دانشگاه علوم کشاورزی و منابع طبیعی گرگان
نشریه حفاظت و بهره‌برداری از منابع طبیعی
جلد چهارم، شماره دوم، ۱۳۹٤
http://ejang.gau.ac.ir

توابع کاپولا و کاربرد آن در هیدرولوژی استوکاستیک

عبدالرضا بهره‌مند[۱]، احسان الوندی[۲]، مرضیه بهرامی[۲]، مریم دشتی مرویلی[۲]، حسام هروی[۲]، غلامرضا خسروی[۲]، *آیدینگ کرنژادی[۲]، حجت‌اله صمدی‌ارقینی[۲]، مریم تاجیکی[۲] و مهدی تیموری[۲]

[۱]دانشیار گروه آبخیزداری و مدیریت مناطق بیابانی، دانشگاه علوم کشاورزی و منابع طبیعی گرگان، ایران،
[۲]دانشجوی دکتری علوم و مهندسی آبخیزداری، دانشگاه علوم کشاورزی و منابع طبیعی گرگان، ایران
تاریخ دریافت: ۱۳۹٤/۰۲/۲۱ ؛ تاریخ پذیرش: ۱۳۹٤/۰٤/۳۱

چکیده

در هیدرولوژی متغیرهای زیادی به‌عنوان نماینده رفتار سیستم جهت مدل‌سازی فرآیندها مورد بررسی قرار می‌گیرند. مستقل فرض کردن این متغیرها صحت نتایج نهایی مدل‌سازی را زیر سوال می‌برد. از طرفی در بعضی از مسائل و پدیده‌های هیدرولوژی (سیل، خشکسالی و غیره) چندین متغیر مؤثر بوده و در حالی پدیده را تحت تأثیر قرار می‌دهند که خود این متغیرها با همدیگر وابستگی و همبستگی دارند. استفاده از توزیع‌های آماری برای چنین متغیرهایی بدون در نظر گرفتن این همبستگی‌ها بر عدم قطعیت‌ها می‌افزاید. بنابراین، دانستن ارتباط بین توزیع‌های حاشیه‌ای متغیرهای مختلف به‌منظور درک قوانین حاکم بر این وابستگی‌ها می‌تواند در شناخت وقایع هیدرولوژیکی مشاهده شده بسیار مؤثر واقع شود. لذا به‌منظور افزایش اطمینان به نتایج تحلیل‌ها باید از رویکردهای آماری چند متغیره استفاده نمود. روش سنتی انجام تحلیل‌های چند متغیره، استفاده از توابع توزیع چند متغیره کلاسیک می‌باشد. اما در استفاده از این توابع مشخص بودن توزیع‌های حاشیه‌ای و یکسان بودن نوع آن‌ها الزامی بوده، لذا استفاده از این روش‌ها با محدودیت مواجه است. روش مناسب‌تر برای

*مسئول مکاتبه: aidin.kornejady@gmail.com

تحلیل‌های چند متغیره که بر محدودیت‌های توابع چند متغیره کلاسیک فائق آمده است، استفاده از توابع مفصل (کاپولا) می‌باشد. پیشرفت‌های اخیر در مطالعات ریاضی کاربردی نشان داده است که توابع مفصل ابزاری مفید برای بررسی رفتار آماری متغیرهای وابسته می‌باشد. لذا این مقاله جهت تحلیل چند متغیره فرآیندهای هیدرولوژیکی به بررسی تاریخچه توسعه و نحوه کارکرد و کاربردهای توابع رایج کاپولا در مدل‌سازی هیدرولوژیکی- تصادفی می‌پردازد.

واژه‌های کلیدی: کاپولا، تابع مفصل، توزیع‌های دومتغیره، هیدرولوژی استوکاستیک

مقدمه

به‌طور کلی اقدامات انجام شده در زمینه علم هیدرولوژی در دو زمینه هیدرولوژی کاربردی (مهندسی) و هیدرولوژی نظری (تحلیلی) می‌باشد. این دو قطب هیدرولوژی محدوده وسیعی از فعالیت‌ها و مطالعات را دربر می‌گیرند. هیدرولوژی نظری نیز به دو بخش قطعی (فیزیکی) و استوکاستیک[1] (آماری- احتمالاتی- تجربی در دو دوره کلاسیک و نوین) تقسیم می‌شود. نخستین محققان هیدرولوژی، بر پایه مشاهدات مستقیم خود از طبیعت اظهار داشته‌اند که تمام متغیرهای هیدرولوژیکی ذاتاً تصادفی‌اند و به مقدار کم یا زیاد تحت قوانین احتمال می‌باشند. بررسی تغییرات مکانی- زمانی متغیرها به‌عنوان روش‌های استوکاستیک معرفی شده است (مارکو و همکاران، ۱۹۹۳). پدیده‌های هیدرولوژیکی مانند بارش، رواناب، سیل، خشکسالی، تصادفی و چند متغیره بوده و توسط مشخصه‌های شدت، مدت و بزرگی بیان می‌شوند. این مشخصه‌ها به‌هم وابسته بوده، به‌صورت مستقل از هم تغییر نکرده و هر کدام بر دیگری تأثیر گذارند. لذا با تحلیل پدیده‌های هیدرولوژیکی به‌صورت تک متغیره نمی‌توان معادله همبستگی بین مشخصه‌های آن را نشان داده و ارزیابی کامل و دقیقی از آن ارائه کرد. برای داده‌های هیدرولوژیکی سعی می‌شود توابع احتمالاتی مناسبی تعیین و ترسیم گردد تا از روی آن‌ها بتوان مقدار متغیر موردنظر را به ازای احتمالات مختلف محاسبه کرد. مبتنی‌بر تحقیقات پیشین، استفاده از روش‌های تک متغیره منجربه برآورد کم یا زیاد مقدار پدیده و خسارات ناشی از آن خواهد شد (رینال و سالاس، ۱۹۸۷؛ برونیو و همکاران، ۱۹۹٤ و یو و همکاران، ۲۰۰۱). از نظر تئوری توابع توزیع احتمالاتی مختلفی وجود دارد که معادله‌های آنان مشخص است. داده‌هایی که به‌صورت

1- Stochastic

تجربی اندازه‌گیری و ثبت شده‌اند، با این توابع توزیع تئوری برازش داده شده و بهترین تابعی که با داده‌ها مطابقت دارد به‌عنوان تابع توزیع احتمال برگزیده می‌شود تا از روی آن به ازای هر احتمال موردنظر مقدار متغیر هیدرولوژی به‌دست آید (علیزاده، ٢٠١٠).

ذکر این نکته ضروری است که هیچ توزیع آماری نمی‌تواند دقیقاً بر داده‌های مشاهده‌ای برازش خوبی داشته باشد و یکی از توابع توزیعی در مقایسه با سایر توزیع‌ها به‌عنوان بهترین توزیع انتخاب می‌گردد (شیائو و همکاران، ٢٠٠٦).

در شبیه‌سازی متغیرهای تصادفی در هیدرولوژی، مفاهیم همبستگی و وابستگی تصادفی میان متغیرها از اهمیت ویژه‌ای برخوردار است. از آنجا که ضرایب همبستگی خطی و رتبه‌ای قادر به اندازه‌گیری این مفاهیم به‌طور کامل نیستند، لذا اخیراً در شبیه‌سازی متغیرهای تصادفی روش‌های جایگزینی برای اندازه‌گیری وابستگی تصادفی میان متغیرها توسعه داده شده است که از جمله مهم‌ترین این روش‌ها، روش کاپولا[1] است. کاپولا در سال ١٩٥٩ توسط اسکلار[2] معرفی شد و پس از آن در تحقیقات فراوانی در زمینه علم آمار و احتمالات، مورد استفاده قرار گرفت. در دهه هشتاد میلادی استفاده کاربردی از این توابع در علم اقتصاد و محاسبات مالی شروع شد. یک کاپولا جهت تهیه توزیع توأم متغیرها از الگوریتم‌های مناسب‌تری استفاده می‌کند در نتیجه واقع بین‌تر می‌باشد. در واقع کاپولا قادر به ساخت توزیع توأم از متغیرهای وابسته دارای توزیع حاشیه‌ای متفاوت می‌باشد و برعکس. با توجه به اهمیت پیش‌بینی در بازارهای بورس، کاربرد این روش بیشتر در محاسبات و تحلیل‌های بازارهای مالی و علوم آماری بوده است (فریس و والز، ١٩٩٨). از سال ٢٠٠٣ به بعد افرادی از قبیل سالوادوری، فاور و گنست در استفاده از توابع کاپولا در بخش هیدرولوژی پیشگام شده‌اند به‌طوری که فاور و همکاران (٢٠٠٤)، از کاپولا در مدل‌سازی متغیرهای هیدرولوژیک دو حوزه آبخیز در کبک کانادا استفاده کردند. آن‌ها بیان نمودند همبستگی بین متغیرهای موجود در مباحث هیدرولوژی را می‌توان با استفاده از این روش مدل‌سازی کرد و نتایج بهتری نسبت به روش‌های پیش‌بینی سنتی یک متغیره داشت. پس از آن مفهوم توابع مفصل به سرعت در زمینه‌های مختلف هیدرولوژی شامل تحلیل فراوانی سیلاب (دی میلچه و همکاران، ٢٠٠٥؛ شیائو و همکاران، ٢٠٠٦؛ گینست و همکاران، ٢٠٠٧؛ چبنا و اوآدرا، ٢٠٠٩؛ چن و همکاران، ٢٠١٢ و عباسیان و همکاران، ٢٠١٣)، تحلیل چند متغیره

1- Copula
2- Sklar

خصوصیات بارش (سالوادوری و دی میچله، ۲۰۰۶؛ ژانگ و سینگ، ۲۰۰۷؛ کائو و گوینداراجو، ۲۰۰۸؛ سردینالدی، ۲۰۰۸؛ سردینالدی، ۲۰۰۹ و ژانگ و همکاران، ۲۰۱۲) تحلیل چند متغیره خصوصیات خشکسالی (سالوادوری و میچل، ۲۰۰۴؛ شیا، ۲۰۰۶؛ ونگ و همکاران، ۲۰۰۸؛ امیدی و همکاران، ۲۰۱۰؛ میراکبری و همکاران، ۲۰۱۲؛ میرعباسی نجف‌آبادی و همکاران، ۲۰۱۴ و عبدالحسینی، ۲۰۱۱)، و احتمال توأم رسوب و رواناب (ژانگ و همکاران، ۲۰۱۴) مورد بررسی قرار گرفته است. تحقیق پیش رو با تکیه بر مرور منابع انجام شده در زمینه کاپولاها، به بررسی مباحث تئوری، تشریح توابع رایج، کاربردها و در نهایت نحوه انتخاب بهترین تابع می‌پردازد.

مواد و روش‌ها

توزیع حاشیه‌ای: اگر فرض شود X و Y متغیرهای تصادفی باشند، مناسب‌ترین توزیع احتمال برازش داده شده بر متغیر تصادفی X بدون در نظر گرفتن توزیع متغیر تصادفی Y، توزیع حاشیه‌ای متغیر X می‌باشد و برای متغیر Y نیز همین‌طور، که به دو شکل تابع توزیع احتمال (PDF[1]) یا تابع توزیع تجمعی (CDF[2]) قابل نمایش است. توزیع حاشیه‌ای متغیرهای تصادفی X و Y به شکل زیر نشان داده می‌شود (روابط ۱ و ۲).

$$F(x) = P[X \leq x]$$ رابطه (۱)

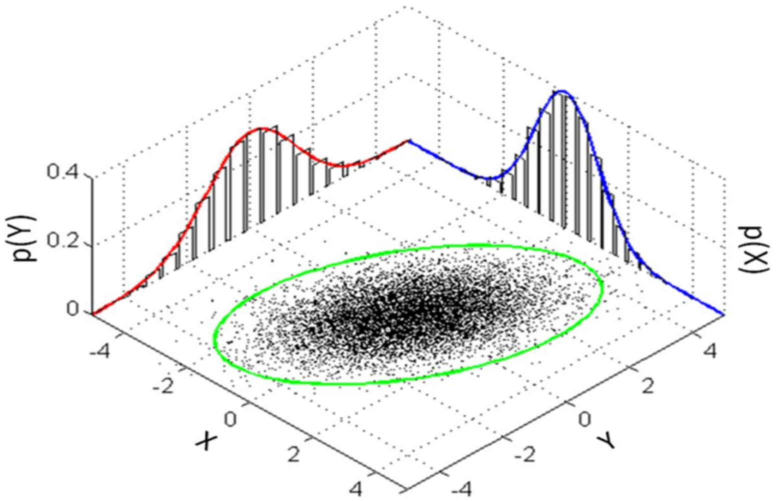
$$G(y) = P[Y \leq y]$$ رابطه (۲)

توزیع توأم: تابع توزیع توأم یک تابع توزیع احتمالاتی است که احتمال این‌که متغیرهای تصادفی X و Y در یک بازه عددی مشخص قرار گیرند را ارائه می‌دهد، به‌عبارت دیگر توزیع توأم از اتصال و ترکیب دو توزیع حاشیه‌ای X و Y حاصل می‌شود که در قالب رابطه ۳ می‌شود.

$$H(x,y) = P[X \leq x, Y \leq y]$$ رابطه (۳)

در شکل ۱ توزیع‌های حاشیه‌ای (رنگ آبی و قرمز) و توزیع توأم (رنگ سبز) قابل مشاهده‌اند.

1- Probability Distribution Function
2- Cumulative Distribution Function



شکل ١- توزیع‌های حاشیه‌ای (رنگ آبی و قرمز) و توأم دومتغیره (رنگ سبز).

کاپولا: کاپولاها به لحاظ کاربرد دکاپولار مدل‌سازی (غیرخطی) روابط چند متغیره در مطالعات هیدرولوژی و هواشناسی مثل آنالیز فراوانی چند متغیره، ارزیابی خطر، مدل‌سازی خشکسالی، سیلاب و درونیابی‌های زمین آماری مورد علاقه محققان واقع شده‌اند.

تاریخچه کاپولا: کلمه کاپولا (یا کوپولا) واژه‌ای لاتین به معنی لینک، اتصال و گره می‌باشد. واژه کاپولا اولین بار در علم آمار و ریاضی توسط اسکلار (١٩٥٩)، به‌عنوان توابع متصل کننده توابع توزیع حاشیه‌ای یک بعدی به‌منظور تشکیل توابع توزیعی توأم چند متغیره، معرفی شدند. اولین مقاله در مورد کاپولا توسط شوایزر و ولف (١٩٨١)، با عنوان وابستگی بین متغیرهای تصادفی منتشر شد. کاپولا توسط جو در سال ١٩٩٧ و نلسون در سال ١٩٩٩ توسعه یافته و همان‌طور که گفته شد از سال ٢٠٠٣ به بعد در علم هیدرولوژی بیشتر به‌کار گرفته شده است.

خصوصیات کاپولا: به‌طور کلی کاپولا یک تکنیک ریاضی انعطاف‌پذیر است که مجموعه‌ای از توابع احتمال تجمعی حاشیه‌ای تک متغیره را به یکدیگر متصل و یک تابع احتمال تجمعی چند متغیره را تولید می‌کند. در واقع کاپولا مبتنی‌بر ارتباط و وابستگی غیرخطی بین متغیرها بوده و پیوند دهنده توزیع توأم و توابع حاشیه‌ای است. از یک نقطه نظر، کاپولا توابعی هستند که توابع توزیعی چند متغیره را به توابع توزیعی حاشیه‌ای یک بعدی آن‌ها اتصال می‌دهند و از طرف دیگر دارای حاشیه‌های یک

بعدی یکنواخت در بازه [۰،۱] می‌باشند. کاپولاها اصولاً قادر به ترکیب هر شکلی از توابع احتمال تجمعی حاشیه‌ای می‌باشند. زیرا برای ساخت یک مدل چند متغیره، توزیع‌های حاشیه‌ای می‌توانند به‌طور مستقل از هم انتخاب شوند و نیازی نیست مانند توابع توزیع دو متغیره، تابع حاشیه‌ای از توزیع خاصی تبعیت کند. مهم‌تر آن‌که کاپولا قادر به تشریح تغییرات درجه همبستگی متغیرها در بخش‌های مختلف توزیع احتمال توأم می‌باشد، که این خصوصیت در سایر روش‌های شبیه‌سازی متغیرهای تصادفی مشاهده نمی‌شود. کاپولا اتصال دهنده توابع توزیع حاشیه‌ای P متغیر تصادفی $\left(F_1(x_1), F_2(x_2), \dots, F_p(x_p)\right)$ و تابع توزیع توأم آن‌ها $(F(x_1, x_2, \dots, x_p))$ بوده، به طوری‌که با داشتن یک تابع توزیع توأم مجموعه توابع حاشیه‌ای محتمل تشکیل دهنده آن قابل ارزیابی است و برعکس (روابط ٤ و ٥).

$$C\left(F_1(x_1), F_2(x_2), \dots, F_p(x_p)\right) = F(x_1, x_2, \dots, x_p) \qquad \text{(٤) رابطه}$$

و

$$F_1(X_1) = U_1 \ , \quad F_2(X_2) = U_2 \ , \quad F_p(X_p) = U_p$$

در نتیجه:

$$C(u_1, u_2, \dots, u_p) = \Pr(U_1 \le u_1, U_2 \le u_2, \dots, U_p \le u_p) \qquad \text{(٥) رابطه}$$

توجه به‌این نکته ضروری است که مقادیر هر کدام از این متغیرها در بازه [۰،۱] قرار می‌گیرند. به‌عبارت دیگر، هر جفت (U_1, U_2) منجربه یک نقطه $F(x), G(y)$ در مربعی به ابعاد واحد [۰،۱] × [۰،۱] می‌شود و این جفت داده، به نوبه خود دارای مقداری در بازه [۰،۱] به‌عنوان توزیع توأم $H(x, y)$ می‌باشند.

خصوصیات و قوانین کلی حاکم بر کاپولا: کاپولا یک تابع توزیعی توأم از متغیرهای تصادفی یکنواخت پیوسته است. یک کاپولای دو متغیره می‌تواند به‌صورت رابطه (٦) بیان شود.

$$C : [0,1]^2 \rightarrow [0,1] \qquad \text{(٦) رابطه}$$

ساده‌ترین نوع کاپولا با مستقل فرض کردن دو متغیر تصادفی دارای شرط‌های زیر می‌باشد که به آن کاپولای مستقل[1] گویند (روابط ٧ و ٨).

1- Independent Copula

$1)\ C(u,v),\ \text{if } v=1 \text{ or } u=1 \Rightarrow C(u,1) \text{ or } C(1,v) = u \text{ or } v$ رابطه (۷)

$2)\ C(u,v),\ \text{if } v=0 \text{ or } u=0 \Rightarrow C(u,0) = C(0,v) = 0$ رابطه (۸)

شرط اول (رابطه ۷ و ۸) بیانگر شرایطی است که تابع کاپولا به کف می‌چسبد و به اصطلاح Grounded می‌شود.

شرط بعدی که صعودی بودن[1] دو متغیر را تضمین می‌کند، بیانگر آنست که در صورت صعودی بودن متغیرهای تصادفی، حجم هر مستطیل ($V_H(B)$) در محدوده تابع کاپولا (I^2 یا I^3) باید بزرگتر یا مساوی صفر باشد؛ در واقع حجم حاصل از یک برش فرضی در فضای دو بعدی یا چند بعدی کاپولا می‌بایست همیشه مثبت باشد. (رابطه ۹).

$3)\ C(u,v)=V_H([0,u]*[0,v])= V(u_2, v_2) - V(u_2, v_1) - V(u_1, v_2) + V(u_1,v_1) \geq 0$ رابطه (۹)

اگر

$u_2 \geq u_1,\ v_2 \geq v_1,\ u_1,\ u_2,\ v_1,\ v_2 \in [0,1]$

اسکلار برای اولین بار از توابع مفصل برای بیان ارتباط توابع توزیع تک بعدی با توابع چند متغیره آن‌ها استفاده کرد (اسکلار، ۱۹۵۹).

قضیه اسکلار: اسکلار نشان داد که برای هر تابع توزیع n بعدی، F می‌تواند به‌صورت رابطه (۱۰) باشد.

$F(x_1, \dots, x_n) = C\,(F_1(x_1), \dots, F_n(x_n))$ رابطه (۱۰)

$F_1,\dots,F_n$ توابع توزیعی حاشیه‌ای هستند. اگر این توابع پیوسته باشند یک تابع کاپولا به نام C وجود دارد که به‌صورت رابطه (۱۱) بیان می‌شود.

$C(u_1,\dots,u_n) = F(F_1^{-1}(u_1), \dots, F_n^{-1}(u_n)),\ 0 \leq u_1, \dots\dots, u_n \leq 1$ رابطه (۱۱)

که در آن $(F_1^{-1}(u_1), \dots, F_n^{-1}(u_n))$ تابع توزیع در مقابل تابع حاشیه‌ای می‌باشد.

1- Increasing

به عبارت دیگر فرض کنید H یک تابع توزیع توأم با توزیع حاشیه‌ای $F(x)$ و $F(y)$ باشد، در این صورت تابع مفصل مانند C وجود دارد به طوری‌که برای هر x و y، توابع توزیع توأم و حاشیه‌ای را به صورت رابطه (۱۲) با هم مرتبط می‌سازد.

$$F(x) = u \quad , \quad F(y) = v \quad \rightarrow$$
$$H(x,y) = C(F(x),F(y))$$ رابطه (۱۲)

$F(x)$ و $F(y)$ توابع توزیع حاشیه‌ای هستند که اگر پیوسته باشند، آنگاه تابع مفصل C یکتاست و به صورت رابطه ۱۳ ارائه می‌شود.

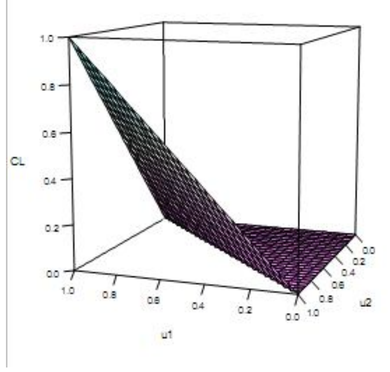
$$C(u, v) = H(F_1^{-1}(u), F^{-1}(v)), \quad 0 \leq u , v \leq 1$$ رابطه ۱۳

که در آن به‌ترتیب u و v توابع حاشیه‌ای $F(x)$ و $F(y)$ می‌باشند (جو، ۱۹۹۷).

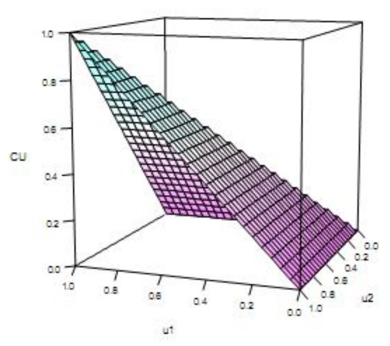
حدود قابل قبول کاپولا: هر تابع کاپولا بایستی توسط دو تابع محدود شود. تمامی توابع کاپولا در فضای خالی بین دو حد بالا مطابق شکل ۲ (رابطه ١٤) و حد پایین مطابق شکل ۳ (رابطه ۱۵) Fréchet - Hoeffding قرار می‌گیرند.

$$C_U(u_1,u_2) = \min\{u_1,u_2\}, u \in [0,1]^2$$ رابطه (١٤)

$$C_L(u_1,u_2) = \max\{0, u_1 + u_2 - 1\}, u \in [0,1]^2$$ رابطه (۱۵)



شکل ۳- حد پایین Fréchet - Hoeffding در کاپولا



شکل ۲- حد بالای Fréchet - Hoeffding در کاپولا

هر کاپولا با حد بالا و پایین Fréchet - Hoeffding به‌صورت رابطه (١٦) محدود می‌شود.

رابطه (١٦) $$C_L(u_1,u_2) \leq C(u_1,u_2) \leq C_U(u_1,u_2), \forall u \in [0,1]^2$$

ضرایب همبستگی: همان‌طور که قبلاً گفته شد از ویژگی‌ها و مزایای مهم استفاده از توابع کاپولا در توزیع‌های چندمتغیره این است که این توابع همبستگی بین متغیرها را در نظر می‌گیرند و در واقع نیازی به استقلال متغیر نیست بلکه حتی این توابع ساختار همبستگی بین متغیرها را نیز لحاظ می‌کنند.

به‌علت این‌که برای برآورد تابع کاپولا، پارامتری به‌نام مولد وجود دارد که مقدار وابستگی را به صورت بدون مقیاس در خود دارد بایستی مقدار ضریب همبستگی مشخص شود. دو ضریب همبستگی تاو کندال و رو اسپیرمن که به‌این منظور مورد استفاده قرار می‌گیرند در ادامه آورده شده‌اند.

ضریب تاو کندال: این ضریب برای اندازه‌گیری درجه تطابق بین دو متغیر، برای مثال مشاهدات جفتی، استفاده می‌شود. اگر تطابق بین دو متغیر عالی باشد، ضریب برابر با یک، و اگر ناهماهنگی بین دو متغیر زیاد باشد ضریب برابر ١- است. برای دسته‌های دیگر، مقادیر بین ١ و ١- می‌باشد. مفهوم ضریب صفر این است که متغیرها کاملا مستقل می‌باشند. محاسبه این ضریب همانند ضریب رو اسپیرمن، در مجموعه‌ای از داده‌ها انجام می‌گیرد. برآورد ضریب تاو کندال برای یک نمونه از مشاهدات n به‌صورت رابطه (١٧) خواهد بود.

رابطه (١٧) $$\hat{\tau} = \frac{C-D}{C+D} = \frac{C \; - \; D}{\frac{n.(n-1)}{2}} = \frac{C \; - \; D}{\left[{n \atop 2}\right]} = \frac{C \; - \; D}{\frac{n!}{2!\,(n-2)!}}$$

که در آن C[1] تعداد جفت‌های هماهنگ و D[2] تعداد جفت‌های ناهماهنگ می‌باشد.

ضریب رو اسپیرمن: یک ضریب غیر پارامتری و بیانگر درجه تطابق بین دو متغیر بوده و مانند تاو کندال در مجموعه‌ای از داده‌ها محاسبه می‌شود. در صورتی‌که تعداد داده‌ها کم و فرض نرمال بودن آن‌ها معقول نباشد، از آن استفاده می‌شود. به‌عبارت دیگر ضریب همبستگی که بر اساس رتبه داده‌ها محاسبه می‌شود، با استفاده از رابطه اسپیرمن محاسبه می‌شود (رابطه ١٨).

رابطه (١٨) $$\hat{\rho} = 1 - \frac{6\sum(di^2)}{n(n^2-1)}$$

که di اختلاف بین رتبه داده‌ها می‌باشد.

1- Concordance
2- Discordance

انواع کاپولا: کاپولاها انواع گوناگونی دارند که به‌طور کلی در دو دسته پارامتریک و ناپارامتریک تقسیم‌بندی می‌شوند. ارجحیت کاپولاهای پارامتریک در استفاده از پارامتر بوده، ازینرو در این مقاله و سایر تحقیقات انجام شده مورد توجه قرار گرفته‌اند. در واقع برازش کاپولا با داده‌های ورودی به کمک تخمین این پارامترها امکان‌پذیر است، نکته‌ای که محدودیتی برای کاپولاهای ناپارامتریک محسوب می‌شود. این پارامترها نماینده شدت وابستگی متغیرها می‌باشند و رابطه ریاضی معینی با آن دارند. بنا به استفاده از پارامترهای متفاوت در کاپولاهای پارامتریک مختلف، نتایج نیز با هم متفاوت خواهند بود (دورلمان، ۲۰۰۰).

هشت نوع از توابع دو متغیره یک پارامتری مرسوم در مباحث هیدرولوژیکی به‌همراه تابع چگالی احتمال مربوطه در جدول ۱ آورده شده‌اند.

جدول ١- توابع کاپولای دومتغیره یک پارامتری پرکاربرد در هیدرولوژی (دورلمان، ٢٠٠٠).

نام تابع مفصل	رابطه تابع مفصل	تابع چگالی احتمال	ملاحظات
Ali-Mikhail-Haq	$C(u,v)=\frac{uv}{1-\theta(1-u)(1-v)}$	$C(u,v)=\frac{[1-\theta(1-u)(1-v)](1-\theta)+2uv}{[1-\theta(1-u)(1-v)]^{3}}$	$-1\le\theta\le1$
Galambos	$C(u,v)=uv\exp\{[(-\ln u)^{-\theta}+(-\ln v)^{-\theta}]^{\frac{1}{\theta}}\}$	$C(u,v)=\frac{C(u,v)}{uv}\{1-[(-\ln u)^{-\theta}+(-\ln v)^{-\theta}]^{-\frac{1}{\theta}-1}\times[(-\ln u)^{-\theta-1}+(-\ln v)^{-\theta-1}]+[(-\ln u)^{-\theta}+(-\ln v)^{-\theta}]^{-\frac{1}{\theta}-2}\times[(-\ln u)(-\ln v)]^{-\theta-1}[1+\theta+[(-\ln u)^{-\theta}+(-\ln v)^{-\theta}]^{\frac{1}{\theta}}]\}$	$\theta\ge0$
Gumble-Hougaard	$C(u,v)=\exp\{-[(-\ln u)^{\theta}+(-\ln v)^{\theta}]^{\frac{1}{\theta}}\}$	$C(u,v)=C(u,v)\frac{[(-\ln u)(-\ln v)]^{\theta-1}}{uv}[(-\ln u)^{\theta}+(-\ln v)^{\theta}]^{\frac{2}{\theta}-2}\times\left\{(\theta-1)[(-\ln u)^{\theta}+(-\ln v)^{\theta}]^{-\frac{1}{\theta}}+1\right\}$	$\theta\ge1$
Clyton	$C(u,v)=(u^{\theta}+v^{-\theta}-1]^{\frac{1}{\theta}}\}$	$C(u,v)=(\theta+1)(u^{-\theta}+v^{-\theta}-1]^{-\frac{1}{\theta}-2}(uv)^{-\theta-1}$	$\theta\ge0$
Frank	$C(u,v)=-\frac{1}{\theta}\ln[1+\frac{(e^{-\theta u}-1)(e^{-\theta v}-1)}{e^{-\theta}-1}$	$C(u,v)=-\frac{\theta e^{-\theta(u+v)}(e^{-\theta}-1)}{[e^{-\theta(u+v)}-e^{-\theta u}-e^{-\theta v}+e^{-\theta}]^{2}}$	$\theta\neq0$
Plackett	$C(u,v)=\frac{1}{2(\theta-1)}\{1+(\theta-1)(u+v)-\times\left[\left(1+(\theta-1)(u+v)\right)^{2}-4\theta(\theta-1)uv\right]^{\frac{1}{2}}\}$	$C(u,v)=[\left(1+(\theta-1)(u+v)\right)^{2}-4\theta(\theta-1)uv]^{\frac{3}{2}}\times\theta[1+(\theta-1)(u+v-2uv)]$	$\theta\ge0$
joe	$C(u,v)=1-[(1-u)^{\theta}+(1-v)^{\theta}-(1-u)^{\theta}(1-v)^{\theta}]^{\frac{1}{\theta}}$	$C(u,v)=[(1-u)^{\theta}+(1-v)^{\theta}-(1-u)^{\theta}(1-v)^{\theta}]^{\frac{1}{\theta}2}\times[(1-u)(1-v)]^{\theta-1}[(\theta-1)+(1-u)^{\theta}+(1-v)^{\theta}-(1-u)^{\theta}(1-v)^{\theta}]$	$\theta\ge0$
Fatlie-Gumble-Morgenstern	$C(u,v)=uv[1+\theta(1-u)(1-v)]$	$C(u,v)=1+\theta(1-2u)(1-2v)$	$-1\le\theta\le1$

کاپولاهای ارشمیدسی، بیضوی و مقادیر حدی، از پرکاربردترین توابع کاپولا به‌شمار می‌روند. در این مقاله با توجه به کاربرد بیش‌تر کاپولاهای ارشمیدسی و بیضوی، به تشریح این دسته از توزیع‌ها پرداخته می‌شود.

کاپولای ارشمیدسی: به‌دلیل سهولت ساختار و متقارن بودن یکی از توابع مهم و کاربردی کاپولا می‌باشند. معادلات آن به‌صورت رابطه (۱۹) می‌باشد.

$$C(u,v) = \varphi^{-1}(\varphi(u) + \varphi(v))$$ رابطه (۱۹)

پارامتر فی (φ) به‌عنوان مولد کاپولا شناخته می‌شود که پیوسته، محدب و غیر منفی می‌باشد. رابطه کلی بین تاو کندال و این مولد برای مجموعه داده‌های دو متغیره می‌تواند به‌صورت رابطه (۲۰) بیان می‌شود.

$$\tau = 1 + 4\int_0^1 \frac{\varphi(t)}{\varphi'(t)}dt$$ رابطه (۲۰)

$$t = \upsilon \ , v$$

توابع کلایتون، فرانک و گامبل مهم‌ترین توابع کاپولای ارشمیدسی می‌باشند.

کاپولای فرانک: از خانواده کاپولای ارشمیدسی و به‌صورت متقارن بوده که توسط یک پارامتر تتا (Θ) تعریف می‌گردد (شکل ٤). پارامتر تتا در کاپولای فرانک با ضرایبی که وابستگی متغیرها را نشان می‌دهد (مثل ضریب همبستگی اسپیرمن)، در ارتباط است.

تابع کاپولای فرانک به‌صورت رابطه (۲۱) محاسبه می‌شود.

$$C(u,v;\theta) = -\frac{1}{\theta}log[1 + \frac{(e^{-\theta u}-1)(e^{-\theta v}-1)}{e^{-\theta}-1}]$$ رابطه (۲۱)

مقدار مولد تابع فرانک از رابطه (۲۲) محاسبه می‌شود.

$$\varphi_\theta(t) = -ln[\frac{e^{-\theta t}-1}{e^{-\theta}-1}]$$ رابطه (۲۲)

که تتا در بازه منفی بی‌نهایت و مثبت بی‌نهایت (به جز صفر) قرار دارد.

رابطه بین ضریب تاو کندال و پارامتر تتای کاپولای فرانک در روابط (۲۳) و (٢٤) نشان داده شده است.

$$\frac{[D_1(\theta)-1]}{\theta} = \frac{1-\tau}{4}$$ رابطه (۲۳)

$$D_1(\theta) = \frac{1}{\theta}\int_0^\theta \frac{t}{e^t-1}dt$$ رابطه (٢٤)

کاپولای کلایتون: کاپولای کلایتون یک کاپولای ارشمیدسی نامتقارن بوده و بیانگر وابستگی بیش‌تر در دنباله منفی نسبت به دنباله مثبت است (شکل ٥).
کاپولای کلایتون از طریق رابطه (٢٥) محاسبه می‌شود.

$$C_\theta(u,v) = max[u^{-\theta} + v^{-\theta} - 1\,,0]^{-\frac{1}{\theta}}$$ رابطه (٢٥)

و مولد آن از رابطه (٢٦) به‌دست می‌آید.

$$\varphi_\theta(t) = \frac{1}{\theta}(t^{-\theta} - 1)$$ رابطه (٢٦)

که تتا در بازه ١- و بی‌نهایت (به‌جز صفر) قرار می‌گیرد.
رابطه بین تاو کندال و پارامتر تتای کاپولای کلایتون به‌صورت معادله (٢٧) محاسبه می‌شود.

$$\theta = \frac{2\tau}{1-\tau}$$ رابطه (٢٧)

کاپولای گامبل: یکی از کاپولاهای ارشمیدسی نامتقارن بوده و نشان‌دهنده وابستگی بیشتر در دنباله مثبت نسبت به دنباله منفی می‌باشد (شکل ٦).
کاپولای گامبل به صورت زیر محاسبه می‌شود (رابطه ٢٨).

$$C_\theta(u,v) = e^{\{-[(-\ln u)^\theta + (-\ln v)^\theta]^{\frac{1}{\theta}}\}}$$ رابطه (٢٨)

و مولد آن از رابطه (٢٩) قابل محاسبه است.

$$\varphi_\theta(t) = (-\ln t)^\theta$$ رابطه (٢٩)

که مقدار تتا در بازه یک و بی‌نهایت قرار می‌گیرد.
رابطه ضریب تاو کندال و پارامتر تتای گامبل از رابطه (٣٠) محاسبه می‌شود.

$$\hat{\theta} = \frac{1}{1-\tau}$$ رابطه (٣٠)

در فرمول‌های فوق θ مربوط به جامعه داده‌ها و $\hat{\theta}$ از روی یک نمونه از داده‌ها محاسبه می‌شود.

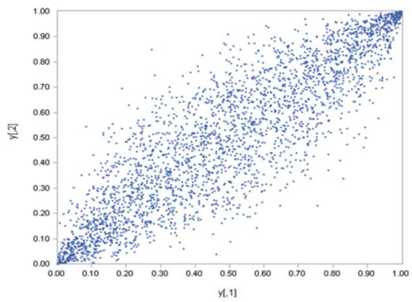
کاپولای بیضوی: مزیت کلیدی این کاپولا این است که قادر به نمایش همبستگی بین توزیع‌های حاشیه‌ای در سطوح مختلف بوده و عیب اصلی آن نامحدود بودن روابط و شکل تقارن شعاعی آن می‌باشند. در کاپولاهای بیضوی روابط بین ضریب همبستگی خطی رو (ρ) و تاو کندال (τ) به‌صورت زیر می‌باشد (رابطه ٣١).

$$\rho(x,y) = sin(\frac{\pi}{2}\tau)$$ رابطه (٣١)

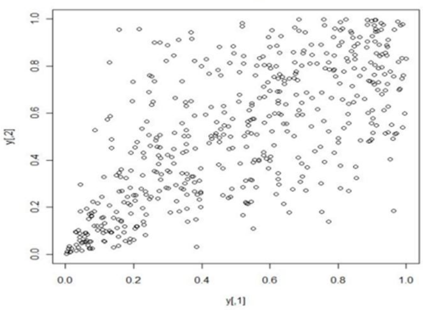
کاپولای نرمال یا گوسن: شایع‌ترین شکل کاپولای بیضوی، کاپولاهای چندمتغیره نرمال بوده که با استفاده از رابطه (۳۲) محاسبه می‌گردد. (شکل ۷).

$$C_\rho(u,v) = \int_{-\infty}^{F^{-1}(u)} \int_{-\infty}^{F^{-1}(v)} \frac{1}{2\pi(1-\rho^2)^{0.5}} exp\left\{-\frac{x^2-2\rho xy+y^2}{2(1-\rho^2)}\right\} dxdy \qquad \text{رابطه (۳۲)}$$

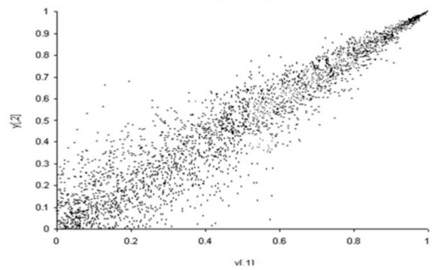
اگر F^{-1} معکوس توزیع نرمال و استاندارد تک متغیره باشد، ρ ضریب همبستگی خطی و پارامتر کاپولای نرمال است.



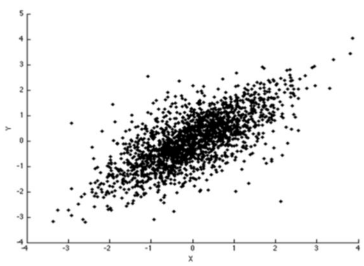
شکل ۴- کاپولای فرانک از خانواده توابع کاپولای ارشمیدسی



شکل ۵- کاپولای کلایتون از خانواده توابع کاپولای ارشمیدسی



شکل ۶- کاپولای گامبل از خانواده توابع کاپولای ارشمیدسی



شکل ۷- کاپولای نرمال یا گوسن از خانواده توابع کاپولای

برآورد پارامتر و آزمون نکویی برازش توابع مفصل: تابع کاپولا یا مفصل مناسب برای یک کاربرد خاص، تابعی است که به بهترین نحو ممکن همبستگی بین داده‌ها را نشان می‌دهد. برای انتخاب مناسب‌ترین مفصل، در مرحله اول، توزیع‌های مختلف تک متغیره بر متغیرهای حاشیه‌ای برازش داده شده و پارامترهای توزیع با روش حداکثر درست‌نمایی برآورد می‌شود. یکی از مهم‌ترین روش‌های

برآورد پارامتر توابع مفصل، روش حداکثر درست‌نمایی است. هر تابع مفصلی که دارای بیش‌ترین مقدار حداکثر لگاریتم درست‌نمایی باشد، مناسب‌ترین تابع شناخته می‌شود.

به‌طور کلی با برازش تابع تئوری توزیع به داده‌های موردنظر و اثبات نکویی برازش، با حل معادله توزیع می‌توان مقدار متغیر را به ازای احتمالات مختلف به‌دست آورد. امروزه نرم‌افزارهای آماری مانند SPSS، SAS و Excel، علاوه‌بر برازش و ارائه نمایه‌های نکویی برازش، مقادیر متغیر را به ازای آن توزیع تئوری در احتمالات مختلف محاسبه می‌کنند. حل توابع تئوری توزیع‌های احتمال نیز به‌دلیل نامعین بودن محدوده انتگرال ساده نبوده و معمولاً به روش‌های عددی یا با استفاده از جداولی که برای این منظور تهیه شده‌اند، صورت می‌گیرد.

برای برازش توزیع‌های تئوری روی داده‌های هیدرولوژیکی روش‌های گوناگون وجود دارد که چهار روش آن در هیدرولوژی کاربرد زیادی دارند. این روش‌ها عبارتند از:

- روش استفاده از پارامترهای توزیع
- روش ضریب فراوانی
- روش گرافیکی
- روش حداقل مربعات

در بین روش‌های فوق ساده‌ترین روش، استفاده از پارامترهای توزیع می‌باشد. برای این کار از روش‌های احتمالات گشتاور[1] و حداکثر درست‌نمایی[2] استفاده می‌شود. روش گشتاور ساده است اما نتایج حاصله از آن به‌خصوص اگر تعداد داده‌ها کم باشد، از دقت کم‌تری برخوردار است، برعکس روش حداکثر درست نمایی از دقت زیادی برخوردار بوده، اما محاسبات آن بسیار پیچیده و وقت‌گیر است.

در روش حداقل مربعات، ابتدا داده‌ها به یک تابع توزیع برازش داده می‌شود. سپس اختلاف بین ارقام حاصل از تابع و داده‌های مشاهده‌ای محاسبه و تابعی که این اختلاف در آن حداقل باشد انتخاب می‌گردد. ممکن است داده‌ها با چند توزیع احتمال برازش داده شده باشند، اما برای انتخاب بهترین

1- Moment
2- Maximum likelihood

آن‌ها بایستی از آزمون‌های نکوئی برازش[1] مثل آزمون کای مربع[2] یا X^2 و آزمون کلموگروف-اسمیرنوف[3] یا K- S استفاده کرد.

نتیجه‌گیری

در این مقاله به بررسی و شناخت الگوریتم‌های محاسباتی، پارامترها و شکل توزیع توابع کاپولای پرکاربرد (توابع ارشمیدسی و بیضوی) و کاربرد و جایگاه آن در هیدرولوژی استوکاستیک پرداخته شد. با توجه به تصادفی و وابسته بودن متغیرهای هیدرولوژیکی مفید بودن استفاده از روش‌های چندمتغیره کاپولا به لحاظ در نظرگرفتن این وابستگی در تمام بخش‌های توزیع مفصل، امری انکارنشدنی به‌نظر می‌رسد. از طرفی بررسی تحقیقات انجام شده نیز موید این موضوع بوده و قطعاً توسعه کاربرد تابع کاپولا به مدیران و برنامه‌ریزان سیستم‌های هیدرولوژیکی در راستای شناخت صحیح فرآیندها و ارائه راهکارها و استراتژی‌های مناسب خواهد کرد.

به‌طور کلی می‌توان مراحل اصلی یک تحقیق مبتنی‌بر کاپولا (در این‌جا کاپولای ارشمیدسی) را به شرح زیر خلاصه کرد:

۱. آنالیز متغیرها و توزیع‌های حاشیه‌ای و بررسی شرایط کاپولا از جمله خصوصیات و حدود آن
۲. تعیین همبستگی غیر پارامتری متغیرها از جمله تعیین تاو کندال
۳. برآورد پارامتر همبستگی تتا با استفاده از تاو کندال برای انواع مختلف کاپولای ارشمیدسی (فرانک، کلایتون، گامبل و غیره)
٤. تعیین تابع مولد کاپولا با استفاده از تتا
۵. تعیین تابع توزیع احتمال کاپولا برای انواع مختلف کاپولای ارشمیدسی
٦. شبیه‌سازی تصادفی با هر یک از کاپولاها که خصوصیات همبستگی بین متغیرها را در خود حفظ کرده‌اند.
۷. برازش کاپولاهای تعیین شده بر داده‌ها
۸. آزمون نکویی برازش جهت انتخاب بهترین کاپولا
۹. استفاده از کاپولای برتر جهت تخمین

1- Goodness of fit test
2- Chi square
3- Kolmogorov- Smirnov

رهیافت‌های ترویجی: پدیده‌های پیچیده در هیدرولوژی معمولاً با متغیرهای تصادفی وابسته متعددی مشخص می‌شوند. روش‌های کلاسیک بیشتر بر اساس تحلیل‌های یک متغیره انجام می‌گرفته که این روش تحلیل منجربه بررسی ناقص پدیده‌ها می‌شود. به‌همین دلیل مدل‌های چندمتغیره کاپولا به‌علت انعطاف‌پذیر بودن و کاربرد آسان در هیدرولوژی مورد استقبال قرار گرفته است، به طوری‌که در موضوعات مختلف مرتبط با هیدرولوژی مورد استفاده قرار می‌گیرد. به‌علاوه نشان دادن تغییرات همبستگی متغیرها در بخش‌های مختلف توزیع احتمال توأم، سبب شده توابع کاپولا به سایر روش‌های تحلیل متغیرهای تصادفی ارجحیت داشته باشد. لذا لازم است در موضوعات مختلف در علم هیدرولوژی همچون بررسی فراوانی بارش، فراوانی سیلاب، فراوانی خشکسالی، هیدروگراف سیلاب، مدلسازی بارش وغیره به‌جای استفاده از توابع تک متغیره، با دقت نظر بیشتر از توابع چندمتغیره کاپولا استفاده شود.

منابع

1. Abbasian, M.S., and Moosavi, S. 2013. Joint analysis of peak discharge and runoff volume using Copula Functions. 7^{th} National Congress on Civil Engineering. Sistan and Baluchestan University. 8p. (In Persian)
2. Abdolhosseini, M. 2011. Multivariate frequency analysis of hydrologic events using copula for both stationary and non-stationary conditions. PhD Thesis. Isfahan University of Technology. 232p. (In Persian)
3. Alizadeh, A. 2010. Principles of applied hydrology. Emam Reza University Press. 26^{th} Edition. 800p. (In Persian)
4. Bruneau, P., Ashkar, F., and Bobe´e, B. 1994. Simplnorm: Un mode`le simple pour obtenir les probabilite´s conjointes de deux de´bits et le niveau qui en de´pend, Canadian Journal of Civil Engineering, 5: 883–895.
5. Chebana, F., and Ouarda, T.B.M.J. 2009. Index flood–based multivariate regional frequency analysis. Water Resources Research, 45(10): 1-15.
6. Chen, L., Singh, V.P., Guo, S., Hao, Z., and Li, T. 2012. Flood coincidence risk analysis using multivariate Copula functions. Journal of Hydrologic Engineering, 17(6): 742-755.
7. De Michele, C., Salvadori, G., Canossi, M., Petaccia, A., and Rosso, R. 2005. Bivariate statistical approach to check adequacy of dam spillway. Journal of Hydrologic Engineering, 10(1): 50–57.
8. Durrleman, V., Nikeghbali, A., and Roncalli, T. 2000. Which copula is the right one? Groupe de Recherche Ope'rationnelle, Cre'dit Lyonnais, France 89p.

9. Favre, A.C., El Adlouni, S., Perreault, L., Thiemonge, N., and Bobee, B. 2004. Multivariate hydrological frequency analysis using Copulas. Water Resource Research, 40: 1-12.
10. Frees, E.W., and Valdez, E.A. 1998. Understanding relationships using Copulas. Actuarial, Journal, 2(1): 1–25.
11. Genest, C., Favre, A.C., Beliveau, J., and Jacques, C. 2007. Metaelliptical copulas and their use in frequency analysis of multivariate hydrological data. Water Resources Research, 43(9): 1-12.
12. Joe, H. 1997. Multivariate models and dependence concepts. Chapman and Hall. New York. CRC Press, 424p.
13. Kao, S.C., and Govindaraju, R.S. 2008. Trivariate statistical analysis of extreme rainfall events via the Plackett family of Copulas. Water Resources Research, 44(2): 1-11.
14. Marco, J.B., Harboe, R., and Salas, J.D. 1993. Stochastic hydrology and its use in water resources systems simulation and optimization, kluwer academic publisher. 237p.
15. Mirabbasi Najafabadi, R., Fakheri-Fard, A., Dinpashoh, Y., and Eslamian, S. 2014. Long term drought monitoring of urmia using joint deficit index (JDI). Journal of Water and Soil Science, 23(4): 87-103p. (In Persian)
16. Mirakbari, M., and Ganji, A. 2012. Bivariate analysis of drought severity–duration (Case study: Kermanshah province). Iranian Journal of Water Research, 6(11): 9-17. (In Persian)
17. Omidi, M., Mohammadzadeh, M., and Morid, S. 2010. The probabilistic analysis of drought severity– duration in Tehran Province using Copula Functions. Iranian Journal of Soil and Water Research, 42(1): 95-101. (In Persian)
18. Raynal-Villasenor, J.A., and Salas, J.D. 1987. Multivariate extreme value distributions in hydrological analyses, in Water for the future. hydrology in Perspective, 111 –119.
19. Salvadori, G., and De Michele, C. 2006. Statistical characterization of temporal structure of storms. Advances in Water Resources, 29(6): 827–842.
20. Salvadori, G., and Michele, C. 2004. Frequency analysis via Copula: Theoretical aspects and applications to hydrological events. Water Resource Research, 40(12): 1-17.
21. Scaillet, O. 2005. Kernel based goodness-of-fit tests for Copulas with fixed smoothing parameters. FAME Research Paper Series rp145. International Center for Financial Asset Management and Engineering. Available at http://ideas. repec. org/p/fam/rpseri/rp145. html.
22. Schweizer, B., and Wolff, E.F. 1981. On nonparametric measures of dependence for random variables. the annals of statistics, Pp: 879-885.

23. Serinaldi, F. 2008. Copula-based mixed models for bivariate rainfall data: An empirical study in regression perspective. Stochastic Environmental Research and Risk Assessment, 23(5): 677-693.
24. Serinaldi, F. 2009. A multisite daily rainfall generator driven by bivariate Copula-based mixed distributions. Journal of Geophysical Research Atmospheres, 114(10): 1795-1782.
25. Shiau, J.T. 2006. Fitting drought duration and severity with two-dimensional Copulas. Water Resources Management, 20: 759-815.
26. Shiau, J.T., Wang, H.Y., and Tsai, C.T. 2006. Bivariate frequency analysis of floods using Copulas. Journal of American Water Resources Association, 42(6): 1549–1564.
27. Sklar, A., 1959. Distribution functions, dimensions and margins. Publications of the Institute of Statistics, Pp: 229-231.
28. Song, S., and Singh, V.P. 2010. Frequency analysis of droughts using the Plackett Copula and parameter estimation by genetic algorithm. Stochastic Environmental Research and Risk Assessment, 24: 783–805.
29. Song, S., and Singh, V.P. 2010. Meta-elliptical Copulas for drought frequency analysis of periodic hydrologic data. Stochastic Environmental Research and Risk Assessment, 24: 425–444.
30. Wang, W., and Wells, M.T. 2000. Model selection and semi-parametric inference for bivariate failure-time data. Journal of the American Statistical Association, 95: 62- 72.
31. Wong, G., Lmbert, M.E., and Metcalfe, A.V. 2008. Trivariate Copulas for characterization of droughts. Journal of Anziam, 49: 306-323.
32. Yue, S., Ouarda, T.B.M.J., and Bobe´e, B. 2001. A review of bivariate gamma distribution for hydrological application. Journal of Hydrology, 246(1): 1–18.
33. Zhang, J., Ding, Z., and You, J. 2014. The joint probability distribution of runoff and sediment and its change characteristics with multi-time scales. Journal of Hydrology and Hydromechanics, 62(3): 218–225
34. Zhang, L., and Singh, V.P. 2007. Bivariate rainfall frequency distributions using Archimedean Copulas. Journal of Hydrology, 332: 93-109.
35. Zhang, Q., Li, J., Singh, V.P., and Xu, C.Y. 2012. Copula-based spatio-temporal patterns of precipitation extremes in China. International Journal of Climatology, 35(5): 1140-1152.

Gorgan University of Agricultural Sciences and Natural Resources
J. of Conservation and Utilization of Natural Resources, Vol. 4 (2), 2015
http://ejang.gau.ac.ir

Copula functions and their application in stochastic hydrology

A. Bahremand[1], E. Alvandi[2], M. Bahrami[2], M. Dashti Marvili[2], H. Heravi[2], *Gh.R. Khosravi[2], *A. Kornejady[2], H. Samadi Arghini[2], M. Tajiki[2] and M. Teimouri[2]

[1]Associate Professor, Department of Watershed and Arid Zone Management, Gorgan University of Agricultural Sciences and Natural Resources, Gorgan, Iran, [2]Ph.D. Student of Watershed Management Science and Engineering, Gorgan University of Agricultural Sciences and Natural Resources, Gorgan, Iran
Received: 2015/05/11 ; Accepted: 2015/07/22

Abstract

In hydrology, numerous variables serve as representative for system behavior to model processes. Assuming these variables as independent may question accuracy of modeling final results. On the other hand several variables may involve in some of the hydrological issues and phenomena (floods, droughts, etc.), and affect the phenomenon in a manner that these variables can be interrelated. Applying statistical distribution for such variables regardless of this correlation increases uncertainty. Therefore, understanding the correlation between the marginal distributions of different variables in order to perceive the laws governing these dependences may be very promising to identify the observed hydrological events. Therefore, In order to increase reliability of analysis results, the multivariate statistical approach must be applied. Conventional methods for multivariate analysis involve classical multivariate distribution functions. But while using such functions, marginal distributions should be known and equal. Hence, applying such methods is limited to some extent. More suitable method for multivariate analysis that has overcome the limitations of classical multivariate functions is the utilization of Joint (Copula) functions. Recent developments in applied mathematics' studies have shown that Joint functions are useful tools to study the statistical behavior of dependent variables. Hence, the present paper aims to shed lights on the history and usages of common Copula functions in stochastic hydrology for multivariate analyzing of hydrological processes.

Keywords: Copula, Joint function, Bivariate distributions, Stochastic hydrology

---

*Corresponding author: aidin.kornejady@gmail.com