



## مدل سازی فضایی-زمانی وقوع و مقدار بارش زمستانه در گستره ایران با استفاده از مدل مارکف پنهان

مهدی قمقامی<sup>۱</sup>، نوذر قهرمان<sup>۲</sup> و جواد بذرافشان<sup>۳</sup>

۱ و ۳- دانشجوی دکتری و استادیار، دانشگاه تهران  
 ۲- دانشیار، دانشگاه تهران، (نویسنده مسوول: nghahreman@ut.ac.ir)  
 تاریخ دریافت: ۹۲/۴/۲۵ تاریخ پذیرش: ۹۲/۹/۱۱

### چکیده

مدل سازی چندمکانی بارش یکی از زمینه‌های مهم در علوم طبیعی است و مدل‌های مختلف آماری برای این مهم توسعه یافته‌اند که نگرشی فضایی به مدل‌سازی و شبیه‌سازی بارش روزانه دارند. مدل مارکف پنهان یکی از انواع مدل‌های چندمکانی بارش روزانه است که علاوه بر شبیه‌سازی بارش روزانه، به بررسی توزیع فضایی و زمانی الگوهای وقوع بارش نیز می‌پردازد. در مطالعه حاضر با بکارگیری مدل مارکف پنهان، اقدام به مدل‌سازی بارش زمستانه (ژانویه تا آوریل) براساس داده‌های ۱۳۰ ایستگاه باران‌سنجی در گستره ایران شده است. طول دوره آماری داده‌های بارندگی روزانه ۲۱ سال (۲۰۱۰-۱۹۹۰) می‌باشد. هدف این تحقیق، دستیابی به توزیع فضایی و زمانی الگوهای جوی و شبیه‌سازی تصادفی وقوع و مقدار بارش به‌طور همزمان است. نیل به این اهداف با بکارگیری الگوریتم‌های مختلف حل مسائل مدل مارکف پنهان شامل الگوریتم برآورد پارامتر بیشینه‌سازی امید (EM) و الگوریتم رمزگشایی ویتربی و یک الگوریتم شبیه‌سازی که با تبدیل احتمالاتی همراه است، میسر است. انتظار می‌رود که الگوهای جوی مستخرج از مدل مارکف پنهان، منطبق بر الگوهای سینوپتیکی باشند که بر این اساس ۸ الگوی مختلف جوی به‌عنوان مجموعه متناهی حالات پنهان به‌دست آمدند و هرکدام گویای الگوی خاص سینوپتیکی هستند. فراوان‌ترین الگوی توزیع بارندگی مستخرج از مدل مارکف پنهان، الگوی خشک (شرایط پایدار) بوده که طی آن احتمال وقوع بارش در اکثر ایستگاه‌های کشور پایین است. این الگو دارای بیشترین احتمال آغازین معادل ۰/۴۲۹ و بیشترین احتمال انتقال مارکفی از حالت مشابه معادل ۰/۶۳۷ می‌باشد. همچنین، شبیه‌سازی چندمکانی بارش زمستانه نیز با حفظ آماره‌های اساسی شامل میانگین و انحراف معیارهای مجموع فصلی ایستگاه‌ها، مقدار صدک‌های مختلف هر ایستگاه و وابستگی‌های فضایی وقوع و عدم وقوع بارش، دارای نتایج قابل قبولی است که در مجموع بکارگیری این رهیافت، سودمند ارزیابی شده است.

واژه‌های کلیدی: مدل مارکف پنهان، بارش زمستانه، الگوی فضایی، شبیه‌سازی

### مقدمه

می‌گردد به‌طوری‌که با اجرای یکبار مدل سری‌های مصنوعی برای ایستگاه‌های متعدد تولید و در اختیار کاربر برای اهداف بعدی قرار می‌گیرد.

برای این نوع مدل‌سازی رهیافت‌های متعددی نظیر رهیافت Wilks (۱۵)، رهیافت نزدیک‌ترین همسایه یا KNN (۱۰)، رهیافت‌های مبتنی بر مارکف پنهان (۸،۱) توسعه یافته است. رهیافت Wilks توسط ویلکز (۱۵) با تعمیم مدل مارکف مرتبه نخست تک‌مکانی به چندین ایستگاه با حفظ وابستگی‌های فضایی توسعه یافته است. مهرترا و همکاران (۱۰) کارکرد این رهیافت را با دو رهیافت دیگر مبتنی بر مارکف پنهان و نزدیک‌ترین همسایه مقایسه و هر سه رهیافت را در شبیه‌سازی ویژگی‌های زمانی و فضایی با وجود معایب و مزایایی برای هر کدام رضایت‌بخش ارزیابی کردند. مدل‌های استوکاستیک مبتنی بر مارکف پنهان نیز در دو نوع کلی همگن و ناهمگن توسط محققین مختلفی مورد مطالعه قرار گرفته است. این مدل‌ها ابتدا توسط رابینر و جوانگ

مدل‌سازی فضایی و زمانی بارش روزانه تاکنون توسط محققین مختلفی مانند هاگز و همکاران (۹،۸)، ویلکز (۱۵)، بلون و همکاران (۳)، مهرترا و همکاران (۱۰)، تامسون و همکاران (۱۴)، ایلپوت و همکاران (۱) مورد مطالعه قرار گرفته است. هدف اصلی از این نوع مدل‌سازی، حفظ وابستگی‌های فضایی در سری‌های شبیه‌سازی شده است. چراکه اگر مدل‌سازی به‌صورت تک ایستگاهی (Single site) انجام گیرد، این وابستگی‌ها که ریشه در ساختار فیزیکی جو دارد، به درستی ساخته نمی‌شوند. این واقعیت از آنجا نشأت می‌گیرد که وقوع بارندگی به‌خصوص بارش‌های زمستانه در ایران، بیشتر تحت تاثیر حرکت و برخورد توده‌های هوا در مقیاس سینوپتیکی است تا فعل و انفعالات جو در مقیاس محلی، که در نتیجه آن بارندگی تحت تاثیر یک الگوی خاص سینوپتیکی در یک گستره به وقوع می‌پیوندد. همچنین دستیابی به یک مدل چندمکانی (Multisite) منجر به صرفه‌جویی در هزینه و زمان

می‌کند. با برآورد پارامترهای بهینه، رمزگشایی حالات پنهان جوی که عبارت است از دستیابی به محتمل‌ترین سری یا توالی زمانی الگوهای فضایی وقوع و میزان بارندگی در یک گستره و فرآیند شبیه‌سازی که به صورت استوکاستیک انجام می‌شود، میسر است. الگوریتم به‌کار رفته برای فرآیند رمزگشایی، الگوریتم ویتربی است که در ابتدا برای رمزگشایی کدهای پیچیده طراحی شد و اکنون نیز در حوزه‌های زیادی مانند پیش‌بینی توالی‌های بیولوژیکی، تشخیص گفتار و غیره به‌کار گرفته می‌شود (۱۱). این الگوریتم به یافتن محتمل‌ترین توالی پنهان به ازای یک توالی قابل رویت معین می‌پردازد. همچنین الگوریتم شبیه‌سازی استوکاستیک شامل یک تبدیل احتمالاتی بردار فضایی بارش به بردار متغیرهای نرمال است که روی هر حالت شرط بسته می‌شوند. علت این تبدیل احتمالاتی، ساده‌تر بودن شبیه‌سازی تصادفی با مجموعه داده‌های نرمال است (۲).

رهیافت به‌کار رفته در تحقیق حاضر برگرفته از روش پیشنهادی ایلپوت و همکاران (۱) است و در آن سه هدف دنبال می‌شود: ۱- دستیابی به بهترین الگوهای توزیع فضایی احتمال وقوع بارش براساس انتخاب مناسب‌ترین تعداد حالات پنهان با معیار BIC، ۲- رمزگشایی توالی این الگوها با استفاده از الگوریتم ویتربی و ۳- شبیه‌سازی استوکاستیک بارش روزانه.

#### مواد و روش‌ها داده‌ها

تمرکز این مطالعه بر بارش زمستانه است، چرا که بخش قابل توجهی از بارندگی‌ها در ایران طی فصل زمستان (ژانویه تا آوریل) اتفاق می‌افتد. در این مطالعه از داده‌های شبکه‌ای ایستگاه‌های بارانسنجی منتخب مشتمل بر ۱۳۰ ایستگاه، استفاده شده است. این داده‌ها در مقیاس روزانه و متعلق به دوره آماری ۲۰۱۰-۱۹۹۰ (۲۱ سال) هستند. بر این اساس، ۲۱ سری زمانی مستقل مربوط به فصل زمستان در هر ایستگاه در دسترس است. به عبارت دیگر، فرض می‌شود که وقوع و مقدار بارش طی زمستان هر سال از سال‌های دیگر مستقل باشد که البته می‌تواند صحیح باشد. موقعیت جغرافیایی ایستگاه‌ها در شکل ۱ آورده شده است. با توجه به حوزه گسترده مطالعه حاضر در سطح کشور، این قابلیت وجود دارد که کارکرد رهیافت در اقالیم مختلف اعم از خشک (کم باران) و مرطوب (پر باران) بررسی شود.

(۱۲) و زوخینی و گاتورپ (۱۶) و هوکاگلو (۷) برای استفاده در مطالعات اقلیمی ارائه و سپس توسط هاگز و گاتورپ (۸) به‌خصوص در شبیه‌سازی الگوهای فضایی و زمانی وقوع و مقدار بارندگی تغییر یافته‌اند. این مدل بعدها توسط هاگز و همکاران (۹)، بلون و همکاران (۳)، تامسون و همکاران (۱۴) و ایلپوت و همکاران (۱) در شبیه‌سازی فضایی و زمانی الگوهای وقوع بارش مورد مطالعه قرار گرفته است. شریکانتن و همکاران (۱۳) نیز از این مدل در شبیه‌سازی بارش سالانه در استرالیا استفاده کردند. در مدل‌های مبتنی بر کار هاگز و گاتورپ (۸) حالات متعدد جوی (weather state) که منجر به وقوع و عدم وقوع بارش در یک گستره مکانی می‌شود به‌عنوان حالات پنهان و سری زمانی روزانه بارش به‌عنوان توالی‌های قابل رویت به‌صورت پیوسته (سری زمانی مقادیر) یا گسسته (سری زمانی وقوع و عدم وقوع بارش) در نظر گرفته می‌شود. واژه "پنهان" اشاره به این واقعیت دارد که فرآیند حالت جوی یک کمیت غیرقابل رویت است (۱۰) و تنها از روی اثر آن (لایه قابل رویت) قابل تخمین است. دستیابی به این حالات، محققین را در درک توزیع الگوهای سینوپتیکی کمک می‌کند. تمایز رهیافت ایلپوت و همکاران (۱) با سایر محققین این است که آنها توالی قابل رویت را یک توالی پیوسته شامل وقوع و مقدار بارش در نظر گرفته‌اند. این مسئله باعث می‌شود که هر دو ویژگی مربوط به متغیر بارندگی روزانه (وقوع و مقدار) به‌طور شرطی به حالات پنهان جوی وابسته باشند که این مطلب به واقعیت نزدیک‌تر است. برخلاف مدل‌های مارکف پنهان کارشده تا قبل از کار این محققین که تنها وقوع را وابسته به حالات پنهان در نظر می‌گرفتند و سپس به شبیه‌سازی مقدار بارندگی بر اساس یک مدل آماری می‌پرداختند.

به‌طور کلی با فرض وجود وابستگی‌های پیایی بین توالی‌های پنهان و وابستگی‌های متقاطع بین توالی‌های پنهان و قابل رویت، هر مدل مارکف پنهان مجموعه‌ای است از ماتریس‌های احتمال انتقال و احتمال انتشار که در صورت دسترسی به هر دو توالی، محاسبه آنها به سادگی انجام می‌شود و در غیر این صورت از الگوریتم‌های برآورد پارامتر نظیر الگوریتم بیشینه‌سازی امید (EM) استفاده می‌شود (۵). این الگوریتم برای برآورد پارامترهای مدلی به کار می‌رود که بخشی از آن مدل، پنهان باشد و یا در دسترس نباشد، و دو فرآیند محاسبه امیدریاضی تابع به ازای پارامترهای به‌دست‌آمده در مرحله قبل و بیشینه‌سازی تابع امید ریاضی را تا رسیدن به همگرایی طی می‌کند و در نهایت پارامترهای بهینه را برآورد



شکل ۱- موقعیت جغرافیایی ایستگاه‌های بارشسنجی در ایران

پنهان باید از این بردار به جای یک عدد استفاده نمود. هر بردار  $Y_t$  به‌عنوان یک عنصر در توالی قابل رویت از یک حالت جوی متعلق به توالی پنهان که با  $S_t \in N$  که نمایش داده می‌شود، گسیل می‌شود که  $N$  تعداد حالات جوی است. فرض‌های مرسوم یک مدل مارکف پنهان که منجر به ساده‌سازی آن شده است، شامل دو فرض اساسی است که در روابط ۱ و ۲ آمده است:

$$P(S_t | S_1, S_2, \dots, S_{t-1}, Y_1, Y_2, \dots, Y_{t-1}) = P(S_t | S_{t-1}) \quad (1)$$

$$P(Y_t | Y_1, Y_2, \dots, Y_{t-1}, S_1, S_2, \dots, S_{t-1}, S_t) = P(Y_t | S_t) \quad (2)$$

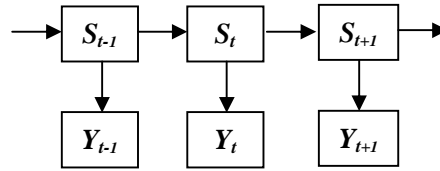
فرض اول اشاره به خاصیت مارکفی مرتبه یک دارد و مربوط به توالی پنهان و فرض دوم اشاره به وجود وابستگی متقاطع و همزمان بین عناصر توالی قابل رویت با توالی پنهان دارد. به عبارت دیگر فرض اول حاکی از آن است که الگوهای جوی دارای وابستگی پیاپی مرتبه یک هستند و محاسبات آن منجر به تولید ماتریس‌های احتمال انتقال (از حالتی به حالت دیگر) و فرض دوم حاکی از آن است که بارش روز معین تنها وابسته به حالت جوی حاکم بر آن روز و مستقل از بارش روز یا روزهای قبل است و محاسبات فرض دوم منجر به تولید ماتریس‌های احتمال انتشار یا گسیل می‌شود. همچنین این دو فرض مدل مارکف پنهان، استقلال فضایی داده‌های بارندگی را به همراه دارند، به‌طوری‌که بارندگی در ایستگاه  $i$  نسبت به بارندگی در ایستگاه  $j$  ( $i \neq j$ ) مستقل است (۴،۱) و وابستگی‌های فضایی به ازای یک حالت معین جوی مدل می‌شوند (معادله ۳). با این

#### مدل مارکف پنهان

رهیافت مورد استفاده مبتنی بر مدل مارکف پنهان است که به شبیه‌سازی الگوهای توزیع زمانی و فضایی احتمال وقوع بارش می‌پردازد. همچنین این رهیافت می‌تواند نقش یک مولد هواشناسی چندممکنی را ایفا نماید و سری‌های تصادفی از وقوع و مقدار بارندگی روزانه را با طول دلخواه در اختیار کاربر قرار دهد. برای نیل به اهداف مذکور، از الگوریتم‌های مختلف حل مسائل مربوط به مدل مارکف پنهان استفاده می‌شود. در گام اول، داده‌های بارندگی به‌عنوان توالی قابل رویت و حالت جوی (weather state) به‌عنوان توالی پنهان در نظر گرفته می‌شود. حالت جوی در یک روز معین اشاره به یک الگوی خاص دارد که کل کشور را فرا گرفته و منجر به وقوع یا عدم وقوع بارندگی در ایستگاه‌های مختلف و با احتمالات متفاوت می‌شود. این الگو بعداً می‌تواند به‌عنوان یک الگوی سینوپتیکی که قابل استخراج از نقشه‌های فشار سطح زمین و سطوح فوقانی است تفسیر شود. به‌طوری‌که انتظار می‌رود در صورت کارکرد صحیح مدل مارکف پنهان، الگوهای فضایی مستخرج از آن، منطبق بر الگوهای سینوپتیکی گرفته شده از نقشه‌ها باشد. پس در یک نگاه کلی، مدل مارکف پنهان حاضر مشتمل بر یک توالی قابل رویت (بارش روزانه در ایستگاه‌ها) و یک توالی پنهان (حالات حاکم بر جو) است. اگر  $Y_t^k$  مقدار بارندگی در روز  $t$  و در ایستگاه  $k$  باشد، آنگاه  $Y_t = (Y_t^1, Y_t^2, \dots, Y_t^K)$  که  $K=130$  بردار فضایی مقدار بارندگی در روز  $t$  به حساب می‌آید. برای دستیابی به الگوی فضایی بارش‌ها در هر مرحله اجرای مدل مارکف

توضیحات نمای کلی مدل مارکف پنهان در شکل ۲ نشان داده شده است.

$$P(Y_t | S_t) = \prod_{k=1}^K P(Y_t^k | S_t) \quad (3)$$



شکل ۲- الگوی کلی مدل مارکف پنهان

زمانی الگوهای جوی است می‌پردازند. الگوریتم ویتربی (Viterbi) به‌عنوان متداول‌ترین روش رمزگشایی توالی پنهان در این تحقیق مدنظر است. دسترسی به توزیع زمانی الگوهای حاکم بر جو، محققین را در درک تغییرات زمانی این الگوها که در مبحث تغییر اقلیم کاربرد دارد، کمک می‌نماید.

#### مساله رمزگشایی و الگوریتم ویتربی

در مساله رمزگشایی مدل مارکف پنهان، هدف، یافتن توالی بهینه از حالات پنهان است. به‌عبارتی، مساله رمزگشایی به‌دنبال تعیین توالی پنهان ازای یک توالی قابل رویت  $O = \{o_1, o_2, \dots, o_T\}$  (در اینجا، الگوهای جوی) به  $Q = \{q_1, q_2, \dots, q_T\}$  (در اینجا، بارش زمستانه) است (یعنی:  $Q^* = \{q_1^*, q_2^*, \dots, q_T^*\}$ ). تعریفی از بهینه بودن، عبارت است از انتخاب محتمل‌ترین توالی پنهانی که به ازای یک توالی قابل رویت وجود دارد (۱۱):

$$Q^* = \arg \max_Q p[Q, O | \#] \quad (4)$$

در این معادله  $\#$  مجموع پارامترهای مدل است. دستیابی به چنین توالی، به صورت تحلیلی مستلزم بررسی احتمال  $N^T$  توالی است که  $N$  تعداد حالات پنهان و  $T$  طول توالی است. با این تفسیر حتی برای توالی‌های کوتاه و با تعداد حالات کم نیز، تعداد توالی‌های قابل بررسی زیاد و کار مشکل می‌شود.

هم‌چنین در حل این مساله باید از انتخاب توالی نامحتمل پرهیز کرد. یک روش کارآمد برای اجتناب از چنین توالی‌های نامحتمل استفاده از الگوریتم ویتربی است که بر مبنای برنامه‌نویسی پویا استوار است. در این الگوریتم دو متغیر  $U_r(i)$  و  $E_r(j)$  تعریف می‌شوند که به ترتیب عبارتند از، بزرگترین احتمال در مسیر منفردی که برای  $t$  مشاهده اول در نظر گرفته و به حالت  $s_i$  ختم می‌شود و

علاوه بر این دو ماتریس، بردار احتمالات اولیه که عبارت است از احتمال وقوع هر حالت در مرحله آغازی یا شرایط مرزی، نیز برای مدل‌سازی باید در نظر گرفته شوند. با وجود چنین مجهولاتی که در اثر ناآگاهی از توالی پنهان است، مسائل متعددی پیش‌روی مدل‌سازی مارکف پنهان واقع می‌شوند که حل به شیوه معمول آنها امکان‌پذیر نیست و محققین را به سمت روش‌های برآورد سوق می‌دهد.

حل مسائل مختلف در مدل‌سازی مبتنی بر مارکف پنهان براساس الگوریتم‌های متعددی میسر است. این الگوریتم‌ها شامل الگوریتم‌های برآورد پارامترهای مدل، الگوریتم‌های رمزگشایی از توالی حالات پنهان و الگوریتم‌های شبیه‌سازی استوکاستیک توالی مشاهده شده می‌باشند. در تحقیق حاضر هرکدام از این الگوریتم‌ها هدف خاصی را دنبال می‌کنند. الگوریتم برآورد پارامتر یا الگوریتم آموزش مدل به برآورد پارامترهای مدل شامل ماتریس احتمالات انتقال، ماتریس احتمالات گسیل، ماتریس احتمالات اولیه و سایر پارامترها (که حسب هر مدل ممکن است متفاوت باشند) می‌پردازد. علت استفاده از این الگوریتم‌ها، در درجه اول عدم دسترسی به توالی حالات پنهان برای آموزش مدل است که محققین به‌طور معمول با آن روبرو هستند. در این‌جا نیز، به ازای یک تعداد معین از حالات جوی، مدل آموزش می‌بیند و الگوهای فضایی احتمال وقوع بارندگی بر اساس این تعداد حالات به‌دست می‌آیند. به علت عدم آگاهی از تعداد حالات مناسب جوی که هرکدام حکایت از الگوی خاصی دارند، مدل به ازای تعداد حالات متفاوت آموزش می‌بیند و بهترین مدل با معیار BIC انتخاب و برای اهداف بعدی دنبال می‌شود. الگوریتم مرسوم برای این نوع کارکرد، الگوریتم EM نام دارد که در جلوتر به تشریح ساختار ریاضی آن پرداخته می‌شود. الگوریتم‌های رمزگشایی، به تولید محتمل‌ترین توالی حالات پنهان که در این‌جا توزیع

نظر می‌گیرند. در حالی که در واقعیت، چندین نوع الگوی جوی بارندگی، مقادیر متفاوتی از مقدار بارش دریافتی به‌وسیله یک ایستگاه معین را به همراه دارند. این مساله با ایجاد یک توزیع چندمتغیره نرمال قابل حل است معادله (۷) یک بردار چندمتغیره نرمال را با استفاده از توان  $S^{(s)}$  به یک بردار تصادفی با توزیع چندمتغیره ترکیبی (پیوسته-گسسته) تبدیل می‌کند. به‌طور دقیق‌تر اگر فرض شود که  $S = s S_r$  سپس:

$$Y_i(k) = \max(X_i(k), 0), \quad (Y)$$

$$X_i(k) = \begin{cases} W_i(k) & (W_i(k) \leq 0) \\ W_i(k)^{s^{(s)}(k)} & (W_i(k) > 0) \end{cases}$$

در این روابط  $W$  متغیری است که از توزیع نرمال تبعیت می‌کند. چنانچه مشاهده می‌شود.  $S^{(s)}$  یک توان مثبت با ازای حالت جوی  $s$  است که توزیع ترکیبی را به توزیع نرمال مرتبط می‌سازد (۲). دلیل استفاده از چنین تبدیلی در فرآیند شبیه‌سازی، ساده‌تر بودن کار با توزیع نرمال است، چراکه داده‌های بارندگی از توزیع بسیار چوله تبعیت می‌کنند و شبیه‌سازی تصادفی با آن مشکل است. اگر  $W_i = (W_i(1), W_i(2), \dots, W_i(K))$  معادل بردار فضایی ایستگاه‌ها باشد، آنگاه:

$$(s = 1, \dots, M) \quad W_i = m^{(s)} + H^{(s)} Z_i \quad (A)$$

$Z$  برداری است با توالی متغیرهای تصادفی نرمال استاندارد،  $m$  و  $H$  به ترتیب بردار  $K$  بعدی میانگین و انحراف معیار به ازای هر حالت معین  $s$  می‌باشند. با توجه به عدم آگاهی از توالی‌های پنهان (حالات جوی) محاسبه پارامترهای روابط (۷) و (۸) شامل  $s$ ،  $m$  و  $H$  به ازای حالت معین جوی، با الگوریتم EM امکان‌پذیر است. بنابراین روند شبیه‌سازی به این صورت است که بردارهای  $K$  بعدی میانگین و انحراف معیار و پارامتر بتا به ازای حالت معین  $s$  با روش برآورد پارامتر (بخش بعدی) برآورد و سپس با تولید تصادفی ابتدا توالی حالات پنهان شبیه‌سازی و در گام بعدی بردارهای  $W$  و  $Y$  در ادامه به ازای حالت معین شبیه‌سازی می‌شوند. با این توصیفات می‌توان شمای کلی مدل مارکف پنهان مورد بحث را از شکل ۲ به شکل ۳ تغییر داد. در واقع رهیافت ارائه شده توسط ایلپوت و همکاران (۱)، یک مدل غیرخطی زمانی- فضایی با دو لایه پنهان است: یک لایه متشکل از متغیر گسسته ( $S_r$ ) و لایه دوم متشکل از متغیر پیوسته ( $W_r$ ).

حالت سیستم در زمان  $t-1$  که منجر به این مسیر محتمل می‌گردد. این دو کمیت در طی زمان با روابط زیر محاسبه می‌شوند (۱۱):

$$u_i(j) = \max_{1 \leq i \leq N} \{u_{t-1}(i) p_{ij}\} b_j(o_t) \quad (5)$$

$$\mathbb{E}_t(j) = \arg \max_{1 \leq i \leq N} \{u_{t-1}(i) p_{ij}\} \quad 1 \leq j \leq N, \quad 2 \leq t \leq T$$

$$u_1(i) = f_i b_i(o_1) \quad 1 \leq i \leq N$$

$$\mathbb{E}_1(i) = 0$$

در این روابط،  $f_i, p_{ij}, b_j(o_t)$  به ترتیب عبارتند از: احتمال گسیل مشاهده  $o_t$  از حالت  $i$ ، احتمال انتقال حالت  $i$  به  $j$  و احتمال اولیه حالت  $i$ . در نهایت با پیمایش معکوس، محتمل‌ترین توالی بر اساس روابط زیر به‌دست می‌آید:

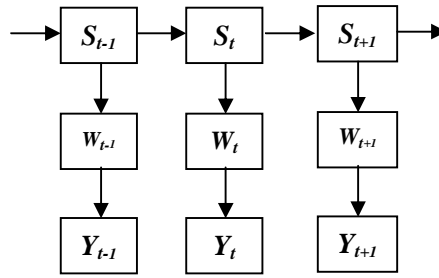
$$q_t^* = \mathbb{E}_{t+1}(q_{t+1}^*) \quad t = T-1, T-2, \dots, 1 \quad (6)$$

$$q_T^* = \arg \max_{1 \leq i \leq N} \{u_T(i)\}$$

الگوریتم شبیه‌سازی استوکاستیک

فرآیند شبیه‌سازی با هدف تطویل دوره‌ی آماری در چندین ایستگاه به‌طور همزمان انجام می‌شود. در این فرآیند سعی بر آن است که ویژگی‌های فضایی و زمانی سری‌های زمانی بارش روزانه اعم از ویژگی‌های مربوط به وقوع و عدم وقوع بارش و مقادیر به‌ازای حالت وقوع به‌خوبی در سری‌های مصنوعی حفظ شوند. رهیافت بکاررفته در مطالعه حاضر برگرفته از ایلپوت و همکاران (۱) است که بطور هم‌زمان وقوع و مقدار را به‌صورت تصادفی شبیه‌سازی می‌کند. این رهیافت مبتنی بر یک تبدیل توانی (power transformation) است که ماهیت گسسته-پیوسته بودن داده‌های بارندگی را مدیریت می‌کند.

چنانچه در مقدمه ذکر شد، تمایز این رهیافت با سایر رهیافت‌های کار شده توسط هاگز و همکاران (۹)، بلون و همکاران (۳)، مهرترا و همکاران (۱۰) و تامسون و همکاران (۱۴) این است که رهیافت‌های دیگر ابتدا متغیر گسسته (وقوع و عدم وقوع) را شبیه‌سازی و سپس به شبیه‌سازی مقدار بارش در حالت وقوع می‌پردازند. عیب اساسی این روش، عدم لحاظ وابستگی مقدار بارش به حالت جوی در فرآیند شبیه‌سازی است (۱). به عبارتی دیگر این رهیافت‌ها تنها حالت وقوع و عدم وقوع بارش (توالی قابل رویت) را وابسته به حالت جوی (توالی پنهان) فرض می‌کنند و مقدار آن را مستقل از نوع آب و هوا در



شکل ۳- الگوی تغییر یافته مدل مارکف پنهان

### نسبت لگاریتمی بخت‌ها

برای ارزیابی توانمندی مدل در ساخت وابستگی‌های فضایی وقوع بارش، از نسبت لگاریتمی بخت‌ها (Log-odds ratio) استفاده شده است. این نسبت که معیاری برای ارزیابی داده‌های باینری است، بازتاب کننده همبستگی فضایی بین وقوع بارش در یک جفت ایستگاه می‌باشد (۱۰). این نسبت به صورت زیر تعریف می‌شود:

$$lr_{i,j} = \log \frac{P_{i,j}^{00} P_{i,j}^{11}}{P_{i,j}^{10} P_{i,j}^{01}} \quad (11)$$

در این معادله،  $lr_{i,j}$  نسبت مذکور بین دو ایستگاه  $i$  و  $j$  ( $i \neq j$ ) است.  $P_{i,j}^{00}, P_{i,j}^{11}, P_{i,j}^{01}, P_{i,j}^{10}$  به ترتیب احتمالات توأم وقوع بارش در ایستگاه  $i$  و عدم وقوع در ایستگاه  $j$ ، عدم وقوع در ایستگاه  $i$  و وقوع در ایستگاه  $j$ ، وقوع در هر دو ایستگاه و عدم وقوع در هر دو ایستگاه است. هرچه مقدار این نسبت بزرگتر باشد، نشان از وابستگی بیشتر بین دو ایستگاه دارد، چرا که احتمالات وضعیت یکسان در دو ایستگاه (هر دو وقوع یا هر دو عدم وقوع) بزرگتر از احتمالات شرایط متفاوت در دو ایستگاه می‌شود. به‌طور معمول این نسبت دامنه‌ای بزرگتر از صفر دارد.

### نتایج و بحث

#### تعیین تعداد الگوهای جوی بهینه

داده‌های بارش روزانه مربوط به فصل زمستان (ژانویه تا آوریل) از سری زمانی روزانه بارش برای ۱۳۰ ایستگاه در سطح کشور استخراج و به فرمت مورد نظر درآمدند. گام نخست در ارائه نتایج، برآورد پارامترهای مدل مارکف پنهان است. با توجه به عدم آگاهی از تعداد حالات مناسب پنهان (تعداد حالات متمایز جوی)، پارامترهای مدل برای تعداد حالات متفاوت  $N=1, \dots, 9$  حالت به‌طور مجزا برآورد شدند و بر اساس معیار BIC مناسب‌ترین تعداد حالات انتخاب شد. در جدول ۱ مقادیر معیار مذکور برای تعداد حالات مختلف آورده شده است.

### مساله آموزش مدل و الگوریتم EM

آموزش مدل در ابتدای کار با هدف برآورد بهینه پارامترهای مدل ( $n$ ) به کار می‌رود. در مدل مارکف پنهان پیش‌رو  $(n^{(0)}, n^{(1)}, \dots, n^{(N)})$  به عنوان مجموعه پارامترهای ناشناخته در نظر گرفته می‌شود.  $n^{(0)}$  شامل احتمالات اولیه یا  $\pi_s = P(S_t = s)$  برای شرایط مرزی یا عناصر اول هر توالی که روزهای ابتدایی فصل زمستان در هر سال هستند و احتمالات انتقال یا  $P_{ij} = P(S_t = j | S_{t-1} = i)$  برای سایر عناصر توالی پنهان که بیانگر پویایی زنجیر مارکف هستند، می‌باشند. پارامترهای  $(s = 1, \dots, N)$  شامل مجموعه‌ای از پارامترهای  $S^{(s)}, H^{(s)}, m^{(s)}$  و ماتریس احتمالات گسیل  $P(Y_t | S_t = s)$  می‌باشد. به‌طور معمول از الگوریتم EM برای برآورد این مجموعه پارامتر استفاده می‌شود. این الگوریتم شامل دو مرحله می‌باشد: مرحله اول محاسبه امید ریاضی (Expectation, E-step) و مرحله دوم بیشینه‌سازی تابع درست‌نمایی (Maximization, M-step). الگوریتم EM ابتدا به‌طور تصادفی، مقادیر اولیه‌ای برای پارامترها در نظر می‌گیرد، سپس دو مرحله ذیل را تا رسیدن به همگرایی به‌طور متناوب تکرار می‌کند (۵). E-step: محاسبه امید ریاضی تابع درست‌نمایی با این شرط که پارامترهای مرحله قبل ( $n_n$ ) و داده‌های مشاهده شده ( $o$ ) وجود دارد. در این مرحله مقادیری برای پارامترها در نظر گرفته می‌شود و سپس تابع توزیعی مبتنی بر این پارامترها بر داده‌های  $o$  و  $q$  (داده‌های پنهان) برازش می‌یابد و امید ریاضی آن محاسبه می‌شود:

$$F(n | n_n) = E_{q|o, n_n} [\log L(n : o, q)] \quad (9)$$

M-step: الگوریتم در این مرحله با بیشینه‌سازی تابع  $F$  به‌دست‌آمده از مرحله قبل، پارامترها را برآورد می‌کند. با هر تکرار مقدار تابع درست‌نمایی افزایش و در نهایت به یک مقدار بیشینه نسبی همگرا می‌شود.

$$n_{n+1} = \arg \max_n F(n | n_n) \quad (10)$$

جدول ۱- معیار BIC برای انتخاب مناسبترین تعداد حالات آب و هوایی

| BIC      | تعداد حالات |
|----------|-------------|
| ۶۷۷۸۸۳/۱ | ۱           |
| ۶۰۹۳۱۶/۳ | ۲           |
| ۵۸۹۰۶۴/۱ | ۳           |
| ۵۷۸۶۲۸/۵ | ۴           |
| ۵۷۱۵۱۱/۷ | ۵           |
| ۵۷۰۲۲۳/۵ | ۶           |
| ۵۶۹۱۰۸/۸ | ۷           |
| ۵۶۸۷۹۰/۹ | ۸           |
| ۵۶۹۴۹۳/۹ | ۹           |

دارای ۸ الگوی متفاوت از نظر احتمال وقوع و مقدار هستند. در جدول ۲ ماتریس احتمالات اولیه و انتقال مربوط به این ۸ حالت پنهان آورده شده است. از جدول ۲ مشاهده می‌شود که بیشترین احتمال انتقال مربوط به حالت ۲ است.

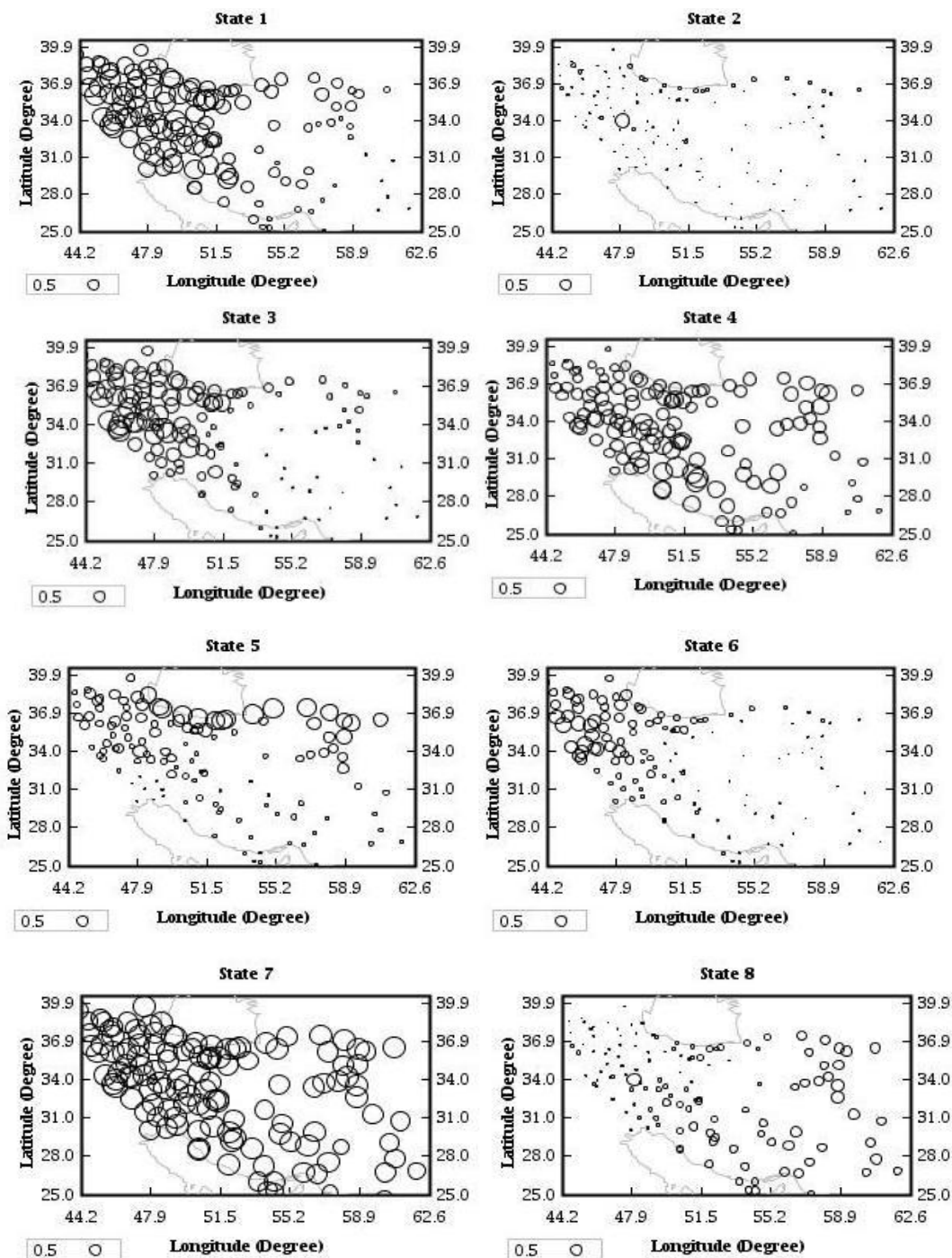
چنانچه مشخص است، معیار BIC از تعداد حالات معادل یک حالت، دارای روند کاهشی است که کمترین مقدار آن در تعداد حالات معادل ۸ حالت است. بر این اساس، مناسبترین تعداد حالات پنهان جوی، ۸ حالت متمایز است. به عبارت دیگر، بارش‌های زمستانه در ایران

جدول ۲- ماتریس احتمالات انتقال و اولیه برای بهترین تعداد حالات (معادل ۸ حالت)

| $f_s$ | ۸     | ۷     | ۶     | ۵     | ۴     | ۳     | ۲     | ۱     | $S_t$ | $S_{t+1}$ |
|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-----------|
| ۰/۰۴۸ | ۰/۰۸۴ | ۰/۰۰۹ | ۰/۰۲۶ | ۰/۲۷۶ | ۰/۳۳۵ | ۰/۰۴۳ | ۰     | ۰/۲۲۷ |       | ۱         |
| ۰/۴۲۹ | ۰/۰۳۹ | ۰     | ۰/۱۹۳ | ۰/۰۳۵ | ۰/۰۱۴ | ۰/۰۷۵ | ۰/۶۳۷ | ۰/۰۰۸ |       | ۲         |
| ۰/۰۹۵ | ۰/۰۸۱ | ۰     | ۰/۰۵۶ | ۰/۲۸۷ | ۰/۱۳۳ | ۰/۱۵۵ | ۰/۰۵۲ | ۰/۲۳۶ |       | ۳         |
| ۰/۱۴۳ | ۰/۳۶۹ | ۰/۰۱۹ | ۰/۰۱۹ | ۰/۱۹۳ | ۰/۲۷۹ | ۰/۰۴۶ | ۰/۰۲۲ | ۰/۰۵۳ |       | ۴         |
| ۰     | ۰/۱۳۷ | ۰/۰۰۴ | ۰/۱۲۷ | ۰/۲۸۸ | ۰/۰۴۴ | ۰/۰۷۴ | ۰/۳۱۳ | ۰/۰۱۴ |       | ۵         |
| ۰/۱۴۳ | ۰/۰۵۸ | ۰     | ۰/۲۵۶ | ۰/۰۹۱ | ۰/۰۷۶ | ۰/۲۶۸ | ۰/۱۵  | ۰/۱۰۲ |       | ۶         |
| ۰/۰۹۵ | ۰     | ۰/۰۹۶ | ۰     | ۰/۰۰۷ | ۰/۰۱۳ | ۰     | ۰     | ۰/۰۰۲ |       | ۷         |
| ۰/۰۴۸ | ۰/۳۲۸ | ۰     | ۰/۰۴۹ | ۰/۰۴۳ | ۰/۰۷۲ | ۰/۰۴۹ | ۰/۴۲۷ | ۰/۰۳۱ |       | ۸         |

همان احتمالات گسیل یک مشاهده از حالت خاص جوی می‌باشد که بر اساس الگوریتم EM برآورد شده‌اند. با داشتن این الگوهای احتمالی، می‌توان به احتمال بارش در هر ایستگاه تحت یک حالت جوی دست یافت. هر چه قطر دوایر بیشتر باشد، احتمال مذکور بزرگتر است. هر الگوی مستخرج از مدل مارکف پنهان، توزیعی خاص از وقوع بارش را در گستره ایران به همراه دارد. به‌عنوان مثال، الگوهای جوی ۲ و ۷ به ترتیب، بیانگر شرایط خشکسالی (پایداری) و ترسالی (ناپایداری) در اکثر قریب به اتفاق ایستگاه‌ها می‌باشند. یا الگوهای ۳ و ۶ حاکی از وجود شرایط ناپایداری تنها در منطقه شمال غرب ایران است.

به‌طوریکه انتظار می‌رود، حالت جوی ۲، تیپ آب و هوایی غالب طی زمستان باشد. این حالت جلوتر به‌عنوان حالت خشک (الگوی هوای پایدار در کل کشور) تفسیر می‌شود. هرکدام از حالات پنهان جوی بیانگر شرایطی از جو هستند که منجر به وقوع بارش در ایستگاه‌های مورد مطالعه با احتمالات مختلف می‌شوند. این موضوع بیانگر ناهمگنی الگوهای بارش جوی طی زمستان در ایران است. بدین معنا که بارش زمستانه از الگوی واحدی تبعیت نمی‌کند. در شکل ۴، هشت حالت جوی شامل الگوهای فضایی احتمال وقوع بارش آورده شده است. دوایر موجود برای هر ایستگاه، میزان احتمال وقوع بارش به ازای حالتی معین از جو را نشان می‌دهند. این احتمال در حقیقت



شکل ۴- الگوهای توزیع فضایی احتمالات وقوع بارش در گستره ایران مبتنی بر مدل مارکف پنهان

گردد. این عوامل تحت تاثیر الگوهای سینوپتیکی در مقیاس وسیعی است که این الگوها می توانند منشا و اثرات متفاوتی در یک منطقه داشته باشند. حجازی زاده و فتاحی (۶)، ۸ الگوی مختلف سینوپتیکی که دارای منشا و تاثیر متفاوت بر ایران هستند، برای بارش های زمستانه بر اساس

تفسیر فیزیکی هر کدام از این الگوها، محققین را در درک بهتر الگوهای سینوپتیکی کمک می کند، به طوریکه انتظار می رود هر کدام از این الگوها منطبق بر یک الگوی خاص سینوپتیکی باشد. از لحاظ فیزیکی بارش زمانی اتفاق می افتد که عامل صعود و اشباع توده ای از هوا فراهم

ایران کشیده شده و در برخی الگوها، یک چرخند و یا واچرخند روی ایران مستقر و منجر به تولید اوضاع ترسالی و یا خشک‌سالی شده است. بنابراین در یک نگاه کلی از مدل مارکف پنهان در درک بهتر الگوهای توزیع فضایی بارش در یک منطقه می‌توان استفاده نمود. به عبارتی دیگر، تحت شرایطی که دسترسی به متغیرهای زیاد برای استخراج الگوهای سینوپتیکی وجود ندارد، می‌توان با مدل مارکف پنهان به این الگوها دست یافت و بهترین تعداد الگوها را با معیار BIC انتخاب نمود.

روش خوشه‌بندی آماری ارائه دادند. داده‌های مورد استفاده در کار آنها شامل ارتفاع تراز ۵۰۰ هکتوپاسکال، فشار سطح زمین و بارش ماهانه بود. تطبیق این الگوهای سینوپتیکی با الگوهای فضایی مستخرج از مدل مارکف پنهان در جدول ۳ آمده است. چنانچه مشخص است، به‌طور تقریبی هر الگوی پنهان جوی، نماینده یک نوع الگوی سینوپتیکی است که بر اثر برخورد توده‌های هوای پرفشار و کم‌فشار منجر به وقوع و یا عدم وقوع بارش می‌شوند. منشا این پرفشارها و کم‌فشارها در عرض‌های بالاتر و مدارهای شرقی‌تر است که زبانه‌های آنها تا درون

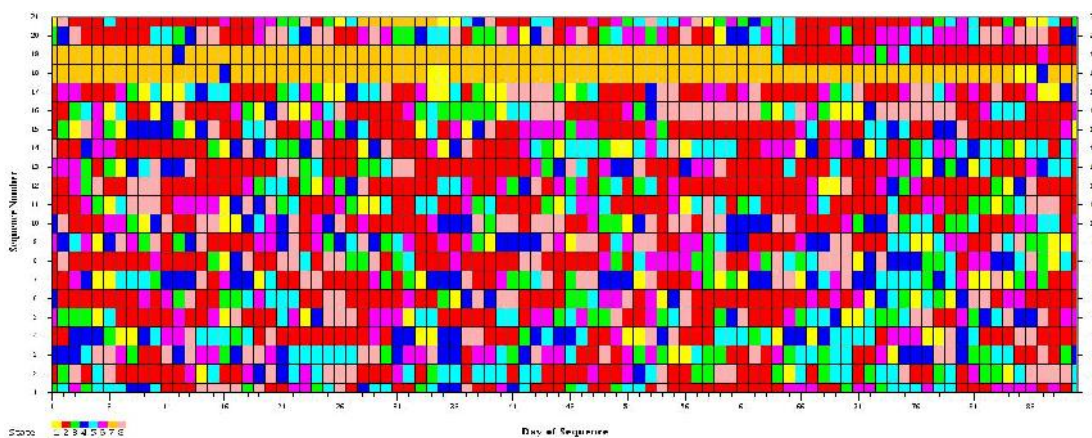
جدول ۳- تطبیق الگوهای مستخرج از مدل مارکف پنهان و الگوهای سینوپتیکی بارش زمستانه در ایران (۶)

| الگوی توزیع فضایی احتمال وقوع بارش، مبتنی بر HMM | منشا سینوپتیکی               | توضیحات  |
|--|------------------------------|--|
| ۱  | کم‌فشار اروپای شمالی         | در این الگو زبانه کم‌فشار مدیترانه ضعیف شده است و تنها بخش شمال غرب ایران از هوای ناپایدار و بارانی برخوردار است.  |
| ۲  | پرفشار مرکزی                 | در این الگو، مرکز پرفشاری در شمال غربی ایران بسته شده است و زبانه‌های این پرفشار در سراسر ایران گسترش دارد. در زمان فعالیت این تیپ آب‌وهوایی در سراسر ایران پایداری حاکم بوده و شرایط خشک و سرد غالب است.  |
| ۳  | پرفشار سیبری                 | در سطح زمین پرفشاری روی سیبری قرار گرفته که زبانه‌های آن تا درون ایران کشیده شده است. این الگو هوای نسبتاً خشک را به‌خصوص برای بخش‌های شمال‌شرقی و مرکزی ایران به‌همراه دارد.  |
| ۴  | پرفشار اروپای شرقی           | در این الگو، ناوهای روی اروپای شرقی و دریای سیاه مستقر است و جریان‌ها پس از عبور از اروپای شرقی، جنوب دریای سیاه، جمهوری آذربایجان و دریای خزر، هوای مرطوب را وارد ایران می‌کند. این الگو سهم عمده را در بارش‌های زمستانه بخش‌های شرقی، شمالی، مرکزی و جنوبی ایران دارد.   |
| ۵  | پرفشار اروپای شمالی          | پرفشاری بر روی اروپای شمالی می‌گیرد و زبانه‌های آن تا شمال ایران کشیده می‌شود. این الگو سهم قابل توجهی در بارش‌های زمستانه نوار ساحلی دریای خزر و بخش‌های شمال غربی دارد.  |
| ۶  | پرفشار شرقی                  | در این الگو، پرفشاری در شرق ایران قرار گرفته است که باعث خشکی هوا به‌خصوص در مناطق شرقی و جنوبی می‌گردد. این شرایط به‌طور نسبی غرب ایران را هم تحت تأثیر قرار می‌دهد.  |
| ۷  | کم‌فشار مدیترانه‌ای          | زبانه کم‌فشار مدیترانه‌ای بسیار قوی بوده و بخش‌های قابل توجه ایران را ناپایدار می‌سازد. کم‌فشار سودان نیز بخش‌های جنوبی را تحت تأثیر قرار داده و شرایط بسیار مساعد برای ریزش‌های جوی در سراسر ایران فراهم می‌شود.  |
| ۸  | تلفیق پرفشار سیبری و اروپایی | با تلفیق دو پرفشار سیبری و اروپایی، بخش‌های شرقی تحت تأثیر پرفشار سیبری و بخش‌ها غربی تحت تأثیر پرفشار اروپایی است. با کشیده شدن زبانه‌های این پرفشار تا دریای عمان و خلیج فارس و فرارفتن هوای مرطوب به سمت درون ایران، بارش‌هایی برای جنوب و تا حدودی بخش‌های جنوب شرقی و خشکی برای سایر نقاط ایران ایجاد می‌شود. |

### توزیع زمانی الگوهای بارندگی

در استفاده از مدل مارکف پنهان علاوه بر دستیابی به توزیع فضایی الگوها، می‌توان به توزیع زمانی آنها یا سری زمانی الگوهای سینوپتیکی نیز دست یافت. دستیابی به توزیع زمانی این الگوها با استفاده از الگوریتم ویتربی مدل مارکف پنهان محقق می‌شود. چنانچه ذکر شد، الگوریتم ویتربی محتمل‌ترین توالی پنهان به‌ازای یک توالی قابل رویت را برآورد می‌کند. توالی پنهان در این مطالعه سری زمانی حالات پنهان جوی یا الگوهای توزیع زمانی احتمالات وقوع بارش است. در شکل ۵، این توالی برای ۹۰ روز زمستان طی ۲۱ سال مشاهده‌شده به‌دست‌آمده با الگوریتم ویتربی آورده شده است. محور افقی شماره روز از ابتدای فصل زمستان و محور عمودی، شماره سال یا توالی است. هر رنگ نماینده یک حالت پنهان جوی است. بیشترین فراوانی بین الگوهای جوی متعلق به حالت ۲ یعنی پرفشار مرکزی (رنگ قرمز) است که تمام

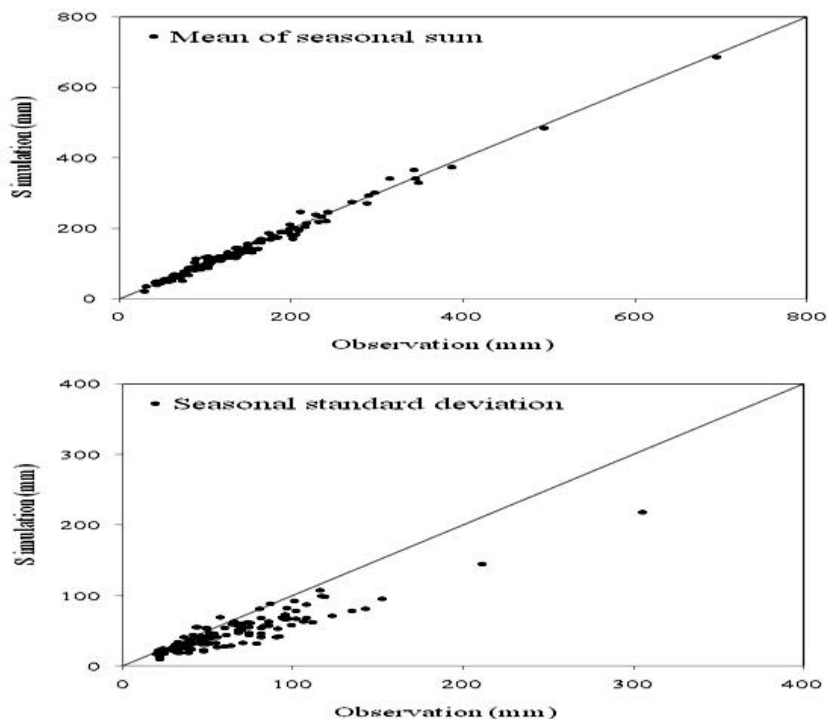
ایستگاه‌های مورد مطالعه تحت این الگو، از یک آب‌وهوای سرد و خشک برخوردار هستند. به‌عبارتی می‌توان گفت که شرایط جوی غالب طی زمستان در گستره ایران، شرایط خشک با احتمال وقوع ریزش‌های جوی اندک است. احتمال اولیه وقوع حالت ۲ نیز طبق جدول ۲، معادل ۰/۴۲۹ است که بیشترین مقدار احتمالات اولیه را به خود اختصاص داده است. با این وجود طبق شکل ۵، طی سال‌های ۲۰۰۷ و ۲۰۰۸ (شماره سال‌های ۱۸ و ۱۹ در محور قائم) الگوی حاکم بر جو ایران، الگوی ۷ (کم‌فشار مدیترانه‌ای) است. براساس این الگو، طی این سال‌ها آب و هوای کشور از شرایط بارانی در تمام ایستگاه‌ها برخوردار است که بارندگی‌های این سال‌ها گواه این مطلب می‌باشد. یکی از کاربردهای دسترسی به توالی زمانی این الگوها، رصد تغییرات فصلی اقلیم یک منطقه است که این تغییرات خود را به شکل تغییر در الگوی غالب حاکم بر آن منطقه نشان می‌دهد.



شکل ۵- توالی حالات پنهان (الگوهای توزیع فضایی احتمال وقوع بارش جوی) مبتنی بر الگوریتم ویتربی

معیار مجموع بارش فصلی در هر ایستگاه، مقادیر صدک‌های جامعه آماری بارش هر ایستگاه، فراوانی طول دوره‌های تر و خشک متوسط ایران و وابستگی‌های فضایی ایستگاه‌ها. در شکل ۶، نمودار ۱:۱ میانگین و انحراف معیار مجموع مقادیر بارش فصلی برای ۱۳۰ ایستگاه مشاهده می‌شود. مدل در شبیه‌سازی میانگین‌ها عملکرد موفق‌تری دارد. اما در شبیه‌سازی انحرافات از معیار، دچار کم‌برآوردی است و تا حدودی قادر به شبیه‌سازی مقادیر بزرگ انحراف معیار نمی‌باشد.

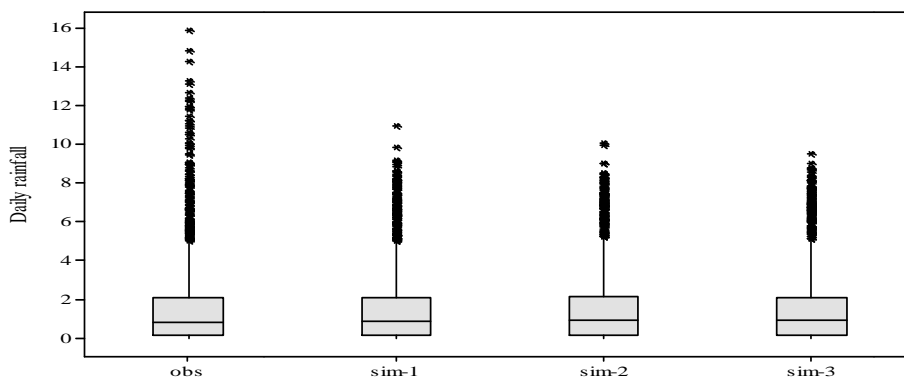
**شبیه‌سازی بارندگی**  
شبیه‌سازی استوکاستیک سری‌های زمانی بارندگی به صورت چندایستگاهی نیز از اهداف دیگر مطالعه حاضر است. یک مدل شبیه‌ساز توانمند در شبیه‌سازی ویژگی‌های آماری و احتمالی و پارامترهای سری‌های مشاهده شده موفق است. برای بررسی کارکرد مدل مارکف پنهان مورد مطالعه در شبیه‌سازی استوکاستیک بارندگی در مقیاس روزانه، ۵۰ سال داده تولید شده است. پارامترهای مورد ارزیابی عبارتند از: میانگین و انحراف



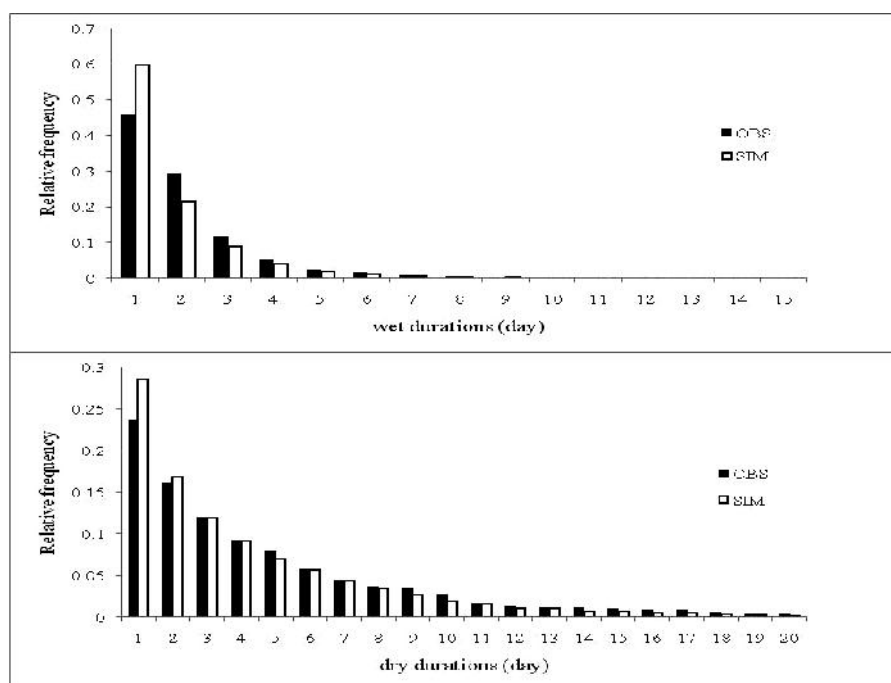
شکل ۶- نمودارهای ۱:۱ میانگین‌ها و انحراف معیارهای فصلی متعلق به ۱۳۰ ایستگاه در گستره ایران

بسیار کمتر از ۵ درصد کل داده‌هاست. همچنین برای بررسی کارکرد مدل در شبیه‌سازی طول دوره‌های تر و خشک، از نمودار میله‌ای فراوانی‌های تجربی تداوم‌های مختلف خشک و تر استفاده شده است. این فراوانی‌ها که برای سری زمانی بارندگی متوسط ایران محاسبه شده است، برای دو سری مشاهده شده و شبیه‌سازی شده در شکل ۸ آورده شده است. چنانچه مشخص است مدل مارکف پنهان در شبیه‌سازی تداوم‌های خشک (تا ۲۰ روز) و تر (تا ۱۵ روز) دارای یک موفقیت نسبی است که این موفقیت برای تداوم‌های بزرگتر به‌خصوص بیشتر است. مقایسه توزیع احتمالاتی تجربی داده‌های بارندگی می‌تواند از دیگر معیارهای ارزیابی باشد.

این مسئله خود را در نمودار جعبه‌ای شکل ۷ نشان داده است. این جعبه‌ها که متعلق به سری زمانی بارش متوسط (روی ۱۳۰ ایستگاه) ایران است، حدود اطمینان ۰/۹۵ داده‌ها را نشان می‌دهد. جعبه اول متعلق به داده‌های مشاهده شده و سه جعبه بعدی حاصل سه بار شبیه‌سازی است. مدل در ساخت حدود اطمینان ۰/۹۵ برای متوسط ایران موفق است ولی چنانچه مشخص است در ساخت مقادیر بارش حدی یا داده‌های پرت در هر سه بار شبیه‌سازی موفق نمی‌باشد. علت اصلی این موضوع را می‌توان در نمودار مربوط به انحراف معیار شکل ۶ یافت. کم‌برآوردی در ساخت واریانس‌های فصلی باعث می‌شود که مدل قادر نباشد تا داده‌های حدی را در یک سری ۲۱ ساله شبیه‌سازی کند. البته فراوانی این اطلاعات



شکل ۷- نمودار جعبه‌ای توزیع داده‌های بارندگی در حدود اعتماد ۹۵ درصد، متوسط ایران



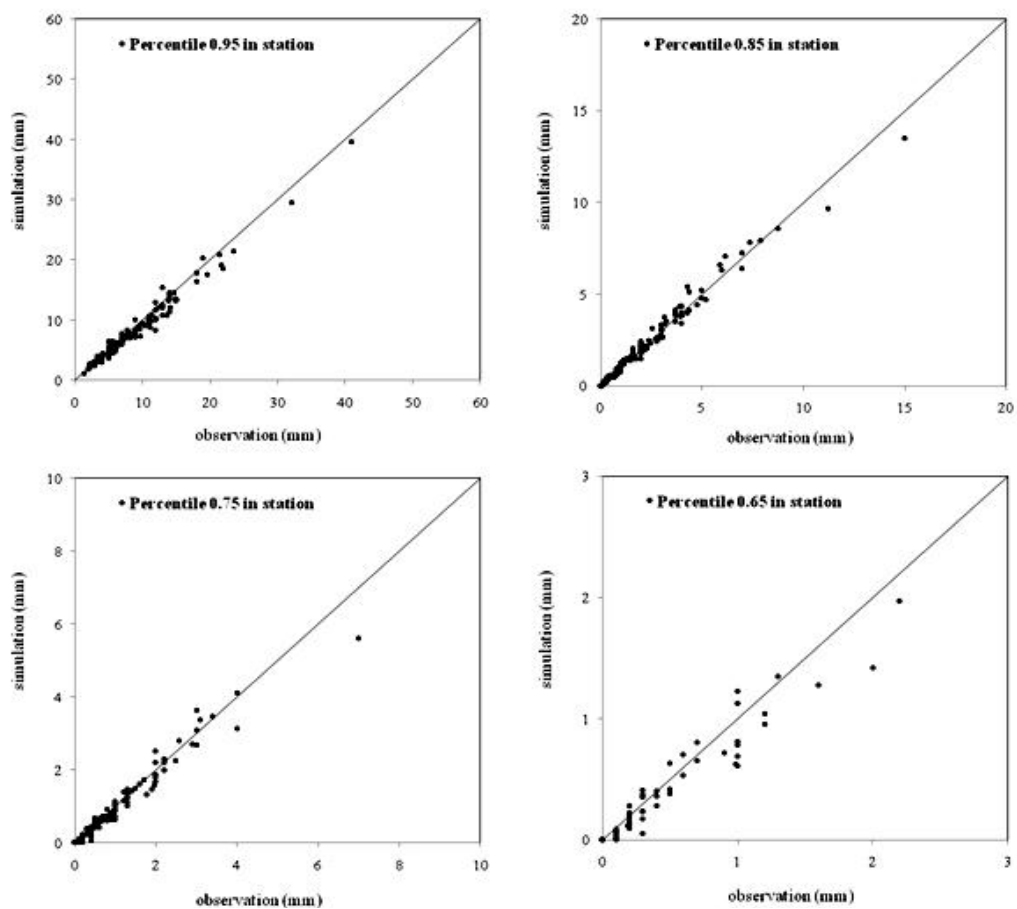
شکل ۸- نمودارهای میله‌ای فراوانی‌های تجربی تداوم‌های خشک و تر برای سری زمانی متوسط ایران

سه نقطه‌نظر مورد مطالعه واقع شده است: انتخاب تعداد الگوهای مناسب جوی به عنوان حالات پنهان برای دستیابی به الگوهای توزیع فضایی بارش، رمزگشایی توالی زمانی این الگوها و شبیه‌سازی استوکاستیک بارش روزانه برای فصل زمستان. با توجه به مشابهت تعداد الگوهای مستخرج از مدل مارکف پنهان با الگوهای سینوپتیکی بارش زمستانه، می‌توان کارکرد این رهیافت را در استخراج الگوهای جوی موفق ارزیابی کرد. مطالعه توالی زمانی این الگوها طی سال‌های آماری می‌تواند تغییر یا ثبات روند یا الگوهای بارشی را نشان دهد که در مطالعات بررسی تغییرات اقلیمی گذشته می‌تواند مفید باشد. با توجه به اینکه یک الگوی بارش طی فصل زمستان، الگوی پایداری حاکم بر کشور و این الگو معرف شرایط خشکی در سرتاسر کشور است، می‌توان از این روش برای پیش‌خشکسالی و برآورد دوره بازگشت این پدیده در تداوم‌های مختلف استفاده نمود. همچنین با توجه به قابلیت شبیه‌سازی استوکاستیک این رهیافت، مدل مارکف پنهان می‌تواند به‌عنوان یک مولد هواشناسی چندمکانی (Multisite Weather Generator) نیز شناخته شود که تولید همزمان بیش از یک ایستگاه منجر به صرفه‌جویی در هزینه و زمان می‌شود.

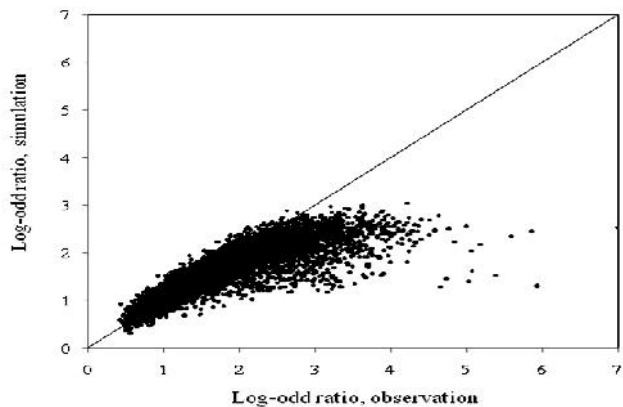
یکی از فرضیات اساسی در انواع مدل‌های مارکف پنهان و مدل مورد بحث در تحقیق حاضر، فرض ایستایی با زمان است. این فرض به‌طور دقیق‌تر اشاره به ثبات احتمال انتقال در طول زمان دارد. به‌طوریکه احتمال انتقال از حالتی به حالت دیگر با زمان تغییر نکند.

از آنجا که مطالعه به‌صورت فصلی انجام می‌شود، مجموعه داده‌های ۲۱ سال، به لحاظ آماری همگن محسوب می‌شوند. چرا که اگر داده‌های روزانه برای کل سال مد نظر بودند، ارزیابی‌ها باید به‌صورت فصلی یا ماهانه انجام می‌شد. بر این اساس با مرتب‌سازی داده‌های هر ایستگاه، مقادیر صدک‌های ۰/۶۵، ۰/۷۵، ۰/۸۵ و ۰/۹۵ برای هر ایستگاه محاسبه شدند. هر صدک معرف فراوانی داده‌های کمتر از خود است. به‌عنوان مثال صدک ۰/۷۵ معادل مقداری در یک سری اعداد است که ۷۵ درصد داده‌ها از آن کمتر یا مساوی با آن هستند. بنابراین ارزیابی این پارامترها می‌تواند در درک کارکرد مدل در شبیه‌سازی توزیع فراوانی تجربی مفید باشد. در نمودارهای شکل ۹، تطابق چشمی این صدک‌ها برای هر ۱۳۰ ایستگاه نشان‌دهنده شده است. این نمودارها حکایت از کارکرد مناسب رهیافت مارکف پنهان در شبیه‌سازی توزیع احتمال تجربی داده‌ها دارد. مقدار نسبت لگاریتمی بخت‌ها برای هر جفت ایستگاه در مطالعه حاضر محاسبه شد و تطابق چشمی نسبت‌های مشاهده‌شده و شبیه‌سازی شده در شکل ۱۰ آورده شده است. در یک نگاه کلی کارکرد مدل در شبیه‌سازی وابستگی‌های فضایی مناسب ارزیابی می‌شود، ولی در ساخت نسبت‌های بزرگتر از ۳ دچار کم‌برآوردی است که کارکرد آن‌را با محدودیت مواجه می‌سازد. مشابه این نتیجه توسط مهرترا و همکاران (۱۰) دریافت شده است.

در تحقیق حاضر عملکرد مدل مارکف پنهان در شبیه‌سازی فضایی و زمانی بارش روزانه فصل زمستان از



شکل ۹- تطابق چشمی صدک‌های دو سری شبیه‌سازی شده و مشاهده شده برای ۱۳۰ ایستگاه



شکل ۱۰- تطابق چشمی نسبت لگاریتمی بخت‌ها بین هر جفت ایستگاه

حالت تنها به حالت پیشین خود)، یک وابستگی متقاطع حالات پنهان به عواملی دیگر را نیز در نظر گرفت. این عوامل می‌توانند متغیرهای خروجی مدل‌های بزرگ مقیاس باشد که چنین نگرشی منجر به توسعه مدل‌های ناهمگن می‌شود. با این نگاه می‌توان، عدم قطعیت موجود در سناریوهای تغییر اقلیم را به خوبی در توالی‌های زمانی الگوهای جوی و در روند شبیه‌سازی تصادفی بررسی نمود و سری‌هایی تصادفی مبتنی بر این سناریوها برای چشم‌انداز آتی شبیه‌سازی نمود.

این نگرش منجر به توسعه مدل‌های همگن می‌شود. این در حالی است که در واقعیت ممکن است، این احتمالات تحت تاثیر عوامل خارجی باشد. به عنوان مثال ممکن است طی دوره‌ای احتمال انتقال به حالت الگوی جوی خشک بیشتر شود، یا به عبارتی تمایل آب‌وهوا به خشکی بیشتر شود که در اینجا به‌طور قطع ممکن است درایه‌های ماتریس احتمال انتقال تغییرات معنی‌داری داشته باشند. برای وارد کردن چنین تغییراتی در روند مدل‌سازی، باید در کنار فرض مارکفی مدل (وابستگی هر

### منابع

1. Ailliot, P., C. Thompson and P. Thompson. 2008. Space time modelling of precipitation using a hidden markov model and censored Gaussian distributions. *Journal of Hydrology*, 420: 27-51.
2. Bardossy, A. and E.J. Plate. 1992. Space-time model for daily rainfall using atmospheric circulation patterns. *Water Resources Research*, 28: 1247-1259.
3. Bellone, E., J.P. Hughes and P. Guttorp. 2000. A hidden markov model for downscaling synoptic atmospheric patterns to precipitation amounts. *Journal of Climate Research*, 15: 1-12.
4. Bogardi, I., I. Matyasovsky, A. Bardossy and L. Duckstein. 1993. Application of a space-time stochastic model for daily precipitation using atmospheric circulation patterns. *Journal of Geophysical Research*, 98: 16653-16667.
5. Ejlali, N. and H. Pezeshk. 2008. A Bidirectional Hidden Markov Model in Linear Memory. *Journal of Statistical Sciences*, 2: 131-148. (In Persian)
6. Hejazizadeh, Z. and A. Fattahi. 2004. Synoptical analysis for winter rainfall in Iran. *Journal of Geography*, 2: 89-107.
7. Hocaoglu, F.O. 2011. Stochastic approach for daily solar radiation modeling. *Journal of Solar Energy*, 85: 278-287.
8. Hughes, J.P. and P. Guttorp. 1994. A class of stochastic models for relating synoptic atmospheric patterns to local hydrologic phenomenon. *Water Resources Research*, 30: 1535-1546.
9. Hughes, J.P., P. Guttorp and S.P. Charles. 1999. A non-homogeneous hidden Markov model for precipitation occurrence. *Applied Statistics*, 48: 15-30.
10. Mehrotra, R., R. Srikanthan and A. Sharma. 2006. A comparison of three stochastic multi-site precipitation occurrence generators. *Journal of Hydrology*, 331: 280- 292.
11. Oliver, C. Ibe. 2009. Markov process for stochastic modeling. University of Massachusetts, Academic press, 505 pp.
12. Rabiner, L.R. and B.H. Juang. 1986. An introduction to hidden markov models. *IEEE Trans. Acoust. Speech Signal Process.* 3: 4-16.
13. Srikanthan, R., M.A. Thyer, G.A. Kuczera and T.A. McMahon. 2002. Application of hidden state markov model to Australian annual rainfall data. CRCC Working Document. (In Persian)
14. Thompson C., P. Thomson and X. Zheng. 2007. Fitting a multisite daily rainfall model to New Zealand data. *Journal of Hydrology*, 340: 25-39.
15. Wilks, D.S. 1999. Multi-site downscaling of daily precipitation with a stochastic weather generator. *Journal of Climate Research*, 11: 125-136.
16. Zucchini, W. and P. Guttorp 1991. A hidden markov model for space-time precipitation. *Water Resources Research*, 27: 1917-1923.

## Spatial-Temporal Modeling of Occurrence and Amount of Winter Rainfall using Hidden Markov Model

Mehdi Ghamghami<sup>1</sup>, Nozar Ghahreman<sup>2</sup> and Javad Bazrafshan<sup>3</sup>

1 and 3- PhD. Student and Assistant Professor, University of Tehran

2- Associate Professor, University of Tehran (Corresponding author: nghahreman@ut.ac.ir)

Received: July 16, 2013 Accepted: December 2, 2013

### Abstract

Multi site modeling of rainfall is one of the most important issues in environmental sciences especially in watershed management. For this purpose, different statistical models have been developed which involve spatial approaches in simulation and modeling of daily rainfall values. The hidden Markov is one of the multi-site daily rainfall models which in addition to simulation of daily rainfall values, explores the spatial and temporal pattern of rainfall events. In this study, the winter (January to April) rainfall pattern of 130 rain gauges have been modeled using hidden Markov approach during a 21 years period (1990-2010). The aim of this study was finding temporal and spatial distribution of weather patterns and stochastic simulation of occurrence and amount of rainfall, simultaneously. To achieve this goal, different hidden Markov algorithms including, *Viterbi* decoding algorithm, *Expectation-Maximization (EM)* algorithm and a stochastic simulation approach with the *probability transformation* were applied. It is expected that extracted patterns, using hidden Markov model, are consistent with synoptic patterns and accordingly eight different weather pattern as the definite set of possible cases were recognized. The most frequent rainfall pattern extracted from hidden Markov model was the dry pattern (stable condition) in which the rainfall occurrence probability is low in most of the stations. This pattern has the maximum initial probability of 0.429 and maximum Markov transfer probability of 0.637 Besides, multi-site simulation of winter rainfall keeping the basic statistic of mean, standard deviation of total seasonal rainfall and percentile values in each station and also spatial correlation of occurrence or non-occurrence of rainfall produced reasonable result. In general this approach can be recommended for regional studies.

**Keywords:** Hidden Markov Model, Winter rainfall, Spatial patterns, Simulation