<sup>3</sup> را برای توزیع‌های با ابعاد بالا ارائه دادند (Bedford and Cooke, 2001, 2002). این نرمال دارای یک ساختار کلی و عمومی می‌باشد که رویکرد این ساختار استفاده از اتصال‌های متفاوت متغیرها در درخت اول به یکدیگر و در نهایت تشکیل جفت-مفصل و در نهایت اتصال جفت-مفصل‌ها در درخت‌های بعدی می‌باشد. منظور از جفت-مفصل-ها در ساختار واین همان مفصل‌های دومتغیره شرطی می‌باشند که در ساختار واین نوع اتصال متغیرها و جفت-مفصل‌های مختلف نقش اصلی را جهت ساختن توزیع‌های حاشیه‌ای توام چند بعدی بازی می‌-

4- Canonical vine  
5- Nodes  
6- Edges

1- Local Minimum Method (LMM)  
2- Hydro Office  
3- Regular vine



شکل 1- ساختارهای مفصل سی- واین چهار بعدی

می‌باشد.  $\theta$  پارامتر تابع مفصل می‌باشد که مقدار وابستگی بین توزیع‌های حاشیه‌ای را نشان می‌دهد. تئوری معکوس نظریه اسکالر بصورت رابطه 3 تعریف می‌شود:

$$C_{12...n}(u_1, u_2, \dots, u_n) = F_{12...n} \{ F_1^{-1}(u_1), F_2^{-1}(u_2), \dots, F_n^{-1}(u_n) \} \quad (3)$$

در رابطه بالا توابع توزیع حاشیه‌ای متغیر  $n$ ام و  $F_i^{-1}(\cdot)$  تابع چندک  $F_i(\cdot)$  می‌باشند. خانواده‌های توابع مفصلی مختلفی برای تعیین ساختمان همبستگی توابع موجود هستند که با انتخاب مناسب آن‌ها می‌توان ساختار وابستگی چند متغیره را به خوبی مدل‌سازی نمود. مطابق با قانون زنجیره‌ای، تابع چگالی چهار بعدی توم بصورت رابطه 4 تعریف می‌شود:

$$f(x_1, x_2, x_3, x_4) = f(x_1) \cdot f(x_2) \cdot f(x_3) \cdot f(x_4) \cdot c_{12} \{ F(x_1), F(x_2) \} \cdot c_{13} \{ F(x_1), F(x_3) \} \cdot c_{14} \{ F(x_1), F(x_4) \} \cdot c_{23|1} \{ F(x_2|x_1), F(x_3|x_1) \} \cdot c_{24|1} \{ F(x_2|x_1), F(x_4|x_1) \} \cdot c_{34|12} \{ F(x_3|x_1, x_2), F(x_4|x_1, x_2) \} \quad (4)$$

مطابق با Joe (1996) توزیع‌های شرطی حاشیه‌ای معادله 4 با استفاده از روابط 5-9 حاصل می‌گردند:

$$y_{12} = F(x_2|x_1) = \frac{\delta C_{22}(F(x_2), F(x_2); \theta_{22})}{\delta F(x_2)} \quad (5)$$

$$y_{13} = F(x_3|x_1) = \frac{\delta C_{23}(F(x_2), F(x_3); \theta_{23})}{\delta F(x_2)} \quad (6)$$

$$y_{14} = F(x_4|x_1) = \frac{\delta C_{24}(F(x_2), F(x_4); \theta_{24})}{\delta F(x_2)} \quad (7)$$

به عبارت دیگر، گره‌های شرطی از اتصال دو متغیر متصل به-یکدیگر در درخت قبلی حاصل می‌شوند. توابع توزیع شرطی یا گره‌ها در درخت دوم و سوم به ترتیب با استفاده از مشتق جزئی مفصل‌های (لبه‌های) درخت اول و دوم صورت می‌پذیرد. در این مطالعه با استفاده از جایگشت (حالت‌های مختلف قرارگیری هر یک از متغیرها در گره‌های  $x_1, x_2, x_3, x_4$  ساختار سی-واین به منظور ایجاد ساختارهای مختلف سی-واین متغیرها تمامی ساختارهای ممکن (ساختارهای گوناگون حاصل از جایگشت متغیرها در ساختار چهاربعدی سی-واین) توسعه یافتند.

### توابع مفصل

مدل‌سازی تابع مفصل بر اساس تئوری اسکالر پایه گذاری شده است. مطابق با نظریه اسکالر تابع توزیع تجمعی  $n$  متغیره پیوسته  $F_{12...n}(x)$  شامل متغیرهای تصادفی  $X_i$  با توزیع حاشیه پیوست  $u_i = F_i(x_i)$  بصورت توابع 1 و 2 تعریف می‌شوند:

$$F_{12...n}(x_1, x_2, \dots, x_n) = C_{12...n} \{ F_1(x_1), F_2(x_2), \dots, F_n(x_n), \theta \} = C_{12...n}(u_1, u_2, \dots, u_n, \theta) \quad (1)$$

$$C_{12...n}(u_1, u_2, \dots, u_n, \theta) = P(X_1 \leq x_1, X_2 \leq x_2, \dots, X_n \leq x_n) \quad (2)$$

$C_{12...n}$  یک تابع مفصل منحصر بفرد  $n$  بعدی می‌باشد که توزیع‌های یکنواخت حاشیه‌ای تک متغیره را به توزیع تجمعی  $n$  بعدی توأم نگاشت می‌کند؛ به طوری که  $C$  نگاشت  $[0,1]^n \rightarrow [0,1]$

شده توسط رابطه  $h$  در زمان  $t$  می باشند. توسط روش های عددی می توان لوگ درست نمائی را در رابطه بالا حداکثری نمود و در نهایت پارامترها را تخمین زد و بدین ترتیب بهترین تابع مفصل را از میان خانواده های مفصل های ارشمیدسی خانواده های (فرانک<sup>1</sup>، جو<sup>2</sup>، گامبل<sup>3</sup> و کلاتون<sup>4</sup>) و بیضوی<sup>5</sup> (تی-استیودنت<sup>6</sup>، گوسین<sup>7</sup>) خواهد خواهد شد.

**انتخاب بهترین ساختارهای سی-واین جهت تعیین تابع توزیع توام چهار بعدی**

به منظور محاسبه توزیع توام ویژگی های سیل (زمان پایه، زمان اوج، حجم کل، دبی اوج) یک ساختار سی-واین چهار بعدی بر توزیع های حاشیه ای ویژگی های سیل برازش داده شد. سه مرحله زیر جهت دستیابی به ساختار سی-واین مناسب اجرا گردید: (1) محاسبه وابستگی تاو-کندال بین همه جفت-متغیرها جهت اتصال متغیرهای همبسته در ساختار سی-واین و همچنین حذف جفت-متغیرهای غیر همبسته. ضریب تاو کندال بصورت اختلاف احتمال انطباق و احتمال عدم انطباق تعریف می شود. اگر فرض کنیم  $(X_1, Y_1), \dots, (X_n, Y_n)$  جفت متغیر تصادفی باشند دو جفت نمونه  $(X_i, Y_i)$  و  $(X_j, Y_j)$  سازگارند اگر  $[(X_j - X_i)(Y_j - Y_i)] > 0$  در غیر اینصورت ناسازگارند. در نهایت ضریب تاو-کندال به صورت  $\tau = (c - d) / \binom{n}{2}$  تعریف می گردد. که  $c$  و  $d$  به ترتیب تعداد جفت های سازگار و ناسازگار و  $n$  تعداد جفت ها می باشد. (2) جایگشت متغیرها (ویژگی های سیل) در ساختار سی-واین به منظور بدست آوردن ساختارهای سی-واین مختلف. بطوریکه در هر جایگشت هر یک از متغیرهای سیل می-توانند در مکان های متغیر  $X_1, X_2, X_3$  و  $X_4$  در ساختار سی-واین جایگزین شوند و در نهایت منجر به ایجاد ساختارهای مختلف گردند. (3) استفاده از خانواده های ارشمیدسی و بیضوی مختلف جهت برازش بر جفت-مفصل های موجود در ساختارهای ایجاد شده سی-واین حاصل از مرحله قبل. عبارت دیگر بر هر یک از جفت-مفصل های موجود در درخت های اول، دوم و سوم ساختار سی-واین همه ی خانواده های مفصل ارشمیدسی و بیضوی برازش می شوند و مورد آزمون قرا می گیرند تا در نهایت بهترین نوع مفصل جهت برازش بر جفت-مفصل های مختلف انتخاب شوند. (4) تخمین پارامترهای جفت-مفصل های خانواده های مختلف مطابق با روش شرح داده

$$F(x_4 | x_1, x_2) = \frac{\delta C_{42|1}(F(x_4|x_2), F(x_2|x_2), \theta_{42|1})}{\delta F(x_2|x_2)} \quad (8)$$

$$F(x_3 | x_1, x_2) = \frac{\delta C_{32|1}(F(x_3|x_2), F(x_2|x_2), \theta_{32|1})}{\delta F(x_2|x_2)} \quad (9)$$

آس و همکاران (2009) روابط توزیع تجمعی شرطی را به صورت رابطه 10 تعریف نمودند:

$$h(u_i, u_j, \theta_{ij}) = F(x_j | x_i) \quad (10)$$

در روابط بالا برای بدست آوردن  $h$ ، از مفصل نسبت به متغیر دوم مشتق گرفته می شود.  $\theta_{ij}$  پارامتر مفصل در توزیع توأم بین  $F(x_i) = u_i$  و  $F(x_j) = u_j$  می باشد و  $\theta_{ij}$  پارامتر مفصل بین دو متغیر  $F(x_j | x_i)$  و  $F(x_i | x_j)$  می باشد. تابع  $h$  یک تابع صعودی و شرطی می باشد که در فاصله  $[0, 1]$  قرار می گیرد.

**تخمین پارامترهای ساختار سی-واین با استفاده از روش ترتیبی**

1- برازش یک مفصل دو بعدی بر  $u_1$  و  $u_2$  جهت تخمین پارامتر  $\theta_{12}$ .  
 2- برازش یک مفصل دو بعدی بر  $u_1$  و  $u_3$  جهت تخمین پارامتر  $\theta_{13}$ .  
 3- برازش یک مفصل دو بعدی بر  $u_1$  و  $u_4$  جهت تخمین پارامتر  $\theta_{14}$ .  
 4- تولید سری  $y_{12}$  با استفاده از مفصل برازش داده شده بر  $u_1$  و  $u_2$  با استفاده از رابطه 4  $h(u_1, u_2, \theta_{12}) = y_{12}$ .  
 5- تولید سری  $y_{13}$  با استفاده از مفصل برازش داده شده بر  $u_1$  و  $u_3$  با استفاده از رابطه 6  $h(u_1, u_3, \theta_{13}) = y_{13}$ .  
 6- تولید سری  $y_{14}$  با استفاده از مفصل برازش داده شده بر  $u_1$  و  $u_4$  با استفاده از رابطه 7  $h(u_1, u_4, \theta_{14}) = y_{14}$ .  
 7- برازش یک مفصل دو بعدی بر  $y_{12}$  و  $y_{13}$  جهت محاسبه پارامتر  $\theta_{23|1}$ .  
 8- برازش یک مفصل دو بعدی بر  $y_{12}$  و  $y_{14}$  جهت محاسبه پارامتر  $\theta_{24|1}$ .  
 9- تولید سری  $y_{23|1}$  با استفاده از مفصل برازش داده شده بر  $y_{12}$  و  $y_{13}$  با رابطه  $h(y_{12}, y_{13}, \theta_{23|1}) = y_{23|1}$ .

$$y_{23|1} = \frac{\delta C_{23|1}(F(y_{23|1}), F(y_{23|1}), \theta_{23|1})}{\delta F(y_{23|1})} \quad (11)$$

10- تولید سری  $y_{43|1}$  با استفاده از مفصل برازش داده شده بر  $y_{14}$  و  $y_{13}$  با استفاده از رابطه  $h(y_{14}, y_{13}, \theta_{43|1}) = y_{43|1}$ .

$$y_{34|1} = \frac{\delta C_{34|1}(F(y_{34|1}), F(y_{34|1}), \theta_{34|1})}{\delta F(y_{34|1})} \quad (12)$$

11- برازش یک مفصل دو بعدی بر  $y_{43|1}$  و  $y_{23|1}$  جهت تخمین پارامتر  $\theta_{24|13}$ . ذکر این نکته ضروری است که در کلیه مراحل بالا تخمین پارامترهای  $\theta$  با استفاده از روش حداکثر لوگ-درست نمائی استفاده شده است به طوری که قرار دادن پارامترهای محاسبه شده منتخب بایستی سبب حداکثری شدن رابطه 13 شود:

$$L(\theta) = \sum_{t=1}^T \{ \log c_{t,j} (y_{t,i}, y_{t,j}; \theta) \} \quad (13)$$

در رابطه بالا  $c_{t,j}$  تابع چگالی مفصل دوبعدی،  $y_{t,i}$  و  $y_{t,j}$  سری های توزیع های حاشیه ای متغیرها و یا توزیع های حاشیه ای تولید

- 1- Frank
- 2- Joe
- 3- Gumbel
- 4- Clayton
- 5- Elliptical
- 6- T-student
- 7- Gaussian

تجربی چهار بعدی می‌باشند و n تعداد داده‌ها و i شماره داده‌ها می‌باشد.

### بحث

در این مطالعه با استفاده از روش حداقلی موضعی 34 رویداد سیل شناسایی شدند و سپس ویژگی‌های مهم این رویدادها یعنی حجم کل (V)، زمان اوج (Tp)، دبی اوج (P) و زمان پایه (B) استخراج شدند که مشخصات آماری ویژگی‌های مهم سیل در جدول 1 ارائه شده‌اند. اولین مرحله تحلیل فرکانس چند بعدی سیل استخراج توزیع‌های حاشیه‌ای چهار ویژگی مهم سیل یعنی حجم کل، دبی اوج، زمان اوج و زمان پایه سیل می‌باشد. در این تحقیق، توابع توزیع پارامتری مختلف به منظور برازش بر توزیع‌های حاشیه‌ای هر کدام از چهار ویژگی سیل به کار گرفته شدند و جهت تخمین پارامترهای توزیع‌های مختلف روش حداکثر درست نمائی<sup>7</sup> با استفاده از نرم افزار ایزی فیت (Easyfit) بکار برده شد. مطابق با تست‌های آماری کولموگروف اسمیرنوف و کای اسکور توزیع ویبول بعنوان مناسب‌ترین مدل جهت مدل‌سازی زمان اوج و حجم سیل تشخیص داده شد در حالی که توزیع‌های رابلی و لوگ لوجستیک به ترتیب به‌عنوان بهترین توزیع‌ها برای برازش بر متغیرهای زمان پایه و دبی اوج در سطح 5 درصد شناخته شدند. پارامترهای توابع توزیع‌های حاشیه‌ای انتخاب شده به‌همراه P- مقدار<sup>8</sup> تست‌های کولموگروف اسمیرنوف و کای اسکور آن‌ها محاسبه شدند که در جدول 2 ارائه شده‌اند در جدول 2  $\alpha$ ،  $\beta$ ،  $\gamma$  پارامترهای توزیع‌ها می‌باشند که مقادیر محاسبه شده آن‌ها ارائه شده است. با توجه به مقادیر p-value در جدول 2 می‌توان مشاهده نمود که توزیع‌های منتخب از لحاظ آماری در سطح 5 درصد معنی دار می‌باشند. محمدپور و خدادادی (1391) نیز در تحلیل فرکانس سیل رودخانه مهاباد، به‌منظور تعیین توزیع حاشیه‌ای حجم سیل و دبی اوج سیل از توزیع ویبول و برای زمان پایه از توزیع پیرسون تپ 3 استفاده نمودند. عباسیان و همکاران (1393) به منظور تحلیل فراوانی چند متغیره سیل حوزه آبریز کسلیان به منظور بدست آوردن توزیع‌های حاشیه‌ای دبی اوج، حجم و زمان پایه سیل توزیع‌های پارامتری مختلف و ناپارامتری را مورد آزمون قرار دادند و در نهایت نتیجه گرفتند که توزیع‌های ناپارامتری بهترین برازش را بر توزیع‌های حاشیه‌ای متغیرهای سیل دارند. اسراج و همکاران (2015) برای تحلیل چند متغیره سیل، توزیع لوگ پیرسون را بهترین توزیع جهت برازش بر دبی اوج و زمان پایه سیل و توزیع پیرسون را مناسب‌ترین توزیع برای حجم سیل یافتند. مرور بسیاری از مطالعات دیگر نشان

شده در بخش 3-2.5) انتخاب مناسب‌ترین خانواده مفصل برای هر جفت- مفصل که در نهایت بزرگ‌ترین لوگ درست نمائی و کم‌ترین مقدار معیار اطلاعات آکائیکه<sup>1</sup> (AIC) و اطلاعات بی‌زین<sup>2</sup> (BIC) را ارائه می‌دهند. معیارهای مذکور در معادلات 15 و 16 شرح داده شده‌اند: (Brechmann, et al. 2012)

$$L(u, y, \theta) = \sum_{i=1}^n \left( \log c_{12}(u_{1i}, u_{2i}, \theta_{12}) + \log c_{13}(u_{1i}, u_{3i}, \theta_{13}) + \log c_{14}(u_{1i}, u_{4i}, \theta_{14}) + \log c_{23}(u_{2i}, u_{3i}, \theta_{23}) + \log c_{24}(u_{2i}, u_{4i}, \theta_{24}) + \log c_{34}(u_{3i}, u_{4i}, \theta_{34}) \right) \quad (14)$$

$$AIC = -2 L(u, y, \theta) + 2k \quad (15)$$

$$BIC = -2 L(u, y, \theta) + 2 \ln(n)k \quad (16)$$

در معادلات بالا  $L(u, y, \theta)$  لوگ درست نمائی، k تعداد پارامترها و n به تعداد داده‌ها اشاره دارد. در نهایت به منظور محاسبه تابع توزیع توام چهار بعدی سی-واین، با استفاده از رابطه 4 می‌توان چگالی توام چهار بعدی را محاسبه نمود و سپس از حاصل جمع چگالی‌ها به تابع توزیع توام دست یافت.

### معیارهای ارزیابی

ارزیابی و انتخاب بهترین تابع توزیع توام چهار بعدی با استفاده از معیارهای ریشه میانگین مربعات خطا (RMSE<sup>3</sup>)، میانگین خطای مطلق (MAE<sup>4</sup>) و ضریب تبیین<sup>5</sup> (R<sup>2</sup>) صورت گرفت. معیارهای ارزیابی مذکور از مقایسه مفصل تجربی و مفصل سی-واین چهار بعدی توام تخمین زده شدند. مفصل تجربی با استفاده از رابطه 17 قابل محاسبه می‌باشد.

$$C_{\mathbb{E}}(p, r, s, w) = \frac{1}{n} \sum_{i=1}^n I \left( \frac{R_i}{n+1} \leq p, \frac{R_i}{n+1} \leq r, \frac{S_i}{n+1} \leq s, \frac{W_i}{n+1} \leq w \right) \quad (17)$$

در رابطه بالا I تابع شاخص<sup>6</sup>، n تعداد داده‌ها،  $R_i$ ،  $P_i$ ،  $S_i$  و  $W_i$  رتبه هر یک از نمونه‌ها در هر یک از چهار سری ویژگی‌های سیل مرتب شده می‌باشد. گنست و ریوست (1993). معیارهای ارزیابی از روابط 18-20 می‌توانند محاسبه شوند:

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (C_{T_i} - C_{\mathbb{E}_i})^2}{n}} \quad (18)$$

$$MAE = \frac{\sum_{i=1}^n |C_{T_i} - C_{\mathbb{E}_i}|}{n} \quad (19)$$

$$R^2 = \left( 1 - \frac{\sum_{i=1}^n (C_{T_i} - C_{\mathbb{E}_i})^2}{\sum_{i=1}^n C_{T_i}^2 - \frac{(\sum_{i=1}^n C_{\mathbb{E}_i})^2}{n}} \right) \quad (20)$$

در روابط بالا  $C_{\mathbb{E}}$  و  $C_T$  به ترتیب مفصل سی-واین و مفصل

- 1- Akaike information criterion
- 2- Bayesian information criterion
- 3- Root Mean of Square Error
- 4- Mean Absolute Error
- 5- Determination Coefficient
- 6- Indicator function

7- Method of Maximum Likelihood

8- P-Value

معنی‌دار نمی‌باشد بنابراین ساختارهایی که شامل جفت-مفصل (P- $T_p$ ) می‌باشند حذف گردیده و بدین ترتیب تنها 12 ساختار به منظور مطالعات بعدی مورد توجه واقع شدند. ابتدا در هر ساختار همه خانواده‌های مفصل ارشمیدسی (فرانک، جو، گامبل و کلایتون) و بیضوی (تی-استیودنت، گوسین) بر هر جفت-مفصل برآزش داده شدند و پارامترها با استفاده از روش ترتیبی تخمین زده شدند. سپس معیارهای ارزیابی برای همه خانواده‌های ارشمیدسی و بیضوی محاسبه شدند و در نهایت مفصل‌هایی که کم‌ترین مقدار معیارهای AIC و BIC و بیش‌ترین لوگ درست‌نمایی را در کل ساختار ارائه می‌دهند انتخاب گردیدند ساختارهای سی-واین مختلف به همراه پارامترهای تخمین زده شده و معیارهای AIC و BIC و لوگ درست‌نمایی در جدول 4 ارائه شده‌اند. به‌عنوان مثال در جدول 4 ساختار (V- $T_p$ -P-B) پارامتر  $\theta_{11}$ ،  $\theta_{12}$  و  $\theta_{13}$  در درخت 1 نشان-دهنده پارامترهای مفصل‌های  $C_{VTP}$ ،  $C_{VTP}$  و  $C_{VTP}$  و  $\theta_{21}$  در درخت 2 بیانگر پارامترهای مفصل‌های  $C_{PTP}$  و  $C_{PTP}$  و بالاخره  $\theta_{31}$  در درخت 3 پارامتر مفصل‌های  $C_{TPB|PV}$  نشان می‌دهد که ساختار سی-واین V- $T_p$ -P-B حداقل مقدار BIC و AIC و بیش‌ترین مقدار لوگ درست‌نمایی را در میان دیگر ساختارها می‌دهد. در بخش آخر این مطالعه توابع توزیع توام چهار بعدی ویژگی‌های مهم سیل محاسبه گردید. در این مطالعه به منظور آزمون نکوتی برآزش پارامترهای بدست آمده از تست انتگرال احتمالی<sup>1</sup> (PIT) (Aas et al. (2009) استفاده شد. مقادیر p-value تست PIT در جدول 5 ارائه شده است. نتایج جدول 5 حاصل از این آزمون نشان داد که p-value کلیه ساختارها بزرگ‌تر از 5 درصد بود و در سطح 5 درصد فرض صفر را نمی‌توان رد نمود. ساختارهای سی-واین بدست آمده همراه با پارامترهای جفت-مفصل‌ها به منظور محاسبه توابع توزیع چهار بعدی به کار برده شدند. ارزیابی و انتخاب بهترین تابع توزیع توام چهار بعدی با استفاده از معیارهای ریشه میانگین مربعات خطا ( $RMSE^2$ )، میانگین خطای مطلق و ضریب تبیین<sup>3</sup> ( $R^2$ ) و ( $MAE^4$ ) صورت گرفت.

معیارهای ارزیابی جدول 4-ساختارهای سی-واین مختلف به همراه پارامترهای تخمین زده شده مذکور از مقایسه مفصل تجربی و مفصل سی-واین چهار بعدی توام تخمین زده شدند و در جدول 6 نشان داده شدند. همچنین تست معنی‌داری تی-استیودنت نیز جهت تست معنی‌داری معیارهای بالا مورد استفاده قرار گرفت.

داد که برای برآزش ویژگی‌های سیل مانند حجم، زمان پایه و دبی اوج سیل و .. در هر حوضه آبریز توزیع‌های متفاوتی بعنوان مناسب-ترین توزیع شناخته می‌شوند و برای یک متغیر خاص هیچ گاه نمی-توان یک توزیع معین را معرفی نمود. قبل از تخمین پارامترهای مفصل ابتدا بایستی جفت متغیرهایی که از لحاظ آماری همبستگی معنی‌دار دارند شناسائی شوند تا در مرحله بعد این متغیرها بعنوان گره‌ها در ساختار مستقیم به یکدیگر متصل شده و ساختارهای مختلف را ایجاد نمایند و همچنین جفت متغیرهای غیر همبسته از ساختارها حذف می‌گردند. پس از استخراج ویژگی‌های سیل، به منظور شناسائی جفت متغیرهایی که از لحاظ آماری با هم وابستگی معنی‌دار دارند همبستگی تاو کندال و p-value آن بین همه جفت-متغیرها محاسبه گردید که مقادیر تاو کندال در جدول 3 ارائه شده‌اند. مقدار p-value همه جفت-متغیرها به استثناء جفت  $T_p$ ، P معنی‌دار و نزدیک به صفر می‌باشند اما برای جفت  $T_p$ ، P 39/0 می‌باشد. بنابراین از آنجا که این مقدار بزرگ‌تر از 5 درصد می‌باشد پس در سطح 5 درصد همبستگی جفت  $T_p$ -P معنی‌دار نمی‌باشد. در این مطالعه به منظور ایجاد ساختارهای مختلف سی-واین چهار بعدی توسط چهار ویژگی مهم سیل، جایگشت‌های مختلف 4 متغیر ویژگی‌های سیل (حجم، دبی اوج، زمان اوج و زمان پایه سیل) در 4 گره ساختار سی-واین چهار بعدی طبق رابطه ( $1 \times 2 \times 3 \times 4 = 24$ ) محاسبه گردید و عدد 24 حاصل می‌شود بنابراین 24 ساختار سی-واین مختلف حاصل گردید.

جدول 1- پارامترهای توابع توزیع‌های حاشیه‌ای و معیارهای نکوتی

برآزش						
ویژگی-های سیل	$\delta$	$\beta$	$\alpha$	$\sigma$	P-value K.S	p-value $\chi^2$
V	0	74/5e+6	9/0	-	0/94	90/0
P	-6/13	-	-	74	0/89	0/89
B	-4/11	7/83	7/3	-	0/97	0/87
$T_p$	-6/1	4/39	1/1	-	0/65	0/75

جدول 2- مقادیر تاو کندال جفت متغیرها

جفت-متغیرها	تاو کندال	p-value
B, $T_p$	0/53	0
B, V	0/61	0
B, P	0/38	0
$T_p$ , V	0/38	0
$T_p$ , P	0/10	0/39
V, P	0/64	0

از آنجا که وابستگی بین دو متغیر دبی و زمان اوج ( $P-T_p$ )

1- probability integral transform (PIT)  
2- Root Mean of Square Error  
3- Determination Coefficient  
4- Mean Absolute Error

جدول 3- معیارهای AIC و BIC و لوگ درست نمائی (در ساختارهای زیر پارامتر کامل با (Gu)، فرانک (F)، گوسین (G)، جو (J)، تی - استیودنت (T).

BIC	AIC	Loglik	$\theta_{T T}$	$\theta_{T G}$	$\theta_{T F}$	$\theta_{T J}$	$\theta_{T Gu}$	$\theta_{T V}$	ساختارهای سی-واین
-111/8	-121	66/5	J(2/1)	G(-0/4)	G(0/7)	Gu(2/1)	G(8/0)	G(6/0)	(B-P-V-T <sub>F</sub> )
-105/9	-115	63/5	Gu(2/2)	G(-0/5)	G(-0/1)	G(0/6)	G(0/8)	Gu(2/1)	(B-T <sub>F</sub> -V-P)
-105/9	-115	63/7	Gu(2/2)	G(-0/1)	G(-0/5)	G(0/8)	G(0/6)	Gu(2/1)	(B-T <sub>F</sub> -P-V)
-120	-121	66/5	J(2/1)	G(0/7)	G(-0/5)	G(0/6)	Gu(2/1)	G(0/8)	(B-P-T <sub>F</sub> -V)
-106/07	-117/6	65/6	T(-0/6,2/5)	G(0/7)	G(-0/1)	G(0/6)	Gu(2/1)	G(0/8)	(B-V-T <sub>F</sub> -P)
-106/7	-117/3	65/6	T(-0/6,2/5)	G(-0/1)	G(0/7)	Gu(2/1)	G(0/6)	G(0/8)	(B-V-P-T <sub>F</sub> )
-106/3	-115/2	63/6	F(5/3)	F(-1/1)	F(-4/4)	G(0/8)	G(0/6)	G(0/8)	(V-P-T <sub>F</sub> -B)
-106/3	-115/2	63/6	F(5/3)	F(-4/4)	F(-1/1)	G(0/6)	G(0/8)	G(0/8)	(V-P-B-T <sub>F</sub> )
--104/7	-113/8	62/9	F(-3/8)	F(-1/1)	J(2/3)	G(0/8)	G(0/6)	G(0/8)	(V-B-T <sub>F</sub> -P)
--104/7	-113/8	62/9	F(-3/8)	J(2/3)	F(-1/1)	G(0/6)	G(0/8)	G(0/8)	(V-B-P-T <sub>F</sub> )
-4/113	-122/5	2/67	F(7/2)	J(3/2)	F(-4/4)	G(8/0)	G(0/8)	G(0/5)	(V-T <sub>F</sub> -P-B)
-4/113	-5/122	2/67	F(7/2)	F(-4/4)	J(3/2)	G(8/0)	G(8/0)	G(6/0)	(V-T <sub>F</sub> -B-P)

جدول 4- p-value برای تست معنی داری PIT ساختارهای سی-واین

ساختارهای سی-واین	PIT	ساختارهای سی-واین	PIT
B-P-V-T <sub>p</sub>	0/67	V-P-T <sub>p</sub> -B	0/65
B-T <sub>p</sub> -V-P	0/69	V-P-B-T <sub>p</sub>	0/65
B-T <sub>p</sub> -P-V	0/69	V-B-T <sub>p</sub> -P	0/97
B-P-T <sub>p</sub> -V	0/67	V-B-P-T <sub>p</sub>	0/97
B-V-P-T <sub>p</sub>	0/61	V-T <sub>p</sub> -P-B	0/75
B-V-T <sub>p</sub> -P	0/61	V-T <sub>p</sub> -B-P	0/75

گره‌ها توسعه یافتند. بطوریکه با استفاده از جایگشت‌ها یا به عبارتی آمون همه حالت‌های قرارگیری متغیرها در مکان‌های مختلف سی-واین چهار بعدی کلیه ساختارها توسعه یافته و مورد بررسی قرار گرفتند.

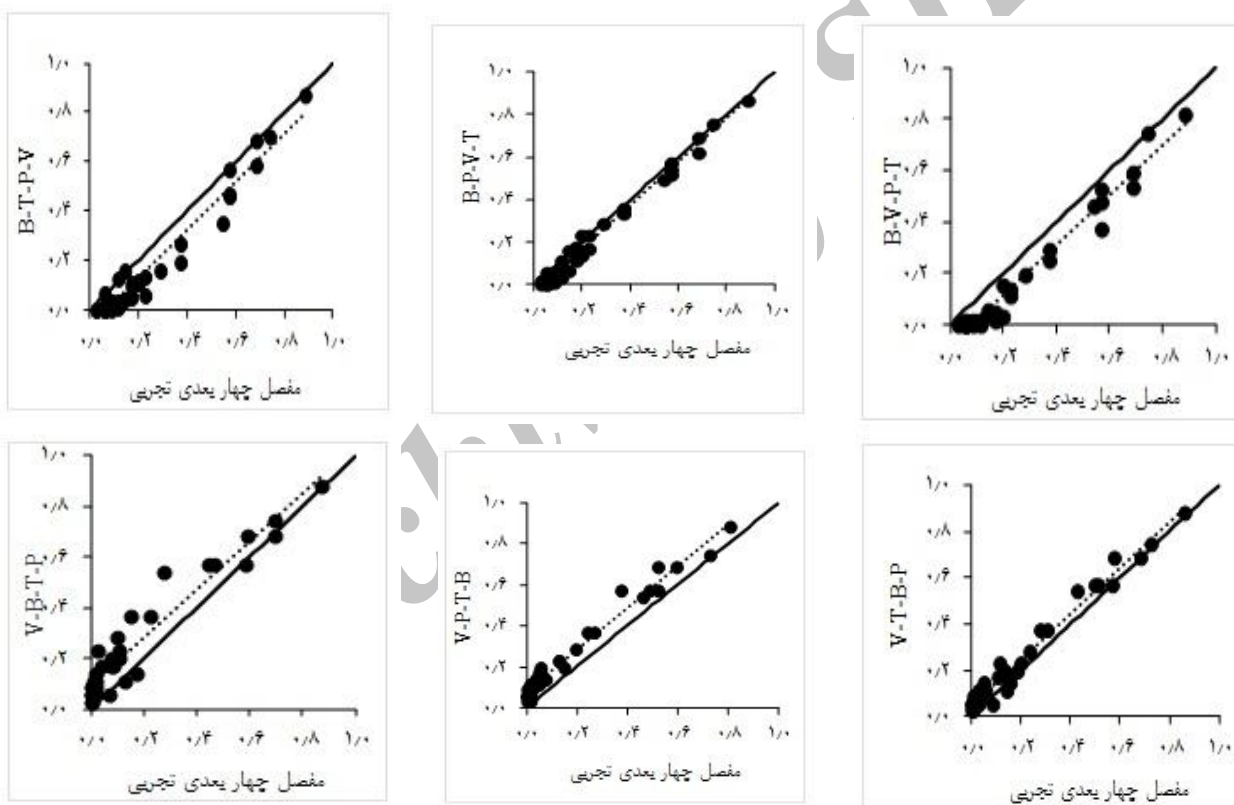
از جدول 4 می‌توان دریافت که بکارگیری وابسته‌ترین جفت-مفصل‌ها در درخت 1 دقیق‌ترین ساختار را لزوما ارائه نمی‌دهد چرا که ساختارهای V-، V-P-T<sub>p</sub>-B، V-P-B-T<sub>p</sub>، V-T<sub>p</sub>-B-P، V-T<sub>p</sub>-P-B، B-P-T<sub>p</sub> و B-V-T<sub>p</sub>-P شامل جفت-مفصل‌های با بیش‌ترین وابستگی  $T_{VP} = 1/44$ ،  $T_{VB} = 1/41$  و  $T_{VT} = 1/23$  می‌باشند و با این حال دارای بیش‌ترین مقدار ضریب همبستگی و کم‌ترین مقدار RMSE و MAE نبوده و در نتیجه دقیق‌تر از بقیه ساختارها نمی‌باشند. با این وجود نکته قابل توجه این است که جفت-مفصل‌های ساختار B-P-V-T<sub>p</sub> دارای بالاترین مقادیر تاو کندال نمی‌باشند.

از جدول 6 می‌توان مشاهده نمود که مقدار p-value کلیه ساختارها از 5 درصد بزرگ‌تر می‌باشند بنابراین بین این ساختارها و مفصل تجربی اختلاف معنی داری وجود ندارد اما با توجه به معیارهای ارزیابی می‌توان نتیجه گرفت که ساختار (B-P-V-T<sub>p</sub>) دقیق‌ترین مدل در بین دیگر مدل‌ها است. از آنجا که دارای بیش‌ترین مقدار ضریب معنی داری تی-استیودنت 0/3، ضریب تبیین یعنی 0/985 و به ترتیب دارای کم‌ترین مقادیر RMSE و MAE یعنی 0/042، 0/034 می‌باشد. در مطالعات گذشته (Aas et al., 2009) و (Vernieuwe et al., 2015) مکان هر یک از متغیرها در ساختار و این با توجه به مقادیر تاو کندال تعیین شده است. به‌طوریکه محققان در گره‌های متصل به هم ساختار و این از متغیرهایی استفاده نموده‌اند که در کنار هم وابسته‌ترین جفت-مفصل‌ها را ایجاد می-نمایند و یا به عبارتی از جفت-متغیرهایی در درخت اول استفاده می-نمایند که دارای بیش‌ترین مقادیر تاو-کندال می‌باشند. اما در این مطالعه همه ساختارهای ممکن با استفاده از جایگشت متغیرها در همه



جدول 5- معیارهای ارزیابی RMSE، MAE و  $R^2$  مربوط به محاسبه توابع توزیع توام چهار بعدی ویژگی‌های سیل

ساختارهای سی-واین	RMSE	$R^2$	MAE	p-value
B-T <sub>p</sub> -P-V	0/089	0/95	0/074	0/11
B-T <sub>p</sub> -V-P	0/075	0/96	0/063	0/16
B-P-V-T <sub>p</sub>	0/042	0/985	0/034	0/30
B-V-T <sub>p</sub> -P	0/102	0/956	0/088	0/06
B-V-P-T <sub>p</sub>	0/095	0/967	0/085	0/07
B-P-T <sub>p</sub> -V	0/086	0/974	0/076	0/09
V-T <sub>p</sub> -P-B	0/092	0/930	0/079	0/10
V-T <sub>p</sub> -B-P	0/060	0/976	0/052	0/20
V-P-B-T <sub>p</sub>	0/080	0/975	0/070	0/11
V-P-T <sub>p</sub> -B	092/0	973/0	083/0	0/07
V-B-P-T <sub>p</sub>	094/0	973/0	085/0	0/07
V-B-T <sub>p</sub> -P	109/0	920/0	089/0	0/07



شکل 2- مقایسه مفصل‌های سی-واین چهار بعدی با مفصل‌های تجربی چهار بعدی (دایره‌های توپر نمایش دهنده نقاط مفصل چهار بعدی تجربی در مقابل ساختار سی-واین می-باشند. خط نقطه چین رابطه خطی بین مقادیر دو مفصل تجربی و سی-واین، و خط مشکی خط 45 درجه را نشان می‌دهند)

بنابراین برای دستیابی به دقیق‌ترین ساختار سی-واین بایستی تمامی ساختارهای ممکن را ایجاد نمود و سپس دقیق‌ترین ساختار را با توجه به معیارهای ارزیابی انتخاب نمود. در این مطالعه همچنین به منظور ارزیابی دقت توزیع‌های توام

اما با این وجود این ساختار بیش‌ترین ضریب تبیین ( $R^2=0.985$ ) و کم‌ترین مقدار معیارهای RMSE و MAE را ارائه می‌نماید. بنابراین می‌توان نتیجه گرفت که استفاده از جفت-مفصل‌های با بیش‌ترین مقادیر وابستگی الزاماً بهترین ساختار را ایجاد نمی‌نماید

(V-Tp) که لزوماً دارای جفت-متغیرهای با بالاترین همبستگی تاو-کندال نمی‌باشد با کم‌ترین خطا نسبت به دیگر ساختارهای سی-واین به‌عنوان دقیق‌ترین ساختار شناخته شد. در ساختار (B-P-V-Tp) مفصل‌های جو، گوسین و گامبل با توجه به معیارهای حداکثر درست نمائی، AIC و BIC به‌عنوان مناسب‌ترین مفصل‌ها در جفت-مفصل‌های تجزیه شده شناخته شدند. همچنین ذکر این نکته ضروری است که در مطالعه موردی داده‌های ایستگاه وینار انتخاب متغیر زمان پایه به‌عنوان متغیر مرکزی در ساختار-سی واین یک انتخاب مناسب می‌باشد که منجر به ارائه دقیق‌ترین ساختار (B-P-V-Tp) جهت تعیین توزیع تجمعی چهار بعدی می‌گردد. پیشنهاد می‌شود که از روش ارائه شده جهت مدل‌سازی توزیع‌های توام دیگر داده‌های هیدرولوژیکی استفاده شود چرا که ساختار سی-واین با در نظر گرفتن روابط بین متغیرها شبیه‌سازی دقیق‌تری از روابط بین متغیرهای هیدرولوژیکی ارائه می‌دهد.

### منابع

محمدپور، ع و خدادادی، ا. 1391. تحلیل فراوانی سیل سه متغیره با استفاده از تابع مفصل خانواده برآورد Plackett، برآورد پارامترها با استفاده از الگوریتم ژنتیک GA. نهمین کنگره بین‌المللی مهندسی عمران، دانشگاه صنعتی اصفهان. اردیبهشت ماه 19-21.

عباسیان، م، موسوی ندوشنی، س، س. 1393. تحلیل فراوانی چند متغیره سیلاب با استفاده از تابع مفصل و توزیع‌های حاشیه‌ای پارامتری و ناپارامتری، مجله علمی-پژوهشی عمران مدرس. 4,14 : 81-92

Aas, K., Czado, C., Frigessi, A and Bakken, H. 2009. Pair-copula constructions of multiple dependence. Insurance: Mathematics and Economics. 44: 182-198.

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.

Bedford, T., Cooke, R. 2002. Vines – A new graphical model for dependent random variables, Annals of Statistics. 30.4: 1031-1068.

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.

De Michele, C., Salvadori, G. 2003. A generalized Pareto intensity-duration model of storm rainfall exploiting 2-copulas. Journal of Geophysical Research. 108: 15-27

سی واین چهار بعدی و مقایسه آن‌ها با مفصل تجربی چهار بعدی از تست گرافیکی نیز علاوه بر تست‌های آماری استفاده شد. بطوریکه در این روش مقادیر توزیع توام چهار بعدی حاصل از ساختار سی-واین در مقابل توزیع توام تجربی حاصل از مفصل تجربی رسم شدند. شکل 2 مقایسه بین شش توزیع توام منتخب سی-واین را با مفصل تجربی چهار بعدی نشان می‌دهد. از شکل 2 می‌توان مشاهده نمود که ساختار B-P-V-Tp انطباق و هم‌پوشانی بیش‌تری با مفصل تجربی دارد این هم آزمون داخلی است. آزمون معنی‌داری لازم دارد از آنجا که خط (نقطه چین) که بیانگر رابطه بین مفصل سی-واین و تجربی می‌باشد کاملاً منطبق بر خط 45 درجه می‌باشد. با توجه به میزان انطباق خطوط نقطه چین با خط 45 درجه می‌توان نتیجه گرفت که ساختار B-P-V-Tp نیز پس از ساختار B-P-V-Tp دارای دقت خوبی می‌باشد و همچنین ساختار V-B-Tp-P دارای کم‌ترین دقت نسبت به دیگر ساختارها می‌باشد. بنابراین با مقایسه شکل 2 و جدول 4 می‌توان دریافت که نتایج حاصل از هر دو تایید کننده یکدیگر می‌باشند.

### نتیجه‌گیری

در این مطالعه برای اولین بار از ساختار سی-واین به منظور تخمین توزیع توام چهار بعدی مهم‌ترین ویژگی‌های سیل ایستگاه وینار در حوضه آبریز آچی چای بهره گرفته شد. بنابراین ساختار سی-واین جهت توصیف ساختار وابستگی میان حجم کل، دبی اوج، زمان پایه و زمان اوج سیل بکار برده شد. ابتدا مناسب‌ترین توزیع‌های حاشیه‌ای تئوری بر توزیع‌های تجربی ویژگی‌های سیل برآزش داده شدند. سپس با استفاده از جایگشت متغیرها در ساختارهای سی-واین چهار بعدی 12 ساختار مختلف توسعه داده شد و خانواده‌های مفصل مختلف از خانواده‌های ارشمیدسی و بیضوی بر جفت-مفصل‌ها مورد آزمون قرار گرفتند و مطابق با معیارهای مختلف مناسب‌ترین مفصل به‌منظور برآزش بر هر جفت-مفصل انتخاب گردید. در مرحله بعدی مناسب‌ترین ساختارها با توجه به معیارهای AIC، BIC و حداکثر لوگ درست نمائی انتخاب شدند و ساختارهای حاصل جهت محاسبه توزیع‌های توام چهار بعدی مورد استفاده قرار گرفتند و در نهایت دقیق‌ترین ساختار با توجه به میزان انطباق توزیع توام بر توزیع تجربی به‌دست آمد که این انتخاب بر اساس روش‌های گرافیکی و آماری صورت پذیرفت. در نهایت نتیجه‌گیری شد که استفاده از جفت-متغیرهای با بالاترین همبستگی تاو-کندال معیار مناسبی برای دستیابی به مناسب‌ترین ساختار نمی‌باشد بلکه بایستی همه جایگشت‌های ممکن انجام شوند و سپس با توجه به معیارهای مختلف اقدام به انتخاب دقیق‌ترین ساختار نمود. در این مطالعه مفصل‌های تجربی و سی-واین چهار بعدی با توجه به معیارهای گرافیکی و آماری (MAE، RMSE و  $R^2$ ) با هم مقایسه شدند و در نهایت ساختار (B-P-

- Mirabbasi, R., Fakheri-Fard, A and Dinpashoh, Y. 2012. Bivariate drought frequency analysis using the copula method. *Theoretical and Applied Climatology*. 108: 191–206.
- Salvadori, G., De Michele, C. 2004. Analytical calculation of storm volume statistics involving Pareto-like intensity duration marginals, *Geophys. Geophysical Research Letters*. 31.4:1-4
- Salvadori, G., De Michele, C. 2006. Statistical characterization of temporal structure of storms, *Advances in Water Resources*. 29: 827–842.
- Serinaldi, F and Grimaldi, S. 2007. Fully nested 3-copula: procedure and application on hydrological data. *Journal of Hydrologic Engineering (ASCE)*. 12: 420–430.
- Singh, V.P., Zhang, L. 2007. IDF curves using the Frank Archimedean copula. *Journal of Hydrologic Engineering (ASCE)* 12: 651–662.
- Sklar, A. 1959. *Fonction de re'partition a' n dimensions et leurs marges*, vol. 8. Publications de L'Institut de Statistique, Universite' de Paris: Paris; 229–231.
- Sloto, R.A., Crouse, M.Y. 1996. HYSEP: A computer program for streamflow hydrograph separation and analysis. U.S. Geological Survey, Water-Resources Investigations, Report 96-4040, Pennsylvania, 46 p.
- Song, S., Singh, V. 2010. Meta-elliptical copulas for drought frequency analysis of periodic hydrologic data, *Stochastic Environmental Research and Risk Assessment*. 24: 425–444.
- Sraj, M, Bezak, N. Brilly, M. 2015. Bivariate flood frequency analysis using the copula function: a case study of the Litija station on the Sava River. *Hydrological Processes* 29: 225–238.
- Vernieuwe, H., Vandenberghe, S., De Baets, B., Verhoest, N.E.C. 2015. A continuous rainfall model based on vine copulas. *Hydrology and Earth System Sciences* 19.6:2685–2699
- Zhang, L., Singh, V.P. 2006. Bivariate flood frequency analysis using copula method. *Journal of Hydrologic Engineering (ASCE)* 11.2: 150–164.
- Zhang, L., Singh V.P. 2007a. Trivariate flood frequency analysis using the Gumbel–Hougaard copula, *Journal of Hydrologic Engineering (ASCE)*, ASCE 12.4: 431–439.
- Zhang, L, Singh, V.P 2007b. Gumbel-Hougaard copula for trivariate rainfall frequency analysis. *Journal of Hydrologic Engineering (ASCE)*. 12.4:409–419.
- 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.
- Genest, C., Favre, A., Beliveau, J., and Jacques, C. (2007), Metaelliptical copulas and their use in frequency analysis of multivariate hydrological data, *Water Resour. Res.*43, W09401, doi:10.1029/2006WR00527.
- Genest, C., Rivest, L.-P., (1993), Statistical inference procedures for bivariate Archimedean copulas. *Journal of the American Statistical Association*. 88, 1034–1043.
- Gräler, B., van den Berg, M.J., Vandenberghe, S., Petroselli, A., Grimaldi, S., De Baets, B 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 Sciences*. 17.4: 1281–1296
- Grimaldi, S, Serinaldi, F. 2006. Asymmetric copula in multivariate flood frequency analysis. *Advances in Water Resources*. 29.8,1115–1167.
- Gyasi-Agyei, Y., Melching, C. 2012. Modelling the dependence and internal structure of storm events for continuous rainfall simulation, *Journal of Hydrology*. 464-465: 249–261.
- Joe, H. 1996. Families of m-variate distributions with given margins and  $m(m-1)/2$  bivariate dependence parameters. In L. Rueschendorf, B. Schweizer, M and Taylor, D (Eds.). *Distributions with fixed marginal and related topics*, pp. 120-141. Hayward: In statute of Mathematical Statistics.
- Kao, S., Govindaraju, R. 2008. Trivariate statistical analysis of extreme rainfall events via the Plackett family of copulas, *Water Resources Researches*. 44. 2, 1-19.
- Kao, S.C., Govindaraju, R.S. 2010. A copula-based joint deficit index for droughts. *Journal of Hydrology*. 380.1–2, 121–134.
- Karmakar, S and Simonovic, S.P. 2009. Bivariate flood frequency analysis. Part 2: a copula-based approach with mixed marginal distributions. *Journal of Flood Risk Management*. 2.1:32–44.
- Kurowicka, D., Cooke, R. 2007. Sampling algorithms for generating joint uniform distributions using the vine-copula method. *Computational Statistics & Data Analysis*. 51: 2889–2906.

## Modeling the Four-Dimensional Joint Distribution Function of Flood Characteristics Using C-Vine Structure

M. Shafaei.<sup>1</sup> A. Fakheri-Fard.<sup>2</sup> Y. Dinpashoh.<sup>3</sup> R. Mirabbasi Najafabadi<sup>4</sup>

Recived: Mar.06, 2016

Accepted: Aug.17, 2016

### Abstract

The recognition of a proper multi-dimensional copula for determination the dependence structure in multi-dimensional data is not easy. Because multi- dimensional copulas such as the multi-dimensional Elliptical and Archimedean copulas lack flexibility to capture dependence structure and include other limitations, such as parameter restrictions. To dominate these limitations, vine copulas have been presented and employed for simulation of hydrology variables. To realize and fully perceive the mixed and hidden dependence patterns between flood characteristics (flood volume (V), peak flow (P), base time (B) and peak time ( $T_p$ )), a mixture of C-vine copulas is proposed describing the dependence structure between the main flood characteristics and then simulates them use the most proper C-vine structures. As a C-vine structure consists several parameters capturing the dependence structure through cascade of pair-copulas, the proposed structure can explain completely complex and hidden dependence patterns in multi- dimensional data. In the present study, to achieve the most suitable C-vine structures, first different structures were created through the permuting of flood characteristics in the C-vine structure and then the most appropriate copula family among Elliptical or Archimedean copula families was selected for each pair-copula according to Log Maximum Likelihood (LML), AIC and BIC evaluation criteria. Finally, the obtained four-dimensional joint distribution functions were compared with empirical copula and according to RMSE, MAE and  $R^2$  criteria and graphical method, structure (B-P-V- $T_p$ ) was known as the most accurate structure. Thus, it was found that the selection of flood base time as central variable is a appropriate selection and plays role of controlling variable for volume, peak discharge and peak time of flood variables.

**Keywords:** Akaike information criteria, Archimedean copula, Elliptic copula, Flood volume, pair-copula.

1- Ph. D Candidate in Water Resources Engineering, Department of Water Engineering, Tabriz University, Tabriz, Iran.  
2- Professor, Department of Water Engineering, Tabriz University, Tabriz, Iran.  
3- Associate Professor, Department of Water Engineering, Tabriz University, Tabriz, Iran.  
4- Assistant Professor. Department of Water Engineering, Shahrekord University, Shahrekord, Iran.  
(\*- Corresponding Author Email: m.shafaei@tabrizu.ac.ir)