

## بررسی ساختار وابستگی و شبیه‌سازی عملکرد گندم دیم شهرستان میانه و متغیرهای آب و هوایی: کاربرد رهیافت تابع مفصل تاکی شکل کانونی (C-Vine)

اسماعیل پیش‌بهار<sup>۱</sup>، سحر عابدی<sup>۲</sup>، قادر دشتی<sup>۳</sup>، علی کیانی‌راد<sup>۴</sup>

تاریخ دریافت: ۱۳۹۴/۶/۸ تاریخ پذیرش: ۱۳۹۵/۶/۲۹

### چکیده

بیمه آب و هوا از بهترین سیستم‌های بیمه‌ای است که در آن به منظور محاسبه حق بیمه، به شبیه‌سازی عملکرد محصول با توجه به متغیرهای آب و هوایی با تعیین ساختار وابستگی آن‌ها نیاز است. بررسی ساختار وابستگی با در نظر گرفتن تأثیرگذاری هم‌زمان متغیرها با استفاده از توابع مفصل تاکی شکل می‌تواند نتایج مناسب‌تری ارائه کند. بنابراین، در پژوهش حاضر، ساختار وابستگی بین متغیرهای آب و هوایی و عملکرد گندم دیم با استفاده از توابع مفصل تاکی شکل کانونی (C-vine) اندازه‌گیری شد، سپس عملکرد گندم دیم در شهرستان

۱. دانشیار اقتصاد کشاورزی، دانشکده کشاورزی، دانشگاه تبریز (نویسنده مسئول) pishbahar@yahoo.com

۲. دانشجوی کارشناسی ارشد اقتصاد کشاورزی، دانشکده کشاورزی، دانشگاه تبریز

۳. استاد اقتصاد کشاورزی، دانشکده کشاورزی، دانشگاه تبریز

۴. استادیار مؤسسه پژوهش‌های برنامه‌ریزی، اقتصاد کشاورزی و توسعه روستایی

## اقتصاد کشاورزی و توسعه - سال بیست و ششم، شماره ۱۰۳

میان‌ه شبیه‌سازی گردید. نتایج مطالعه حاضر نشان داد که اکثر متغیرهای مورد بررسی دارای وابستگی دنباله‌ای در دم پایین‌اند؛ به عبارتی، تأثیرپذیری آن‌ها در مقادیر کوچک بیشتر از مقادیر بزرگ است. به این ترتیب، مقدار متوسط شبیه‌سازی شده عملکرد گندم دیم ۸۷۱/۶۸۳۹ کیلوگرم در هکتار به دست آمد که این نتایج می‌تواند در محاسبه خسارت مورد انتظار و حق بیمه مورد استفاده قرار گیرد.

طبقه‌بندی JEL: G22، J65، N55، O13، Q10

**کلیدواژه‌ها:** بیمه آب و هوا، شبیه‌سازی عملکرد، تابع مفصل تاکی‌شکل، گندم دیم، میان‌ه

### مقدمه

گندم به عنوان محصولی محوری و کلیدی جایگاه ویژه‌ای در تولید و مصرف مواد غذایی در ایران دارد و تأمین این محصول برای کشور به معنی ایجاد امنیت غذایی و رفاه اجتماعی می‌باشد. لذا این محصول از تولیدات اساسی و راهبردی کشور و محور سیاست‌های خودکفایی دولت به شمار می‌آید و در نتیجه، بررسی اقتصادی آن اهمیت فراوانی دارد. براساس گزارش وزارت جهاد کشاورزی سطح زیرکشت گندم در ایران حدود ۶/۲۴۱ میلیون هکتار برآورد شده که بیشترین سطح اراضی زراعی را دربرمی‌گیرد. سهم اراضی آبی ۳۶/۷۵ درصد و سهم اراضی دیم ۶۳/۲۵ درصد است. استان آذربایجان شرقی با سهم ۶/۸ درصدی از کل سطح زیرکشت گندم رتبه دوم کشور را بعد از استان خوزستان به خود اختصاص داده است. در این میان شهرستان میانه، بزرگ‌ترین تولیدکننده گندم دیم در استان آذربایجان شرقی است و ۲۰ درصد سطح زیرکشت و ۲۰ درصد تولید این استان مربوط به این شهرستان است (۲۲). در استان آذربایجان شرقی و شهرستان میانه بیش از ۷۰ درصد اراضی گندم به صورت دیم کشت می‌شوند، بنابراین طبیعی است که کشت این محصول مانند دیگر نقاط کشور بیش از هر عاملی به شدت تحت تأثیر شرایط آب و هوایی قرار گیرد. همین امر سبب

بررسی ساختار وابستگی .....

گردیده کشت این محصول با طیف گسترده‌ای از خطرات و حوادث طبیعی مواجه باشد و فعالیتی پرخطر و توأم با ریسک محسوب گردد.

علی‌رغم اینکه بیمه ابزار مناسبی برای کاهش ریسک است ولی هم هزینه‌بر می‌باشد و هم مشکلاتی نظیر اطلاعات نامتقارن<sup>۵</sup> (شامل مشکلات ناشی از مخاطرات اخلاقی و انتخاب نامساعد) باعث می‌شود که کارایی لازم در مدیریت ریسک را نداشته باشد. تجربیات کشورهای پیشرفته و در حال توسعه همچون هلند، سوئیس، هند، اوکراین، اتیوپی، تایلند و آمریکای مرکزی نشان داده است که تغییر نوع بیمه از سنتی به بیمه شاخص محور بسیاری از مشکلات موجود در نظام سنتی را برطرف کرده و منجر به کاهش ریسک کشاورزان شده است. یکی از انواع بیمه شاخص محور «بیمه محصولات کشاورزی براساس شاخص‌های آب و هوایی» است. در بیمه آب و هوا برخلاف طرح‌های سنتی حق بیمه و غرامت براساس مقدار شاخص و تأثیر آن بر افت محصول تعیین می‌گردد و از آنجا که این شاخص‌ها براساس منابع داده‌های شفاف و مشخص می‌باشند، بسیاری از مشکلات ناشی از اطلاعات نامتقارن را برطرف می‌کنند. این نوع بیمه دارای آستانه و حد است؛ یعنی برای متغیر یا شاخص آب و هوایی یک محدوده مشخص می‌شود و پرداخت خسارت در این محدوده صورت می‌گیرد. آستانه شاخص آب و هوا نقطه شروع خسارت بوده که با نزدیک شدن به مقدار حد مبلغ خسارت زیاد می‌شود. پرداخت خسارت متناسب با توزیع شرطی توأم متغیر عملکرد و شاخص‌های آب و هوایی صورت می‌گیرد. به بیان دیگر در این سیستم بیمه‌ای به منظور محاسبه خسارت مورد انتظار و حق بیمه نیاز به «شناخت متغیرهای آب و هوایی مختلف مؤثر در رشد گیاه و شبیه‌سازی و پیش‌بینی عملکرد محصول قبل از فصل برداشت با توجه به متغیرهای آب و هوایی می‌باشد» (۵).

انتخاب متغیرهای آب و هوایی در منطقه برای محصول مورد نظر معمولاً براساس مطالعات قبلی و اندازه‌گیری همبستگی با استفاده از ضریب همبستگی تاو کندال صورت می‌گیرد. در مرحله بعد نیاز است که مشخص شود با توجه به مقادیر مختلف متغیر آب و

5. Asymmetric Information

6. Weather-Based Crop Insurance Scheme (WBCIS)

#### اقتصاد کشاورزی و توسعه - سال بیست و ششم، شماره ۱۰۳

هوایی چه عملکردی اتفاق می‌افتد؛ به عبارتی باید مقادیر عملکرد را با توجه به ساختار وابستگی که با متغیرهای آب و هوایی دارد، قبل از فصل برداشت شبیه‌سازی و پیش‌بینی نمود. به منظور شبیه‌سازی و پیش‌بینی محصولات کشاورزی با استفاده از متغیرهای آب و هوایی، روش‌های مختلفی در مطالعات گسترده در داخل و خارج از کشور اتخاذ شده است که از جمله مطالعات داخلی می‌توان به مطالعه فرج‌زاده و زرین (۱۳) اشاره نمود که با استفاده از چند پارامتر اقلیمی، عملکرد گندم دیم در استان آذربایجان غربی را مدل‌سازی نمودند. عزیزی و یاراحمدی (۲)، کمالی و بازگیر (۱۹)، فرج‌زاده و همکاران (۱۴) و زارع ایبانه (۲۶) نیز در مطالعات جداگانه و مشابهی اقدام به پیش‌بینی عملکرد محصولات کشاورزی نمودند. در این مطالعات از مدل‌های رگرسیونی ساده و غیرترکیبی به منظور پیش‌بینی استفاده گردید. اما در مطالعات خارجی مخصوصاً در دهه گذشته از توابع مفصل (کاپیولا)، به منظور تبیین ساختار وابستگی عملکرد محصولات با متغیرهای آب و هوایی و در نهایت پیش‌بینی عملکرد استفاده شده است. در این زمینه ژو و همکاران (۲۷) ساختار وابسته بین نوسانات درآمد ناشی از ریسک سیستماتیک و عملکرد را با استفاده از توابع مفصل بررسی نمودند. همچنین بوکوشوا (۵) ساختار وابستگی شاخص‌های آب و هوایی و عملکرد محصول را در قزاقستان طی سال‌های ۱۹۶۱ تا ۲۰۰۳ با استفاده از توابع مفصل سلسله مراتبی بررسی نمود. شولت و برگ (۲۵) نیز در مطالعه مشابهی به مقایسه دو روش ضریب همبستگی و توابع مفصل ارشمیدسی سلسله مراتبی پرداختند و نشان دادند که توابع مفصل ارشمیدسی سلسله مراتبی بهتر از سایر روش‌ها می‌تواند ارتباط بین عملکرد محصولات مختلف و شاخص‌های آب و هوایی را مدل‌سازی نماید. همچنین گودوین (۱۷) در مطالعه‌ای با هدف مدل‌سازی ریسک سیستماتیک در کشاورزی آمریکا بر پایه توابع مفصل، اقدام به نرخ‌گذاری حق بیمه پوشش درآمدی محصول با استفاده از توابع مفصل نمود. وی ریسک موجود در تولید دو محصول ذرت و سویا را مدل‌سازی و توابع مفصل مختلف بیضوی و ارشمیدسی را با روش «مفصلی تاکی شکل

بررسی ساختار وابستگی .....

کانونی<sup>۷</sup> (C-vine) مقایسه کرد تا ساختار وابستگی ریسک محصول را محاسبه کند. او دریافت که روش C-vine برای تبیین ساختار وابستگی بسیار برتر از روش‌های سنتی می‌باشد. با توجه به آنچه بیان شد، اگرچه ریسک عنصری گریزناپذیر است اما می‌توان آن را مدیریت کرد که یکی از سیاست‌های مهم اتخاذی در بخش کشاورزی در جهت مدیریت ریسک، بیمه محصولات کشاورزی می‌باشد. ولی سیستم‌های بیمه‌ای سنتی به علت وجود اطلاعات نامتقارن مشکلاتی در اجرا دارند (۵). با توجه به تجربیات دیگر کشورها، به نظر می‌رسد استفاده از روش‌های نوین بیمه مانند «بیمه شاخص آب و هوا» و «بیمه عملکرد منطقه» بتواند مشکل بیمه سنتی را حل نماید. در این راستا استفاده از روش جدید مفصل تاکی شکل در تبیین ساختار وابستگی بین عملکرد و شاخص‌های آب و هوایی می‌تواند نتایج قابل اعتمادتری را ارائه نماید؛ لذا هدف اصلی این پژوهش تبیین ساختار وابستگی بین متغیرهای آب و هوایی و عملکرد گندم دیم میانه با استفاده از تابع مفصل تاکی شکل است تا در نهایت عملکرد گندم دیم با توجه به شاخص‌های آب و هوایی شبیه‌سازی گردد. این روش مشکل همبستگی خطی و ثابت بین عملکرد و شاخص‌های آب و هوایی را ندارد؛ لذا نتایج آن می‌تواند مطمئن‌تر از روش‌های سنتی و مبنای عمل «صندوق بیمه کشاورزی» در طراحی بیمه شاخص آب و هوا و بیمه عملکرد منطقه باشد.

## مواد و روش‌ها

روش‌های کلاسیک مانند رگرسیون ساده و ضریب همبستگی خطی به منظور بررسی ساختار وابستگی در ادبیات موضوع مخصوصاً مطالعات داخلی به‌وفور استفاده شده است؛ اما این روش‌ها اشکالات اساسی مانند در نظر گرفتن رابطه یک طرفه یا دوجه‌دوی متغیرها و پایه‌گذاری بر توزیع نرمال را دارند (۲۵). بنابراین به‌نظر می‌رسد تعیین توزیع توأم متغیرها برای تبیین تأثیرگذاری هم‌زمان آنها بتواند به نتایج قابل اعتمادتری منجر شود. در همین راستا جستجو برای توزیع‌های چندمتغیره انعطاف‌پذیر، مدل‌های مفصل را اخیراً بسیار محبوب کرده است (۸). تابع مفصل یک ابزار آماری کارآمد است که اجازه مطالعه ساختار وابستگی را

---

7. Canonical Vine Copula (C-Vine)

می‌دهد. کلمه «مفصل» یا «کاپیولا» از واژه لاتین برای «Connect یا Join، Link» گرفته شده است (۹). توزیع‌های حاشیه‌ای می‌توانند از طریق مفصل به هم متصل شوند و به شکل توزیع توأم چندمتغیره در آیند.

توابع مفصل ساده نسبت به دیگر روش‌های اندازه‌گیری ساختار وابستگی، بهتر عمل می‌کنند، ولی در تعداد زیاد متغیر، محدودیت‌های ساختاری به رفتار وابستگی بین متغیرها اعمال می‌کنند. تعمیم این مدل‌ها (مثلاً به حالت سلسله مراتبی) بهبودهایی را ایجاد می‌کند، ولی محدودیت‌هایی مانند محدودیت‌های پارامتری را منجر می‌شود (۷). تحلیل داده‌ها با ابعاد بیشتر نیازمند مدل‌های تصادفی چند متغیره انعطاف‌پذیر است که بتواند الگوهای وابستگی ذاتی خوبی ارائه نماید. رهیافت توابع مفصل در دهه گذشته میلادی به خوبی بسط یافته است. مفصل تاکی شکل<sup>۸</sup> (واین کاپیولا) در زمره بهترین این نوع از بسط‌هاست (۱۱). مفصل‌های تاکی شکل در ابتدا توسط جو (۱۸) معرفی شدند و سپس با جزئیات بیشتر توسط بدفورد و کوک (۳ و ۴) و کوروویکا و کوک (۲۰) توسعه داده شدند.

شولزل و فریدریچز (۲۴) در تحقیقات مربوط به اقلیم، برچمن و زادو (۶) در مطالعه‌ای با هدف بررسی مدیریت ریسک با مفصل‌های واین برای تحلیل شاخص‌های بورس اروپایی، دیب من و همکاران (۱۲) در پژوهشی با هدف انتخاب و تخمین مفصل‌های رگیولار واین در امور مالی، برچمن و اسکپسمیر (۷) در تحقیقی با هدف مدل‌سازی وابستگی با C-vine و D-vine نشان دادند که این مدل‌ها نسبت به مدل‌های مفصلی ساده از انعطاف و کارایی بیشتری برای نشان دادن ساختار وابستگی برخوردار هستند.

واژه مفصل برای نخستین بار توسط اسکالار (۱۹۵۹) در قضیه‌ای که بعدها به نام وی نامیده شد، به کار رفته است. قضیه اسکالار بیان می‌کند اگر یک تابع توزیع توأم با توابع توزیع حاشیه‌ای  $F_1, \dots, F_d$  وجود داشته باشد، تابع مفصل در فضای  $[0,1]^d \rightarrow C$  وجود دارد به گونه‌ای که به ازای تمامی متغیرهای  $X_1, \dots, X_d$  در فضای  $\bar{R} = [-\infty, \infty]$  خواهیم داشت (۲۳):

بررسی ساختار وابستگی .....

$$F(X_1, \dots, X_d) = C(F_1(X_1), \dots, F_d(X_d)) \quad (1)$$

با توجه به رابطه ۱، یک مفصل، یک توزیع تجمعی چند متغیره  $d$  بعدی در  $[0, 1]^d$  با توابع توزیع حاشیه‌ای یکنواخت به صورت زیر است:

$$C(u_1, \dots, u_d) = F(F_1^{-1}(u_1), \dots, F_d^{-1}(u_d)) \quad (2)$$

تابع رابطه ۱ یک تابع توزیع  $d$  متغیره با حاشیه‌های  $F_1, \dots, F_d$  دارای تابع چگالی به شکل زیر می‌باشد:

$$f(x_1, \dots, x_d) = c(F_1(x_1), \dots, F_d(x_d)) \prod_{i=1}^d f_i(x_i) \quad (3)$$

تابع چگالی مفصل  $c$  از طریق رابطه ۲ و تابع‌های چگالی حاشیه‌ای  $f_i$  قابل استنتاج است (۱۶):

$$c(u_1, \dots, u_d) = \frac{\partial^d C(u_1, \dots, u_d)}{\partial u_1 \dots \partial u_d} = \frac{f(F_1^{-1}(u_1), \dots, F_d^{-1}(u_d))}{\prod_{i=1}^d f_i(F_i^{-1}(u_i))} \quad (4)$$

توابع مفصل پارامتریک به دو دسته «ضمنی و صریح»<sup>۹</sup> تقسیم می‌شوند. توابع مفصل ضمنی (مانند توزیع گوسی (نرمال) و توزیع  $t$  استیودنت) فقط می‌توانند وابستگی دنباله‌ای متقارن را اتخاذ کنند (۱۵). لازم به ذکر است که مفصل‌های ضمنی (بیضوی) به طور مستقیم با معکوس کردن تئوری اسکالار به دست می‌آیند (۷). در نقطه مقابل، توابع مفصل صریح (ارشمیدسی) دارای فرم بسته و مشخصی نیستند. تابع‌های مفصل «کلایتون، گامبل، فرانک و جو»<sup>۱۰</sup> از انواع توابع مفصل صریح به شمار می‌روند که به وسیله تابع مولد  $\varphi$  تولید می‌شوند. توابع مفصل ارشمیدسی با تابع مولد  $\varphi$  به صورت زیر تعریف می‌شوند:

$$C(u, v) = \varphi^{-1}(\varphi(u) + \varphi(v)) \quad (5)$$

در رابطه ۵ اگر  $C$  مطلقاً پیوسته باشد، تابع چگالی آن به صورت زیر است (۱۵):

$$c(u, v) = -\frac{\varphi''(C(u, v))\varphi'(u)\varphi'(v)}{[\varphi'(C(u, v))]^3} \quad (6)$$

9. Implicit and Explicit

10. "Clayton", "Gumble", "Frank" and "Joe"

اقتصاد کشاورزی و توسعه - سال بیست و ششم، شماره ۱۰۳

علاوه بر خانواده‌های ذکر شده، خانواده توزیع‌های ارشمیدسی با دو پارامتر مانند کلایتون-گامبل، جو-گامبل، جو-کلایتون، و جو-فرانک وجود دارند که جو (۱۹۹۷) به طور ساده آنها را BB1، BB6، BB7 و BB8 نام‌گذاری کرد. نسخه‌های چرخشی ۹۰، ۱۸۰ و ۲۷۰ درجه کلایتون، گامبل، جو و خانواده‌های BB نیز وجود دارد. در جدول ۱ خلاصه‌ای از انواع توابع مفصل‌ها ارائه شده است:

جدول ۱. خلاصه‌ای از انواع توابع مفصل

اسم	تابع مولد	شکل دو متغیره
گوسی	-	$\Phi_p(\Phi^{-1}(u), \Phi^{-1}(v))$
t	-	$t_{v,p}(t_v^{-1}(u), t_v^{-1}(v))$
کلایتون	$(t^{-\theta} - 1) / \theta$	$(\max\{u^{-\theta} + v^{-\theta} - 1; 0\})^{-1/\theta}$
گامبل	$(-\log(t))^\theta$	$\exp(-((-\ln u)^\theta + (-\ln v)^\theta)^{1/\theta})$
فرانک	$-\log\left(\frac{\exp(-\theta t) - 1}{\exp(-\theta) - 1}\right)$	$-\frac{1}{\theta} \ln\left\{1 + \frac{(e^{-\theta u} - 1)(e^{-\theta v} - 1)}{e^{-\theta} - 1}\right\}$
جو	$-\log(1 - (1-t)^\theta)$	$1 - ((1-u)^\theta + (1-v)^\theta) - (1-u)^\theta (1-v)^\theta)^{1/\theta}$
کلایتون-گامبل (BB1)	$(t^{-\theta} - 1)^\delta$	$(1 + [\bar{u}^\delta + \bar{v}^\delta]^{1/\delta})^{-1/\theta}$
جو-گامبل (BB6)	$(-\log[1 - (1-t)^\theta])^\delta$	$1 - (1 - 10^{-[(-\log[1 - (1-u)^\theta])^\delta + (-\log[1 - (1-v)^\theta])^\delta]^{1/\delta}}})^{1/\theta}$
جو-کلایتون (BB7)	$(1 - (1-t)^\theta)^{-\delta} - 1$	$1 - (1 - [(1 - \bar{u}^\theta)^{-\delta} + (1 - \bar{v}^\theta)^{-\delta} - 1]^{-1/\delta})^{1/\theta}$
جو-فرانک (BB8)	$-\log\left(\frac{1 - (1 - \delta t)^\theta}{1 - (1 - \delta)^\theta}\right)$	$1 - [1 - 10^{-\frac{[-\log(\frac{1 - (1 - \delta u)^\theta}{1 - (1 - \delta)^\theta}) - \log(\frac{1 - (1 - \delta v)^\theta}{1 - (1 - \delta)^\theta})]}]{1/\theta}$
		$\delta$



بررسی ساختار وابستگی .....

ادامه جدول ۱.

وابستگی دنباله‌ای (بالا، پایین)	کندال $\tau$	پارامتر	اسم
0	$\frac{2}{\pi} \arcsin(\rho)$	$\rho \in (-1,1)$	گوسی
$2t_{v+1}(-\sqrt{v+1}\sqrt{\frac{1-\rho}{1+\rho}})$	$\frac{2}{\pi} \arcsin(\rho)$	$\rho \in (-1,1)$ $, v > 2$	t
$(2^{-1/\theta}, 0)$	$\frac{\theta}{\theta+2}$	$\theta > 0$	کلاتون
$(0, 2-2^{1/\theta})$	$1-\frac{1}{\theta}$	$\theta \geq 1$	گامبل
(0,0)	$1-\frac{4}{\theta} + 4 \frac{\int_0^{\theta} \frac{c/\theta}{\exp(x)-1} dx}{\theta}$	$\theta \in \mathbb{R} \setminus \{0\}$	فرانک
$(0, 2-2^{1/\theta})$	$1-\frac{4}{\theta^2} \int_0^1 t \log(t)(1-t)^{2(1-\theta)/\theta} dt$	$\theta \geq 1$	جو
$(2^{\frac{1}{\delta\theta}}, 2-2^{\frac{1}{\delta}})$	$1-\frac{2}{\delta(\theta+2)}$	$\theta > 0,$ $\delta \geq 1$	کلاتون-گامبل (BB1)
$(0, 2-2^{\frac{1}{\delta\theta}})$	$1+\frac{4}{\delta\theta} \int_0^1 -\log[1-(1-t)^\theta] \times (1-t)(1-(1-t)^\theta) dt$	$\theta \geq 1,$ $\delta \geq 1$	جو-گامبل (BB6)
$(2^{\frac{1}{\delta}}, 2-2^{\frac{1}{\theta}})$	$1+\frac{4}{\delta\theta} \int_0^1 (-[1-(1-t)^\theta]^{\delta+1} \times \frac{(1-(1-t)^\theta)^{-\delta}-1}{(1-t)^{\theta-1}}) dt$	$\theta \geq 1,$ $\delta > 0$	جو-کلاتون (BB7)
(0,0)	$1+\frac{4}{\delta\theta} \int_0^1 (-\log[\frac{(1-\delta t)^\theta-1}{(1-\delta)^\theta-1}] \times (1-t\delta)(1-(1-t\delta)^\theta)) dt$	$\theta \geq 1,$ $\delta \in (0,1]$	جو-فرانک (BB8)

منبع: برچمن و اسکیمیر (۷) و فیشر (۱۵)

اگر  $\delta = 1$  باشد ضریب وابستگی دنباله‌ای بالا برای BB8 به صورت  $2-2^{1/\theta}$  خواهد بود. در رابطه‌های بالا t تابع

توزیع حاشیه‌ای است که در حالت دو متغیره توزیع‌های u و v جایگزین آن می‌شوند.

**توابع مفصل تاکی شکل (یا واین کاپیولا<sup>۱۱</sup>):** این نوع مفصل‌ها یک مدل گرافیکی انعطاف‌پذیر برای توصیف ساختار مفصل‌های چندمتغیره با استفاده از مفصل‌های دو متغیره که «ساختار جفت مفصل»<sup>۱۲</sup> (یا PCC) نامیده می‌شوند را فراهم می‌آورند. همچنین ساختار جفت مفصل یک چگالی احتمال چندمتغیره را به مفصل‌های دو متغیره تجزیه می‌کند. مفصل‌های تاکی شکل به این ترتیب مزیت‌های مدل‌سازی‌های مفصلی چند متغیره را ترکیب می‌کنند و با انعطاف‌پذیر کردن مفصل‌های دو متغیره، مدل‌سازی وابستگی را کارتر می‌نمایند. مفصل‌های تاکی شکل اجازه می‌دهند رفتار ساختارهای مختلف جفت متغیره‌ها به طور مناسب مدل‌سازی شود، مخصوصاً این اجازه را می‌دهند که به تقارن و عدم تقارن، قدرت وابستگی و وابستگی دنباله‌ای در مدل‌سازی توجه شود. این انعطاف‌پذیری به انتخاب یک مدل طراحی خوب برای تحقق بخشیدن به پتانسیل مفصل‌های تاکی شکل در مدل‌های وابستگی نیاز دارد (۱۱). مطابق مطالعه آس و همکاران (۱) تابع چگالی چندمتغیره توأم رابطه ۳ برای مجموعه‌ای از  $d$  متغیر تصادفی به صورت زیر می‌تواند بازنویسی شود:

$$f(x_1, x_2, \dots, x_d) = f_d(x_d) \cdot f(x_{d-1} | x_d) \cdot f(x_{d-2} | x_{d-1}, x_d) \dots \cdot f(x_1 | x_2, \dots, x_d) \quad (7)$$

اگر یک حالت سه متغیره  $F \sim (X_1, X_2, X_3)^T$  با توابع توزیع حاشیه‌ای  $F_1$

و  $F_2$  و  $F_3$  و چگالی‌های  $f$  در نظر گرفته شود، با حالت عطفی<sup>۱۳</sup> می‌توان رابطه زیر را نوشت:

$$f(x_1, x_2, x_3) = f_1(x_1) \cdot f(x_2 | x_1) \cdot f(x_3 | x_1, x_2) \quad (8)$$

به وسیله قضیه اسکالر خواهیم داشت:

$$f(x_2 | x_1) = \frac{f(x_1, x_2)}{f_1(x_1)} = \frac{c_{12}(F_1(x_1), F_2(x_2)) \cdot f_1(x_1) \cdot f_2(x_2)}{f_1(x_1)} = c_{12}(F_1(x_1), F_2(x_2)) \cdot f_2(x_2) \quad (9)$$

11. Vine Copula

12. Pair-Copula Construction (PCC)

13. Recursive

و

$$f(x_3 | x_1, x_2) = \frac{f(x_2, x_3 | x_1)}{f(x_2 | x_1)} = \frac{c_{23|1}(F(x_2 | x_1), F(x_3 | x_1)) \cdot f(x_2 | x_1) \cdot f(x_3 | x_1)}{f(x_2 | x_1)} \quad (10)$$

$$= c_{23|1}(F(x_2 | x_1), F(x_3 | x_1)) \cdot f(x_3 | x_1) = c_{23|1}(F(x_2 | x_1), F(x_3 | x_1)) c_{13}(F_1(x_1), F_3(x_3)) \cdot f_3(x_3) \quad (3)$$

روابط ۸ تا ۱۰ را می‌توان به صورت زیر نوشت:

$$f(x_1, x_2, x_3) = \underbrace{f_1(x_1)}_{\text{Marginal}} \cdot \underbrace{f_2(x_2)}_{\text{Marginal}} \cdot \underbrace{f_3(x_3)}_{\text{Marginal}} \times \underbrace{c_{12}(F_1(x_1), F_2(x_2))}_{\text{Unconditional pairs}} \cdot \underbrace{c_{13}(F_1(x_1), F_3(x_3))}_{\text{Unconditional pairs}} \quad (11)$$

$$\times \underbrace{c_{23|1}(F(x_2 | x_1), F(x_3 | x_1))}_{\text{Conditional pairs}}$$

بنابراین چگالی توأم سه بعدی رابطه ۱۱ را می‌توان به صورت توابع چگالی حاشیه‌ای، مفصل‌های دو متغیره  $C_{12}$ ،  $C_{13}$  و  $C_{23|1}$  با چگالی‌های  $c_{12}$ ،  $c_{13}$  و  $c_{23|1}$  - که «جفت مفصل»<sup>۱۴</sup> نامیده می‌شوند- نشان داد. از آنجا که تجزیه در رابطه ۸ یکتا نیست و تعداد زیادی از PCC های تکراری وجود دارد، برای طبقه‌بندی آنها بدفورد و کوک (۳ و ۴) مدل‌های گرافیکی که «تاکی شکل منظم»<sup>۱۵</sup> (R-vine) نامیده می‌شوند را معرفی کردند که جزئیات آن توسط کوروویکا و کوک (۲۰) و کوروویکا و جو (۲۱) مورد بحث قرار گرفته است (۷). به عبارت دیگر یک مفصل تاکی شکل منظم با  $n$  مؤلفه، یک مجموعه متداخل از  $n-1$  درخت است که یال‌های درخت  $z$  گره‌های درخت  $z+1$  می‌باشند. شرط مجاورت تأکید می‌کند که دو گره در صورتی می‌توانند به وسیله یک یال در درخت  $z+1$  به یکدیگر متصل شوند که این گره‌ها در درخت  $z$  در یک گره مشترک، سهمیم باشند. توجه شود که مجموعه گره‌ها در درخت اول شامل همه  $1, \dots, n$  عضو است، در حالی که مجموعه یال‌ها یک مجموعه از  $n-1$  جفت از این

14. Pair Copula

15. Regular Vine Copula

اعضا می‌باشد. در درخت دوم مجموعه گره‌ها شامل مجموعه جفت‌های اعضا و مجموعه یال‌ها دوباره جفت‌هایی از جفت‌های اعضای قبلی را می‌سازند (۱۱).

خانواده‌های مختلفی از مفصل‌های تاکی شکل منظم وجود دارد که سه مورد از مهم‌ترین آنها شامل R-vine، C-vine و مفصل D-vine است که دو نوع C-vine (مفصل تاکی شکل کانونی) و D-vine (مفصل تاکی شکل قابل رسم<sup>۱۶</sup>) به‌طور گسترده‌ای در ادبیات استفاده شده‌اند (۱۱) در ادامه به تشریح مفصل ترکیبی C-vine پرداخته می‌شود.

یک C-vine یک مفصل تاکی شکل منظم است که ساختار درختی ستاره‌ای مانند دارد. در ساختار C-vine عموماً یک گره ریشه‌ای در هر درخت انتخاب می‌شود و وابستگی‌های جفتی با توجه به این گره و شرط در نظر گرفتن همه گره‌های ریشه‌ای قبلی، مدل‌سازی می‌شوند که در شکل ۱ نمونه‌ای از آن نمایش داده شده است. براساس این مطالب تجزیه یک تابع چگالی چندمتغیره با گره‌های ریشه‌ای  $1, \dots, d$  به صورت زیر است که یک تابع چگالی C-vine نامیده می‌شود:

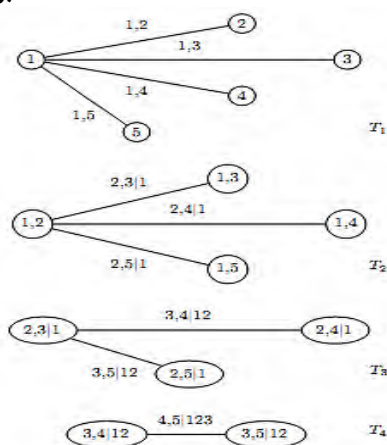
$$f(x) = \prod_{k=1}^d f_k(x_k) \prod_{i=1}^{d-1} \prod_{j=1}^{d-i-1} c_{i,i+j|l:(i-1)}(F(x_i | x_1, \dots, x_{i-1})), \quad (12)$$

در اینجا  $f_k$ ،  $k=1, \dots, d$ ، نشان‌دهنده چگالی‌های حاشیه‌ای و  $c_{i,i+j|l:(i-1)}$  چگالی مفصل‌های دو متغیره با پارامترهای  $\theta_{i,i+j|l:(i-1)}$  (منظور از  $i_k: i_m$ ،  $i_k, \dots, i_m$  است) است. برای

حالت  $d$  بعدی، جفت‌ها در یک C-vine برای هر درخت به صورت زیر تعیین می‌شوند:

$$\left\{ \begin{array}{ll} (1,2), (1,3), (1,4), \dots, (1, d), & \text{Tree1} \\ (2,3 | 1), (2,4 | 1), \dots, (2, d | 1), & \text{Tree2} \\ (3,4 | 1,2), (3,5 | 1,2), \dots, (3, d | 1,2), & \text{Tree3} \\ \vdots & \vdots \\ (d-1, d | 1, \dots, d-2). & \text{Tree } d-1 \end{array} \right. \quad (13)$$

بررسی ساختار وابستگی .....



شکل ۱. مثالی از حالت ۵ بعدی یک C-vine (منبع: برچمن و اسکپسیر (۷))

به منظور ایجاد درختان بعدی باید ورودی برای آنها ایجاد نمود بنابراین می توان از قضیه اسکالر استفاده کرد و یک توزیع توأم یکتای  $F$  با  $n$  متغیر با چگالی رابطه ۱۴ را در نظر گرفت:

$$f_{1...n}(x) = \prod_{k=1}^n f_k(x_k) \prod_{i=1}^{n-1} \prod_{e \in E_i} c_{C_{e,a}, C_{e,b} | D_e} (F_{C_{e,a} | D_e}(x_{C_{e,a}} | x_{D_e}), F_{C_{e,b} | D_e}(x_{C_{e,b}} | x_{D_e})) \quad (14)$$

که  $x = (x_1, \dots, x_n)$  و  $e = \{a, b\}$  و  $x_{D_e} = \{x_i | i \in D_e\}$  است و  $f_k$  نشان دهنده چگالی  $F_k$  برای  $i=1, \dots, n$  است.  $D_e$  مجموعه مشروط کننده و  $C_{e,a}$  و  $C_{e,b}$  مجموعه های مشروط شونده هستند. به منظور تعیین توزیع های شرطی  $F_{C_{e,a} | D_e}(x_{C_{e,a}} | x_{D_e})$  و  $F_{C_{e,b} | D_e}(x_{C_{e,b}} | x_{D_e})$  فرض می شود که  $e = \{a, b\} \in E_i$  با  $a = \{a_1, a_2\}, b = \{b_1, b_2\}$  یالی است که  $C_{e,a}$  را به

$C_{e,b}$  با متغیرهای مفروض  $D_e$  متصل می کند. جو (۱۹۹۶) نشان داد که:

$$F_{C_{e,a} | D_e}(x_{C_{e,a}} | x_{D_e}) = \frac{\partial C_{C_a | D_a}(F_{C_{a,a_1} | D_a}(x_{C_{a,a_1}} | x_{D_a}), F_{C_{a,a_2} | D_a}(x_{C_{a,a_2}} | x_{D_a}))}{\partial F_{C_{a,a_2} | D_a}(x_{C_{a,a_2}} | x_{D_a})} = h(F_{C_{a,a_1} | D_a}(x_{C_{a,a_1}} | x_{D_a}), F_{C_{a,a_2} | D_a}(x_{C_{a,a_2}} | x_{D_a})) \quad (15)$$

به صورت عطفی و درخت به درخت  $F_{C_{a,a_2} | D_a}(x_{C_{a,a_2}} | x_{D_a})$  و  $F_{C_{a,a_1} | D_a}(x_{C_{a,a_1}} | x_{D_a})$  به دست آمده اند. به طور مشابه  $F_{C_{e,b} | D_e}(x_{C_{e,b}} | x_{D_e})$  به دست می آید.  $F_{C_{e,a} | D_e}(x_{C_{e,a}} | x_{D_e})$  و  $F_{C_{e,b} | D_e}(x_{C_{e,b}} | x_{D_e})$  متغیرهای تغییر شکل یافته هستند که از آنها به منظور ایجاد داده در

درختان بعدی و شبیه‌سازی استفاده می‌شود.  $(v, B, \theta)$  نماینده یک مفصل تاکی شکل منظم است.  $v$  ساختار یک مفصل تاکی شکل منظم با  $n$  بعد،  $B = \{B_e \mid i=1, \dots, n-1; e \in E_i\}$  یک مجموعه از مفصل‌های  $B_e$  دو متغیره که جفت مفصل نامیده می‌شود و  $\theta$  پارامترهای جفت مفصل‌ها هستند. این مشخصات بیان می‌کند که برای کار با یک مفصل تاکی شکل منظم، باید سه انتخاب صورت گیرد: انتخاب اول مربوط به ساختار درختی، انتخاب دوم مربوط به انتخاب جفت مفصل‌ها و انتخاب سوم نحوه تخمین پارامترهای مربوط به این جفت مفصل‌ها هستند (۱۲).

برای یک ساختار درختی مفصل تاکی شکل منظم مشخص  $(v)$  و خانواده‌های جفت مفصلی  $B = B(v)$  (که جفت مفصل تابعی از ساختار درختی است) چالش اساسی، تخمین پارامترهای  $\theta = \theta(B(v))$  مفصل تاکی شکل منظم است. به طور کلاسیکی، پارامترهای یک مدل آماری اغلب با استفاده از تکنیک حداکثر درست‌نمایی تخمین زده می‌شود. زادو و همکاران (۱۱) تابع درست‌نمایی یک مفصل تاکی شکل منظم را به صورت زیر تعریف کردند:

$$L(v, B, \theta) = \prod_{i;d}^N f_{i;d}(x_k \mid v, B, \theta) \quad (16)$$

برای یک C-vine با پارامترهای  $\theta_{Dv}$ ، لگاریتم درست‌نمایی به صورت زیر است:

$$L_{cv}(\theta_{cv}) = \sum \sum_{d-1} \sum_{d-i} \log [c_{i,i+jll:(i-1)}(F_{ill:(i-1)}, F_{i+jll:(i-1)} \mid \theta_{i,i+jll:(i-1)})] \quad (17)$$

برای مفصل‌ها، مخصوصاً اگر  $n > 2$  باشد، تعداد پارامترهای برآوردی ممکن است بیش از حد زیاد باشد و تخمین حداکثر درست‌نمایی توأم (JML) پارامترهای مفصل تاکی شکل منظم می‌تواند چالش زیادی داشته باشد. از این رو، روش SM (یا توالی) <sup>۱۷</sup> پیشنهاد می‌شود. این روش با درخت اول شروع می‌کند و پارامترهای جفت مفصل‌های درخت اول را با استفاده از حداکثر درست‌نمایی تخمین می‌زند. سپس در درخت دوم باید پارامترهای مفصل‌های شرطی را تخمین زد. برای این منظور از مشاهدات ساختگی رابطه ۱۵ و پارامترهای تخمین زده

بررسی ساختار وابستگی .....

شده جفت مفصل‌های درخت اول استفاده می‌شود. به طور مشابه، این کار برای دیگر مفصل‌ها انجام می‌پذیرد.

با فرض اینکه یک ساختار مناسب و این  $v$  انتخاب شده است، پرسشی که مطرح می‌شود این است که چگونه جفت مفصل مناسب  $C_{i(e),j(e)|D_e}$  انتخاب شود. انتخاب خانواده مفصل می‌تواند به صورت اختیاری باشد، یعنی از طبقه‌های پرکاربرد ارضمیدسی و ضمنی باشد (۱۱) یا می‌توان از ابزارهای گرافیکی و ابزارهای تحلیلی استفاده کرد که ابزارهای تحلیلی مانند آماره‌های AIC و BIC از اعتبار بیشتری برخوردار می‌باشد.

خانواده‌های جفت مفصلی  $B = B(v)$  و پارامترهایشان  $\theta = \theta(B(v))$  هر دو به ساختار درختی و این  $v$  وابسته هستند. برای یک حالت  $d$  بعدی، تعداد مفصل تاکی شکل منظم می‌توان تعریف کرد. این بدان معنی است که تعداد مدل‌های مختلف بسیار زیاد است. حتی برای زیرطبقه‌های C-vine و D-vine‌ها هنوز  $d!/2$  نوع مختلف از این دو مفصل وجود دارد. مفصل‌های C-vine و D-vine به خصوص C-vine‌ها مناسب‌ترین انواع هستند. برای انتخاب ساختار درختی دو روش کلی توالی و انتخاب درخت با توالی بیزین وجود دارد. روش توالی که در ادبیات کاربرد بیشتری نسبت به روش توالی بیزین دارد یک رویه توالی درخت به درخت را با توجه به شرط مجاورت در هر مرحله در پیش می‌گیرند (۱۱). در رویه توالی درخت به درخت معمولاً براساس حداکثر مجموع قدر مطلق تاو کندال تصمیم‌گیری می‌شود. با توجه به آنچه گفته شد، به منظور شبیه‌سازی عملکرد گندم دیم با توجه به متغیرهای آب و هوایی، مراحل زیر دنبال خواهد شد:

۱. ایجاد داده‌های مفصل: مجموعه داده‌هایی که با آنها کار می‌شود شامل عملکرد گندم دیم برای رقم سرداری در شهرستان میانه و متغیرهای آب و هوایی شامل «دمای هوای میانگین»، «بارش تجمعی» و «رطوبت نسبی میانگین» برای دوره رشد (اول فروردین تا آخر تیر) و «سرعت باد حداکثر» برای دوره برداشت (اول تیر تا آخر مرداد) باید توزیع‌های حاشیه‌ای یکنواخت در  $[0, 1]$  داشته باشند که داده‌های مفصل نامیده می‌شوند. برای این منظور می‌توان از

اقتصاد کشاورزی و توسعه - سال بیست و ششم، شماره ۱۰۳

تابع توزیع تجمعی تجربی<sup>۱۸</sup> (ECDF) استفاده کرد و برای ساده‌سازی محاسبات به متغیرهای استاندارد شده فوق یا توزیع‌های حاشیه‌ای یکنواخت با تابع توزیع تجمعی تجربی به صورت زیر یک شماره نسبت داد:

$$Y_{ECDF}=1, T_{ECDF}=2, CRI_{ECDF}=3, RH_{ECDF}=4, U_{ECDF}=5$$

که در آن توزیع‌های حاشیه‌ای یکنواخت  $Y_{ECDF}$  عملکرد گندم دیم،  $T_{ECDF}$  دمای هوای میانگین،  $CRI_{ECDF}$  بارش تجمعی،  $RH_{ECDF}$  رطوبت نسبی و  $U_{ECDF}$  سرعت باد حداکثر می‌باشند.

۲. تخمین توزیع توأم متغیرهای آب و هوایی و عملکرد گندم دیم با استفاده از

مفصل C-vine.

۳. شبیه‌سازی با توجه به توزیع توأم: شبیه‌سازی از مفصل‌های مفصل تاکی شکل منظم توسط بدفورد و کوک (۴) مورد بحث قرار گرفته است. الگوریتم کلی برای نمونه برداری  $n$  متغیر وابسته یکنواخت  $[0, 1]$  برای C-vine به صورت زیر است:

اول، نمونه برداری از  $u'_1, \dots, u'_n$  متغیر یکنواخت مستقل در  $[0, 1]$  صورت می‌گیرد و سپس در روابط زیر قرار داده می‌شود:

$$\begin{cases} x_1 = u'_1 \\ x_2 = F^{-1}(u'_2 | x_1) \\ x_3 = F^{-1}(u'_3 | x_1, x_2) \\ \vdots \\ x_n = F^{-1}(u'_n | x_1, \dots, x_{n-1}) \end{cases} \quad (18)$$

به منظور تعیین  $F(x_i | x_1, \dots, x_{i-1})$  برای هر  $z$  از تابع  $h$  به صورت عطفی برای هر دو ساختار و این استفاده می‌شود. به هر حال، انتخاب متغیر  $v_j$  برای یک C-vine به صورت زیر می‌باشد (۱):

$$F(x_j | x_1, \dots, x_{j-1}) = \frac{\partial C_{j,j-1|1, \dots, j-2} \{F(x_j | x_1, \dots, x_{j-2}), F(x_{j-1} | x_1, \dots, x_{j-2})\}}{\partial F(x_{j-1} | x_1, \dots, x_{j-2})} \quad (19)$$



بررسی ساختار وابستگی .....

۴. تعیین توزیع حاشیه‌ای متغیرها: وقتی متغیرهای شبیه‌سازی شده به صورت داده‌های مفصل در فاصله بسته  $[0, 1]$  به دست آمدند، به منظور برگرداندن آنها به فرم واقعی می‌توان از «وارون تابع توزیع تجمعی» هر متغیر استفاده کرد. لذا لازم است ابتدا توزیع حاشیه‌ای متغیرها را مشخص کرد. برای این منظور از آزمون‌هایی مانند «کولموگروف-اسمیرنف»<sup>۱۹</sup>، «اندرسون-دارلینگ»<sup>۲۰</sup> و «خی-دو»<sup>۲۱</sup> استفاده می‌شود. متغیرهای عملکرد و آب وهوایی استفاده شده در این تحقیق در طول سال‌های زراعی ۹۲-۱۳۶۶ به ترتیب از جهاد کشاورزی و سازمان هواشناسی استان آذربایجان شرقی گردآوری گردیدند.

## نتایج و بحث

پنج متغیر نام‌گذاری شده مجموعه گره‌های  $\{N_i = \{1, 2, 3, 4, 5\}\}$  درخت اول را تشکیل می‌دهند. در یک ساختار درختی C-vine باید گره مرکزی (ریشه‌ای) در هر درخت مشخص شود. برای انتخاب گره مرکزی ضریب همبستگی تاو کندال تجربی همه جفت متغیرهای  $\{j, k\}$  محاسبه می‌گردد و در نهایت گره‌ی به عنوان گره مرکزی انتخاب خواهد شد که بیشترین مجموع قدر مطلق ضریب همبستگی تاو کندال را داشته باشد. نتایج مربوط به محاسبه ضریب همبستگی تاو کندال برای تعیین گره مرکزی در درخت اول در جدول ۲ گزارش شده است. در ردیف آخر این جدول مجموع قدر مطلق ضریب همبستگی تاو کندال یا همان درخت پوشا گزارش شده است. با توجه به اینکه بیشترین عدد حاصل مربوط به متغیر ۴ یا رطوبت نسبی میانگین می‌باشد، بنابراین در درخت اول این متغیر به عنوان گره ریشه‌ای انتخاب می‌شود. به این ترتیب مجموعه یال‌ها در درخت اول به صورت  $\{E_1 = \{(4, 1), (4, 2), (4, 3), (4, 5)\}$  خواهد بود. در ادامه لازم است برای هر یال یک جفت مفصل انتخاب شود سپس با استفاده از روش حداکثر درست‌نمایی پارامترهای آن تخمین زده شوند.

---

19. Kolmogorov-Smirnov test

20. Anderson-Darling test

21. Chi-Squared test

اقتصاد کشاورزی و توسعه - سال بیست و ششم، شماره ۱۰۳

یال‌های انتخاب شده در درخت اول، گره‌های درخت دوم خواهند بود. برای انتخاب گره مرکزی دوم لازم است ابتدا ورودی‌های درخت دوم محاسبه شوند. برای این منظور، از مشاهدات ساختگی  $F_{C_{e,a}|D_e}(x_{C_{e,a}} | x_{D_e})$  یا توزیع‌های شرطی رابطه ۱۵ استفاده می‌شود. مجموعه گره‌های درخت دوم به صورت  $N_2 = \{(4,1), (4,2), (4,3), (4,5)\}$  خواهد بود. در این حالت، چهار نماینده برای گره مرکزی در درخت دوم وجود دارد. بنابراین در گام دوم با ورودی‌های جدید ایجاد شده، ضریب همبستگی تاو کندال محاسبه می‌شود که نتایج مربوط به این ضرایب همبستگی در جدول ۳ ارائه شده است. همان گونه که نتایج جدول ۳ نشان می‌دهد، براساس حداکثر مجموع قدر مطلق ضریب همبستگی تاو کندال گره مرکزی در درخت دوم، گره (۴،۲) است. مجموعه یال‌ها در درخت دوم به صورت  $E_2 = \{(2,14), (2,314), (2,514)\}$  می‌باشد. مطابق حالت قبل برای درخت دوم نیز مفصل انتخاب کرده و پارامترهای آن تخمین زده می‌شود و با استفاده از رابطه ۱۵ دوباره لازم است ورودی برای درخت سوم تعیین گردد. به همین ترتیب درخت‌های سوم و چهارم نیز ایجاد می‌شوند. نتایج مربوط به قدر مطلق ضریب همبستگی تاو کندال برای درخت سوم و چهارم در جدول ۴ و ۵ آورده شده است. نتایج جدول ۴ نشان داد که گره ریشه‌ای سوم نیز (۴،۲،۵) می‌باشد. بنابراین رتبه متغیرها در یک الگوی C-vine به ترتیب متغیر ۴ (رطوبت نسبی)، ۲ (دما)، ۵ (سرعت باد)، ۱ (عملکرد) و ۳ (بارش تجمعی) می‌باشد. مجموعه درخت‌های C-vine به صورت شکل ۲ گزارش شده است.

جدول ۲. نتایج مربوط به ضریب همبستگی تاو کندال در درخت اول

متغیر	۱	۲	۳	۴	۵
۱	-	-۰/۲۵۳	۰/۱۴۸	۰/۲۷۷	-۰/۱۳۱
۲	-۰/۲۵۳	-	-۰/۲۱۸	-۰/۴۸۳	-۰/۱۶۴
۳	۰/۱۴۸	-۰/۲۱۸	-	۰/۴۶۵	-۰/۰۷۹
۴	۰/۲۷۷	-۰/۴۸۳	۰/۴۶۵	-	۰/۰۰۷
۵	-۰/۱۳۱	-۰/۱۶۴	-۰/۰۷۹	۰/۰۰۷	-
	۰/۸۰۹	۱/۱۱۸	۰/۹۱۰	*۱/۲۳۲	۰/۳۸۱

مأخذ: یافته‌های تحقیق

بررسی ساختار وابستگی .....

جدول ۳. نتایج مربوط به ضریب همبستگی تاو کندال در درخت دوم

۴/۵	۴/۳	۴/۲	۴/۱	
-۰/۰۵۲	-۰/۰۳۴	-۰/۱۲۰	-	۴/۱
-۰/۱۷۵	-۰/۰۳۴	-	-۰/۱۲۰	۴/۲
-۰/۰۵۲	-	-۰/۰۳۴	-۰/۰۳۴	۴/۳
-	-۰/۰۵۲	-۰/۱۷۵	-۰/۰۵۲	۴/۵
۰/۲۸۰	۰/۱۲۰	*۰/۳۲۹	۰/۲۰۶	جمع

مأخذ: یافته‌های تحقیق

جدول ۴. نتایج مربوط به ضریب همبستگی تاو کندال در درخت سوم

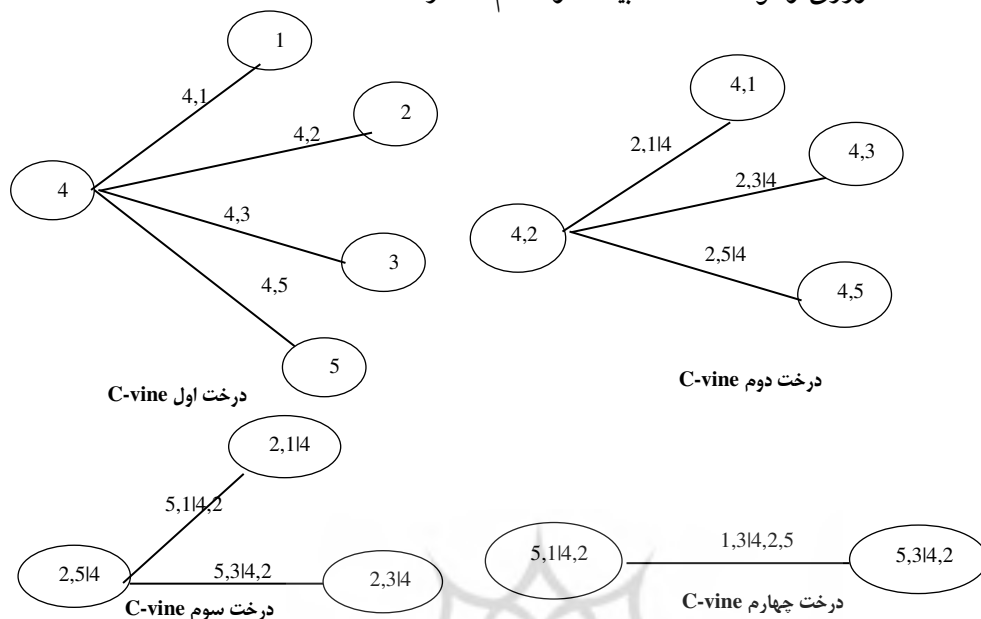
2,514	2,314	2,114	
-۰/۱۲۶	-۰/۰۸۹	-	2,114
-۰/۱۰۸	-	-۰/۰۸۹	2,314
-	-۰/۱۰۸	-۰/۱۲۶	2,514
*۰/۲۳۴	۰/۱۹۷	۰/۲۱۵	جمع

مأخذ: یافته‌های تحقیق

جدول ۵. نتایج مربوط به ضریب همبستگی تاو کندال در درخت چهارم

5,314,2	5,114,2	
-۰/۱۲۶	-	5,114,2
-	-۰/۱۲۶	5,314,2
۰/۱۲۶	۰/۱۲۶	جمع

مأخذ: یافته‌های تحقیق



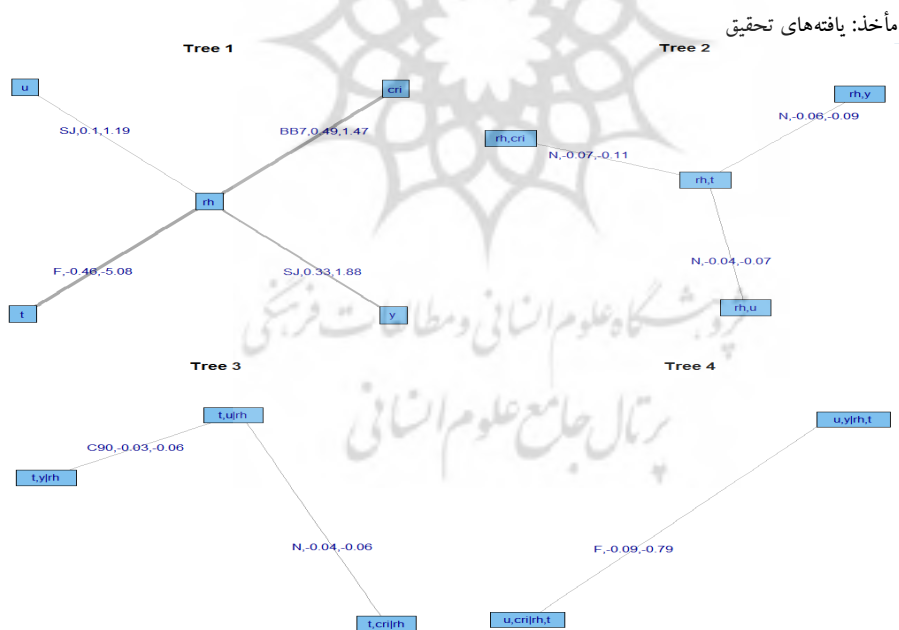
شکل ۲. مجموعه درختان C-vine

به منظور انتخاب خانواده مناسب جفت مفصلی برای یک ساختار درختی C-vine از آماره‌های AIC و BIC استفاده شد و بعد از تعیین جفت مفصل‌ها تخمین پارامترهای آنها با استفاده از روش SM و JML صورت گرفت که نتایج به دست آمده از دو روش یکسان می‌باشد. نتایج حاصل از تخمین پارامترها در جدول ۶ گزارش شده است. با توجه به نتایج جدول ۶ متغیرهای 4,1، 4,5 و 5,1,14,2 وابستگی متقارن نداشته و وابستگی قوی در دنباله پایین دارند. به عبارتی به مقادیر کوچک هم واکنش بیشتری نشان می‌دهند و تأثیرپذیری آنها در مقادیر کوچک بیشتر از مقادیر بزرگ است. مفاصل فرانک و گوسی نمی‌توانند وابستگی دنباله‌ای را اندازه‌گیری کنند و مفصل جو-کلایتون که برای جفت متغیر 4,3 انتخاب شده است، وابستگی دنباله‌ای متقارن را برای این متغیرها نشان می‌دهد.

بررسی ساختار وابستگی .....

جدول ۶. نتایج تخمین پارامترهای جفت مفصل‌ها

درخت	پارامترها	خانواده انتخاب شده	ضرایب برآوردی	انحراف معیار
اول	$P_{4,2}$	فرانک	-۵/۰۸۳	۱/۵۲
	$P_{4,5}$	جویقا	۱/۱۹۴	۰/۳۱
	$P_{4,1}$	جویقا	۱/۸۷۷	۰/۳۹
	$P_{1,4,3}$	جو-کلایتون	۱/۴۶۷	۰/۰۱
	$P_{2,4,3}$		۱/۴۹۸	۰/۵۶
دوم	$P_{2,5,4}$	گاوسی	-۰/۰۶۵۲	۰/۰۸
	$P_{2,1,4}$	گاوسی	-۰/۰۸۸	۰/۱
	$P_{2,3,4}$	گاوسی	-۰/۱۰۶	۰/۱۲
سوم	$P_{5,1,4,2}$	کلایتون ۹۰ درجه گاوسی	-۰/۰۵۶	۰/۰۳
	$P_{5,3,4,2}$		-۰/۰۵۶	۰/۱
چهارم	$P_{1,3,4,5,2}$	فرانک	-۰/۷۹۰	۱/۲۲
			۴۰/۹۱۷	
Log-likelihood				



شکل ۳. ساختار درختی C-vine برآورد شده بین عملکرد گندم دیم و شاخص‌های آب و هوایی (خروجی نرم افزار R)

اقتصاد کشاورزی و توسعه - سال بیست و ششم، شماره ۱۰۳

ساختار درختی مدل C-vine یا تابع توزیع چگالی توأم متغیر عملکرد و شاخص‌های آب و هوایی بعد از تخمین پارامترهای جفت مفصلی در شکل ۳ که خروجی نرم‌افزار R3.1.1 می‌باشد، نشان داده شده است. در ساختار درختی شکل ۳ عبارت اول خانواده مفصل، عبارت دوم ضریب همبستگی تاو کندال و عبارت آخر ضریب پارامتر برآورد شده می‌باشد.

قبل از شبیه‌سازی داده‌های عملکرد با تابع توزیع توأم مدل C-vine لازم است توزیع حاشیه‌ای متغیر عملکرد گندم تعیین گردد. به منظور بررسی انطباق توزیع تجربی یک متغیر بر توزیع‌های نظری از آزمون‌های «کولموگروف-اسمیرنف»، «اندرسون-دارلینگ» و «خی-دو» استفاده می‌شود. جهت انجام دقیق این بررسی از میان توزیع‌های نظری مشهور، ۶۵ توزیع نظری در نظر گرفته شد. پس از محاسبه آماره‌های آزمون‌های سه گانه فوق برای این ۶۵ توزیع با استفاده از نرم‌افزار EasyFit 5.5، بهترین گزینه‌ها از نظر انطباق توزیع نظری بر توزیع تجربی متغیر عملکرد گندم دیم شهرستان میانه به ترتیب به قرار جدول ۷ به دست آمد.

جدول ۷. توابع توزیع نظری که بیشترین انطباق را با توزیع تجربی متغیر عملکرد گندم دیم دارند

خی-دو	اندرسون- دارلینگ	کولموگروف- اسمیرنف	توزیع			
			آماره	P-value	توزیع	
۰/۰۷۵	۰/۱۳۱	۰/۰۵۸	آماره	Wakeby (ویکبای)	اولین گزینه	
۰/۹۹۴	---	۰/۹۹۹	P-value		مقدار بحرانی ( $\alpha = 5\%$ )	
۷/۸۱۴	۲/۵۰۱	۰/۲۵۹				
۰/۳۲۱	۰/۳۰۸	۰/۱۱۲	آماره	Weibull (وایبل)	دومین گزینه	
۰/۹۵۶	---	۰/۸۵۸	P-value		مقدار بحرانی ( $\alpha = 5\%$ )	
۷/۸۱۴	۲/۵۰۱	۰/۲۵۹				

مأخذ: یافته‌های تحقیق

بررسی ساختار وابستگی .....

مطالعات صورت گرفته در این زمینه نظیر مطالعه گودوین، (۱۷) و بوکوشوا (۵) توزیع وایبل را به عنوان توزیع مناسب برای عملکرد محصول انتخاب نموده‌اند. در حالی که در این پژوهش توزیع وایبل دومین گزینه مناسب می‌باشد. از این رو با توجه به نتایج فوق و نیز مطالعات قبلی هر دو توزیع ویکبای و وایبل برای متغیر عملکرد گندم دیم میانه در نظر گرفته شد که مشخصات این دو توزیع به قرار زیر است:

$$yeild \sim WAK(\xi = 308.31, \alpha = 1475.6, \beta = 3.2585, \gamma = 327.58, \delta = -0.47761)$$

$$yeild \sim WEB(\alpha = 3.2723, \beta = 960.82)$$

که در توزیع ویکبای،  $\xi$  پارامتر موقعیت<sup>۲۲</sup>،  $\alpha$ ،  $\beta$  پارامترهای مقیاس<sup>۲۳</sup> و  $\delta$ ،  $\gamma$  پارامترهای شکل<sup>۲۴</sup> می‌باشند. در توزیع وایبل نیز  $\alpha$  و  $\beta$  به ترتیب پارامترهای شکل و مقیاس اند.

پس از به دست آوردن توزیع توأم مناسب، لازم است که متغیر عملکرد شبیه‌سازی شود؛ لذا ۱۰۰۰۰ متغیر تصادفی از داده‌ها با استفاده از الگوی C-vine تولید می‌شوند. با توجه به اینکه داده‌های شبیه‌سازی شده به صورت داده‌های مفصل یعنی در فاصله بسته صفر و یک می‌باشند، لازم است به سطح تجربی خود برگردانده شوند که برای این منظور از وارون تابع توزیع تجمعی هر توزیع استفاده می‌شود. مقدار متوسط شبیه‌سازی شده عملکرد با توجه به توزیع وایبل برای سال آبی ۸۶۰/۵۷۷۱ کیلوگرم در هکتار و با توجه به توزیع ویکبای ۸۷۱/۶۸۳۹ کیلوگرم در هکتار می‌باشد. با توجه به نتایج ارائه شده در جدول ۷ و آزمون‌های سه‌گانه به نظر می‌رسد نتایج مقادیر شبیه‌سازی شده با توجه به توزیع ویکبای دقیق‌تر از مقادیر مربوط به توزیع وایبل باشد. برای بررسی بیشتر وجود تفاوت معنی‌دار بین مقادیر شبیه‌سازی شده از طریق دو توزیع و مقادیر واقعی عملکرد محصول از آزمون t استفاده شد. فرض صفر در این آزمون صفر بودن میزان تفاوت بود. مقدار آماره t محاسباتی و p-value برای بررسی وجود تفاوت معنی‌دار بین مقادیر شبیه‌سازی شده از طریق توزیع وایبل و مقادیر واقعی به

---

22. Location

23. Scale

24. Shape

ترتیب برابر ۳/۹۸ و صفر بود در حالی که این مقادیر در مورد مقادیر شبیه‌سازی شده از طریق توزیع ویکبای و مقادیر واقعی عملکرد ۰/۵۷- و ۰/۵۶ بود. نتایج مربوط به این آزمون نشان می‌دهد که مقادیر شبیه‌سازی شده از طریق توزیع ویکبای به مقادیر واقعی عملکرد نزدیک‌تر هستند و از نظر آماری تفاوت معنی‌داری بین مقادیر آنها وجود ندارد؛ در حالی که در مطالعه بوکوشوا (۵) و گودوین (۱۷) توزیع وایبل به دلیل چولگی بیش از حد متغیر عملکرد، توزیع مناسب‌تری بود.

### نتیجه‌گیری و پیشنهادها

آذربایجان شرقی از قطب‌های تولید گندم در کشور است و شهرستان میانه با اختصاص ۲۰ درصد سطح زیر کشت و ۲۰ درصد تولید استان بزرگ‌ترین تولیدکننده گندم در این استان می‌باشد. در مناطقی که کشاورزی به صورت دیم صورت می‌گیرد، مهم‌ترین ریسک‌هایی که باعث کاهش عملکرد محصول می‌شود مربوط به تغییرات آب و هوایی می‌باشد که قسمت عمده آن مربوط به خشکسالی و یا سرمازدگی است. با توجه به جایگاه گندم در تأمین امنیت غذایی در کشور، به کارگیری روش‌های نوین برای مدیریت ریسک در آن امری ضروری می‌باشد. یکی از راه‌های مدیریت ریسک گندم دیم، بیمه آب و هوایی می‌باشد. در این سیستم بیمه‌ای به منظور محاسبه حق بیمه نیاز به شناخت متغیرهای آب و هوایی مؤثر در رشد گیاه و مدل‌سازی ساختار وابستگی عملکرد و این متغیرها به منظور شبیه‌سازی عملکرد می‌باشد. لذا در این پژوهش ساختار وابستگی بین عملکرد گندم دیم و متغیرهای آب و هوایی مدل‌سازی شد. برای این منظور داده‌های مربوط به عملکرد و متغیرهای آب و هوایی در سال‌های زراعی ۹۲-۱۳۶۶ به ترتیب از جهاد کشاورزی و سازمان هواشناسی استان آذربایجان شرقی گردآوری گردید. در چند سال اخیر، توابع مفصل تاکی شکل به عنوان مناسب‌ترین روش برای اندازه‌گیری ساختار وابستگی و بیان تابع توزیع توأم در زمینه‌های مختلف مالی و بیمه معرفی شده‌اند. بنابراین در این مطالعه برای شبیه‌سازی و پیش‌بینی عملکرد



بررسی ساختار وابستگی .....

گندم دیم از این توابع مفصل استفاده گردید. بررسی ساختار وابستگی متغیرهای آب و هوایی و عملکرد نشان داد که این متغیرها وابستگی متقارن نداشته و وابستگی دنباله‌ای قوی‌تری در دنباله پایین دارند. به عبارتی وقتی مقادیر این متغیرها بالا باشد زیاد تحت تأثیر همدیگر قرار نمی‌گیرند ولی در مقادیر کم به شدت بر هم اثر می‌گذارند. این نتایج مطابق با یافته‌های مطالعه بوکوشوا (۵) و شولت و برگ (۲۵) است که نشان دادند متغیرهای آب و هوایی در صورت کاهش میزان به شدت بر یکدیگر اثر می‌گذارند. با تبیین تابع توزیع توأم متغیرهای آب و هوایی و عملکرد گندم دیم مقادیر شبیه‌سازی شده عملکرد به دست آمد. نتایج آزمون t نشان داد که مقادیر شبیه‌سازی شده از طریق توزیع ویکبای در مقایسه با توزیع وایبل به مقادیر واقعی عملکرد نزدیک‌تر هستند. توزیع وایبل در مقایسه با توزیع ویکبای دم کشیده‌تری دارد و پراکنش نسبت به میانگین بیشتری را نشان می‌دهد در حالی که داده‌های عملکرد گندم دیم در شهرستان میانه پراکنش زیادی نداشتند و به مقدار میانگین خود نزدیک بودند. به این ترتیب میزان متوسط شبیه‌سازی شده عملکرد گندم دیم با استفاده از روش مفصل تاکی شکل کانونی و با توجه به توزیع ویکبای ۸۷۱/۶۸۳۹ کیلوگرم در هکتار به دست آمد. صندوق بیمه محصولات کشاورزی در طراحی بیمه شاخص آب و هوا برای محصولات دیم نیاز به روشی دارد که به طور دقیق خسارت مورد انتظار عملکرد را در اثر متغیرهای آب و هوایی تعیین نماید. استفاده از روش تابع مفصل تاکی شکل این امکان را می‌دهد تا متغیرهای آب و هوایی بیشتری را در کنار عملکرد محصول مورد بررسی قرار داد و تابع توزیع توأم شرطی واقعی و منعطف‌تری را برآورد نمود. در نتیجه خسارت مورد انتظار محاسبه شده و حق بیمه برآوردی واقعی‌تر خواهد بود. این امر به کاهش هزینه اجرایی در محاسبه خسارت و کاهش حق بیمه منجر می‌شود و از آنجا که بیش از ۶۵ درصد حق بیمه را دولت می‌پردازد، به این ترتیب هزینه‌های دولت نیز کاسته می‌شود. از طرف دیگر پایین بودن حق بیمه منجر به رغبت بیشتر کشاورز در تقاضای این سیستم بیمه‌ای خواهد شد. لازم به ذکر است که شرایط آب و هوایی در یک منطقه به مرور زمان دچار تغییرات اساسی می‌شود، بنابراین برای کاهش ریسک پایه

توصیه می‌شود این مطالعات هر چند سال یک‌بار برای منطقه تکرار شود. برای کنار گذاشتن سایر عوامل مانند کارایی مدیران، رقم مورد استفاده محصول و بافت خاک لازم است مناطق همگن در این طرح گنجانده شوند. برای افزایش دقت، لازم است میزان اثرگذاری هر متغیر آب و هوایی بر عملکرد محصول دیم در مراحل مختلف فنولوژیک در یک شرایط آزمایشگاهی بررسی گردد.

#### منابع

1. Aas, K., Czado, C., Frigessi, A. and Bakken, H. (2009). Pair-copula constructions of multiple dependence. *Insurance: Mathematics and Economics*, 44: 182-198.
2. Azizi, Gh. And Yarahmadi, D. (2003). Investigation of climatic parameters and dry farmed wheat relationship using regression equation. *Geographical Research Quarterly*, 35:23-29. (Persian)
3. Bedford, T. and Cooke, R.M. (2002). Vines: a new graphical model for dependent random variables. *Annals of Statistics*, 30: 1031-1068.
4. Bedford, T. and Cooke, R.M. (2001). Probability density decomposition for conditionally dependent random variables modeled by vines. *Annals of Mathematics and Artificial Intelligence*, 32: 245-268.
5. Bokusheva, R. (2010). Measuring the dependence structure between yield and weather variables. ETH Zurich, Institute for Environmental Decisions.

6. Brechmann, E.C. and Czado, C. (2011). Risk management with high-dimensional vine copulas: an analysis of the Euro Stoxx 50. Submitted for publication.
7. Brechmann, E.C. and Schepsmeier, U. (2012). Modeling dependence with C- and D-vine copulas: the R- package CDVine. To appear in the Journal of Statistical Software.
8. Brechmann, E.C., Czado, C. and Aas, K. (2010). Truncated regular vines and their applications. *Canadian Journal of Statistics*, 40(1): 68–85.
9. Chen, S., Wilson, W.W., Larsen, R. and Dahl, B. (2013). Investing in agriculture as an asset class. Department of Agribusiness and Applied Economics Agricultural Experiment Station North Dakota State University.
10. Cooke, R.M., Morales, O. and Kurowicka, D. (2007). Vines in overview. Invited Paper Third Brazilian Conference on Statistical Modeling in Insurance and Finance Maresias.
11. Czado, C., Brechmann, E.C. and Gruber, L. (2014). Selection of vine copulas. Technische Universität München.
12. Dibmann, J., Brechmann, E.C., Czado, C. and Kurowicka, D. (2013). Selecting and estimating regular vine copulae and application to financial returns. *Computational Statistics & Data Analysis*. 59: 52–69.
13. Farajzadeh, M. and Zarrin, A. (2002). Modeling the amount of dry wheat yield with respect to agricultural climatic criteria in West Azerbaijan Province. *Tarbiat Modares University Press*, 25: 71-96. (Persian)

14. Farajzadeh, M., Khorani, A., Bazgeer, S. and Zeaeian, P. (2011). Modeling and predicting of rainfed wheat yield in attention to phenological phases of plant growth (A case study for Kurdistan Province). *Physical Geography Research Quarterly*, 76: 21-34. (Persian)
15. Fischer, M. (2002). Tailoring copula-based multivariate generalized hyperbolic secant distributions to financial return data: an empirical investigation. Institute of Statistics and Econometrics University of Erlangen-Nurnberg, Lange Gasse 20, D-90403 Nurnberg, Germany.
16. Goodwin, B.K., Holt, M.T., Onel, G. and Prestemon, J.P. (2011). Copula-based nonlinear models of spatial market linkages. *American Journal of Agricultural Economics*, in press, 2011.
17. Goodwin, BK. (2012). Copula-based models of systemic risk in US. agriculture: implications for crop insurance and reinsurance contracts. The NBER conference on Insurance Markets and Catastrophe Risk in Boston.
18. Joe, H. (1997). Multivariate models and dependence concepts. Chapman and Hall, London
19. Kamali, Gh. And Bazgir, S. (2008). Dry wheat yield prediction using meteorological indices in some parts of Iran western. *Journal of Agricultural Science and Natural Resources*, 2(64): 113-121. (Persian)
20. Kurowicka, D. and Cooke, R.M. (2006). Uncertainty analysis with high dimensional dependence Modeling. John Wiley & Sons Ltd.
21. Kurowicka, D. and Joe, H. (2011). Dependence modeling: vine copula handbook. World Scientific Publishing Co, Singapore.

بررسی ساختار وابستگی .....

22. Ministry of agriculture-Jahad. (2015). Center for Statistics and Information, production cost of crops Available at: <http://www.maj.ir/Index.aspx?page=form&lang=1&PageID=11583&tempname=amar&sub=65&methodName>ShowModuleContent>.
23. Nelsen, R.B. (2005). An introduction to copulas. Second Edition. Springer-Verlag, Berlin.
24. Scholzel, C. and Friederichs, P. (2008). Multivariate non-normally distributed random variables in climate research introduction to the copula approach. *Nonlinear Processes in Geophysics*, 15: 761–772.
25. Schulte, G.M. and Berg, E. (2011). Modeling farm production risk with copula instead of correlations. Institute of Food and Resource Economics, University of Bonn, Germany.
26. Zare abyaneh, H. (2013). Evaluating roles of drought and climatic factors on variability of four dry farming yields in Mashhad and Birjand. *Water and Soil Science*, 23(1):39-56. (Persian)
27. Zhu, Y., Ghosh, S. and Goodwin, B. (2008). Modeling dependence in the design of whole farm insurance contract a copula-based approach. Contributed paper at the Annual Meeting of the AAEA, Orlando, USA, July 27-29.