

تعیین متغیرهای زمینه‌ای در طبقه‌بندی اقلیمی ایران: معرفی و کاربرد روش تحلیل عاملی و تجزیه مؤلفه‌های اصلی در تحلیل مطالعات جغرافیایی و اقلیم‌شناسی

چکیده

روشهای آماری چند متغیره، کاربردهای بی‌شماری در مطالعات جغرافیایی و به‌خصوص اقلیم‌شناسی دارند. طبقه‌بندی اقلیمی ایران که در سال ۱۳۸۰ توسط مؤلفان این مقاله با استفاده از روش تحلیل خوشه‌ای انجام گردید، یکی از این کاربردها می‌باشد. از طرفی با توجه به تنوع زیاد متغیرهای مورد استفاده برای طبقه‌بندی مذکور، انجام مطالعه حاضر جهت معرفی متغیرهای اصلی که در مرزبندی نواحی اقلیمی نقش داشته‌اند، و تعیین میزان تأثیر هر کدام از آنها در این طبقه‌بندی ضرورت پیدا کرد. برای این منظور داده‌های ماهانه ۴۱ ایستگاه هواشناسی سینوپتیک (۱۹۹۳-۱۹۹۵) مورد استفاده قرار گرفت. روش تحلیل عاملی و تجزیه مؤلفه‌های اصلی که از کاربردی‌ترین روشهای چند متغیره آماری می‌باشند برای تجزیه و تحلیل داده‌ها انتخاب شد.

نتایج نشان می‌دهند که مهمترین مؤلفه نم نسبی می‌باشد که بیش از ۴۰٪ از کل واریانس را به خود اختصاص داده است ($Pct\ of\ Var = 40/90$). دومین مؤلفه اصلی مربوط به دماهای حداقل و حداکثر است ($Pct\ of\ Var = 23/37$) که بیش از ۲۳٪ از کل تغییرات را تبیین نموده‌اند. سایر متغیرها در مابقی مؤلفه‌ها قرار دارند. یافته‌ها گویای آن است که طبقه‌بندی انجام شده دارای مبنای اصولی بوده و بر اساس قوانین علمی استوار گردیده است به‌نحوی که مرزهای نواحی اقلیمی به‌طور متعادل از کلیه متغیرها تأثیر پذیرفته‌اند. این مطالعه همچنین کاربرد روش تحلیل عاملی را به‌عنوان روشی کارآمد، برای انجام مطالعات تحلیلی و مفصلی که نیاز به کنکاش در خصوص روابط بین

متغیرها دارند، نشان می‌دهد. در این مطالعه همچنین سعی شده است تا چگونگی بارگذاری متغیرها در مؤلفه‌های اصلی و نحوه محاسبه مقادیرهای ویژه و بردارهای ویژه بصورت ساده شرح داده شود. کلید واژه‌ها: متغیر زمینه‌ای، تجزیه مؤلفه‌های اصلی، تحلیل عاملی، مقدار ویژه، بردار ویژه.

مقدمه

امروزه کاربرد روشهای کمی در قلمرو تحلیل موضوعات و مسائل جغرافیایی بنا به ضرورت بسیار رایج گردیده است. مروری بر مطالعات انجام شده در چند سال اخیر در لابلای منابع جغرافیایی و بخصوص اقلیم‌شناسی نشان می‌دهد که این روشها به‌طور وسیعی مورد استفاده قرار گرفته و نتایج به‌دست آمده از آنها ارزش و کاربرد مطالعات جغرافیایی را بیش از پیش نمایان ساخته‌اند. از سوی دیگر با توجه به اینکه موضوعات و مسائل جغرافیایی دارای ابعاد وسیعی بوده و تحقیق پیرامون یک موضوع جغرافیایی متضمن در نظر گرفتن متغیرها و عوامل متعددی است، امروزه در بسیاری از شاخه‌های تخصصی علم جغرافیا از روش‌های آماری چند متغیره استفاده می‌شود.

طبق تعریف فرهنگ جدید جغرافیایی روش تجزیه مؤلفه‌های اصلی (PCA)^۱ و روش تحلیل عاملی (FA)^۲ دو نمونه از مهمترین روشهای آماری چند متغیره محسوب می‌شوند. بسیاری از روشهای آماری مشروح در کتب درسی آمار مقدماتی شامل روشهای یک متغیره یا دو متغیره می‌باشند زیرا فقط با تجزیه و تحلیل یک یا دو متغیر تصادفی سروکار دارند. اما در روشهای آماری چند متغیره پژوهشگر به‌طور همزمان چندین متغیر را برای تحلیل یک موضوع به کار می‌گیرد و گاه تعداد متغیرهای به کار رفته به بیش از چند صد نیز می‌رسند (Small and Witherick, 1990).

ویلموت^۳ (۱۹۷۸) در مقاله‌ای با عنوان "گروه‌بندی مناطق بارش کالیفرنیا با روش PCA"، با استفاده از داده‌های مجموع بارش ماهانه ۹۰ ایستگاه هواشناسی، کاربرد این روش را عملاً نشان داده است. به نظر نامبرده روش مؤلفه‌های اصلی روش عملی چندان مناسبی نیست زیرا با این روش نمی‌توان بیان نمود که از لحاظ اقلیم‌شناسی به چه دلیل برخی ایستگاهها در مؤلفه اول یا به عبارت دیگر در ناحیه بارشی اول قرار گرفته‌اند. اما از سوی دیگر اذعان می‌دارد اگرچه ثابت کردن دلایل در این خصوص به لحاظ پیچیدگی ذاتی فرآیندهای اقلیمی قدری مشکل است اما نواحی که در هر یک از مؤلفه‌ها قرار می‌گیرند، دارای اهمیت می‌باشند.

از نمونه‌های دیگری که در خصوص کاربرد روشهای مذکور می‌توان ذکر نمود مطالعه‌ای است که ویتون^۴ (۱۹۸۸) در مورد تحلیل اقلیم‌شناسی سینوپتیکی تغییرپذیری بارش در جنوب شرقی استرالیا با روش PCA انجام داده است. همچنین شاو و ویلر^۵ (۱۹۸۵) تحلیلی مورفومتریک بر روی مشخصات هیدرولوژیکی ۳۰ حوضه رودخانه‌ای در کشور انگلستان انجام داده‌اند. در این تحلیل مشخصات مختلف حوضه‌ها در قالب ۱۳ متغیر، مورد مطالعه قرار گرفته است. پس از انجام تحلیل عاملی این مشخصات در قالب سه فاکتور اصلی شامل شاخص خاک، خشکی دره و فراوانی رودخانه‌ها در حوضه خلاصه شده‌اند.

از دیگر تحقیقات انجام شده با روشهای فوق می‌توان به مطالعات مستاز- نونز^۶ (۲۰۰۰)، کلی و جونز^۷ (۱۹۹۹)، استبان- پارا و همکاران^۸ (۱۹۹۵)، کرایپالانی و همکاران^۹ (۱۹۹۸)، درودووسکی^{۱۰} (۱۹۹۳) و اوگالو^{۱۱} (۱۹۸۹) اشاره نمود. در مطالعات سنجش از راه دور نیز داده‌های خام سنجنده‌ها، با کمک روش تجزیه مؤلفه‌های اصلی قابلیت تفسیر پیدا نموده و امکان بررسی تصاویر ماهواره‌ای را فراهم می‌کنند. زیرا تصاویر ماهواره‌ای در حقیقت تصاویر رقومی بوده و هر تصویر دارای حجم عظیمی از داده‌های خام می‌باشد و اساسا بدون استفاده از روشهای آماری امکان تحلیل و تفسیر آنها میسر نیست. به‌علاوه از این روشها می‌توان برای فشرده کردن محتوای اطلاعات چندین باند مختلف (مثلا هفت باند تصاویر سنجنده‌های TM) به دو یا سه تصویر خلاصه شده استفاده کرد (Jenson, 1986; Mather, 2001).

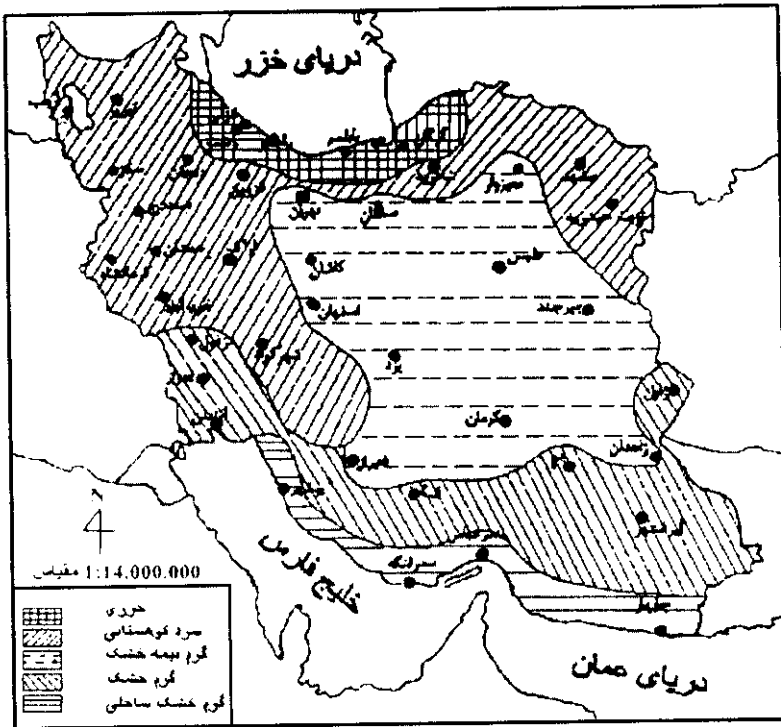
روش تحلیل خوشه‌ای نیز یکی دیگر از روشهای چند متغیره است که کاربردهای مشابهی با روشهای یاد شده به‌خصوص در جهت طبقه‌بندی‌هایی نظیر تعیین قلمروهای آب و هوایی یا رژیم‌های بارشی و دمایی دارد. طبقه‌بندی اقلیمی ایران با روش تحلیل خوشه‌ای توسط ترابی و همکاران، (۱۳۸۰) انجام شده است. طبقه‌بندی یاد شده نیز به‌خوبی منعکس کننده یکی از کاربردهای مؤثر این روشها می‌باشد. در مطالعه فوق از مهمترین متغیرهای اقلیمی (دما، بارش و نم نسبی) ۴۱ ایستگاه سینوپتیک کشور در یک دوره ۳۰ ساله آماری (۱۹۶۶ تا ۱۹۹۵) استفاده شده است. شکل ۱ موقعیت ایستگاههای منتخب و محدوده نواحی اقلیمی ایران را که بر اساس مطالعه مذکور تعیین گردیده‌اند نشان می‌دهد.

4. Whetton.
8. Esteban-parra et al.

5. Shaw & Wheeler.
9. Kripalan et al.

6. Mestas-Nunez.
10. Drosdowsky.

7. Kelly and Jones.
11. Ogalló. SID.ir



شکل ۱. موقعیت ایستگاهها و محدوده نواحی اقلیمی تعیین شده با روش تحلیل خوشه‌ای

یکی از پرسشهایی که در مورد طبقه‌بندی فوق مطرح گردیده این است که آیا متغیرهای به کار رفته دارای سهم یکسانی در طبقه‌بندی هستند؟ به عبارت دیگر کدام متغیرها مؤثرتر از سایرین بوده‌اند و چرا؟ اهمیت این پرسش در آن است که در پاسخ می‌توان متغیرهایی را که دارای بیشترین تأثیر در طبقه‌بندی بوده‌اند تعیین کرد.

طبق نظریات ترجونگ^{۱۲} (۱۹۹۶) و استرالر^{۱۳} (۱۹۹۴) مهمترین موضوع در بحث طبقه‌بندی اقلیمی، نوع متغیرها و عواملی است که با کمک آنها طبقه‌بندی انجام می‌شود. لذا به دنبال انجام طبقه‌بندی اقلیمی فوق، انجام تجزیه و تحلیل دیگری نیز در این زمینه ضرورت پیدا کرد تا بر اساس نتایج آن بتوان سهم هر یک از متغیرها را در طبقه‌بندی یاد شده نشان داد.

مطالعه حاضر در جهت رسیدن به هدف ذکر شده و با استفاده از روش تحلیل عاملی بر مبنای روش PCA، طراحی و انجام گردیده است. در ضمن مروری کوتاه بر مفاهیم روشهای مذکور و توضیح آماری آنها نیز صورت گرفته است.

مواد و داده‌ها

همانگونه که ذکر شد هدف از انجام مطالعه حاضر تعیین سهم هر کدام از متغیرهای بدکار رفته در طبقه‌بندی اقلیمی است که قبلاً با روش تحلیل خوشه‌ای انجام شده است. روش مؤلفه‌های اصلی چنانکه اشاره خواهد شد بر مبنای تعیین همبستگی بین متغیرها عمل می‌کند بدین معنا که همبستگی هر یک از متغیرها با تک تک متغیرها محاسبه می‌شود. لذا از لحاظ محاسباتی انتخاب نمونه‌ای از داده‌ها برای این منظور کفایت می‌کند (Kline Paul, 1994; Mather, 1976).

به همین لحاظ در این بررسی داده‌های ۳ ساله تعداد ۴۱ ایستگاه هواشناسی (۱۹۹۳-۱۹۹۵) مورد استفاده قرار گرفته است.

متغیرهای مورد بررسی در این تحقیق عبارتند از:

۱- میانگین حداقل دما (°C) ۲- میانگین حداکثر دما (°C)

۳- متوسط رطوبت نسبی در ساعت ۳ (%) ۴- متوسط رطوبت نسبی در ساعت ۹ (%)

۵- متوسط رطوبت نسبی در ساعت ۱۵ (%) ۶- مجموع بارش ماهانه (mm)

۷- بیشترین بارندگی روزانه (mm) ۸- تعداد روزهای با بارش مساوی یا بیشتر از ۱۰mm

لازم به ذکر است که داده‌ها به صورت ماهانه در نظر گرفته شده‌اند بنابراین ابتدا ماتریس اولیه‌ای به ابعاد ۴۱×۲۸۸ تهیه و تنظیم گردیده است (عدد ۴۱ نمایانگر تعداد ایستگاهها و عدد ۲۸۸ نشانگر تعداد متغیر یاد شده در ۱۲ ماه و در مدت ۳ سال می‌باشد).

الف- بررسی کیفیت داده‌ها: ایستگاههای مورد مطالعه از پوشش کاملی برخوردار بوده و لذا داده‌ها از نظر همگنی با روش آزمون توالی مورد بررسی قرار گرفته‌اند. در این آزمون با شمارش تعداد دنباله‌ها در هر یک از متغیرها و تعیین تعداد دنباله‌های مجاز، تصادفی بودن آنها مورد بررسی قرار می‌گیرد (Ebdon, 1977). در این مطالعه تصادفی بودن داده‌ها با احتمال خطای ۰/۰۵ مورد بررسی قرار گرفته و از همگن بودن آنها اطمینان حاصل شده است. احتمال همگونی داده‌ها و تصادفی بودن آنها با توجه به $P\text{-Value} < 0.05$ تأیید می‌شود.

ب- استاندارد نمودن داده‌ها: از آنجا که متغیرهای مختلف در این مطالعه دارای مقیاسهای اندازه‌گیری متفاوتی می‌باشند لذا بدست آوردن نمره Z و استاندارد نمودن آنها ضروری می‌نمود. بدین منظور مقدار میانگین و انحراف معیار متغیرها در هر یک از ماهها و برای هر یک از ایستگاهها تعیین گردیده و سپس استاندارد شده‌اند.

روش تجزیه و تحلیل داده‌ها

روش تحلیل عاملی اولین بار توسط چارلز اسپیرمن^{۱۴} در سال ۱۹۰۴ ابداع گردید. مدل پیشنهادی او که در واقع بر مبنای مدل خطی ساده استوار بود به خوبی می‌توانست روابط بین متغیرهای مختلف را توجیه کند (مقدم و همکاران، ۱۳۷۳).

روش تجزیه مؤلفه‌های اصلی برای اولین بار توسط کارل پیرسون در سال ۱۹۰۱ شرح داده شد. اما روشی که او ارائه نمود تنها برای محاسبه دو یا سه متغیر بود. روشهای عملی سالها بعد توسط هتلینگ در سال ۱۹۹۳ توضیح داده شد (Johnson & Dean, 1988).

نظر به اهمیت زیاد و کاربردهای وسیع روشهای مذکور در این قسمت به طور خلاصه به مهمترین کاربردهای آنها، که از لحاظ تکنیکهای آماری مهم می‌باشند، اشاره می‌شود:

۱- ساده کردن توصیف گروهی از متغیرهای مربوط به هم.
۲- جستجوی مقادیر حداکثر و حداقل با رسم نمودار هیستوگرام از تک تک مؤلفه‌های اصلی (Mather, 1976).

۳- تشخیص نرمال بودن حجم عظیمی از داده‌ها (Mulaik, 1972).

۴- محدود کردن ابعاد مسأله‌ای وسیع بدون از دست دادن متغیرها و داده‌های مهم.
۵- در مواردی که مشکل "همبستگی خطی چندگانه"^{۱۵} " میان متغیرهای مورد مطالعه وجود داد، تجزیه مؤلفه‌های اصلی امکان تحلیل رگرسیونی را فراهم می‌کند (جانسون و دین، ۱۹۸۸).
۶- پس از تعیین مؤلفه‌های اصلی متغیرهای زمینه‌ای به دست می‌آیند و این مؤلفه‌ها در حقیقت پیش‌فرض‌های جدیدی ایجاد می‌کنند که می‌توانند مکمل تحلیل‌های دیگر قرار گرفته و اساس مطالعات بعدی باشند (شاو و ویلر، ۱۹۸۵).

در تجزیه مؤلفه‌های اصلی کلیه متغیرها به طور مساوی در نظر گرفته می‌شوند. در این روش برخلاف روش رگرسیون چندگانه متغیرها به دو گروه متغیرهای مستقل و وابسته تقسیم نمی‌شوند بلکه هر مؤلفه اصلی مقداری از واریانس کل را تبیین می‌کند. بنابراین اولین مؤلفه اصلی حاوی بیشترین اطلاعات و دارای بالاترین واریانس است و آخرین مؤلفه دارای کمترین مقدار واریانس می‌باشد. در تعیین مهمترین مؤلفه‌های اصلی، تعیین مقدار ویژه و بردارهای ویژه حائز اهمیت هستند از این رو به شرح مختصری در مورد آنها می‌پردازیم:

- "مقادیر ویژه^{۱۶}" و "بردارهای ویژه^{۱۷}": مقدار ویژه به زبان ساده عبارت است از تعیین واریانس و انحراف معیار در ماتریس. از آنجا که اولین مرحله در تجزیه مؤلفه‌های اصلی ایجاد ماتریس همبستگی است، لذا گام بعدی پس از تشکیل ماتریس اولیه، تعیین مقادیر ویژه و بردارهای ویژه می‌باشد. محاسبات این بردارها از طریق محاسبات تکراری صورت می‌گیرد.

- مراحل محاسباتی تجزیه مؤلفه‌های اصلی: کلیه محاسبات با استفاده از نرم‌افزار SPSS انجام گردیده و چنانچه ذکر شد، اولین گام تهیه ماتریس داده‌های خام بوده است. با توجه به تعداد متغیرها (۲۸۸)، عملاً ارائه جدولی بزرگ از داده‌ها در این متن امکان‌پذیر نبود لذا با استفاده از چهار متغیر و همچنین مقادیر فرضی، محاسبات به صورت ساده شرح داده شده است. بدیهی است مراحل که توضیح داده خواهند شد قابل تعمیم به همه متغیرهای مورد مطالعه می‌باشند. جدول ۱ ماتریس اولیه‌ای را نشان می‌دهد که در نرم‌افزار SPSS جهت محاسبات تنظیم شده است.

جدول ۱ ماتریس داده‌های خام تنظیم شده

نام ایستگاه هوا شناسی	دمای حداقل ماه ژانویه (°C)	دمای حداقل ماه فوریه (°C)	دمای حداقل ماه مارس (°C)	دمای حداقل ماه آوریل (°C)
مشهد	۱۴	۹	۱۳	۵
تبریز	۱۹	۷	۱۹	۷
اصفهان	۱۲	۳	۲۰	۸
تهران	۱۱	۸	۲۵	۹

مرحله بعد پس از تعیین ماتریس اولیه، محاسبه همبستگی هریک از متغیرها با خودشان و با سایر متغیرها می‌باشد. این کار منجر به تولید ماتریس همبستگی می‌گردد. ضرایب همبستگی متغیرهای جدول ۱ با خودشان به شکل ماتریسی مقارن در می‌آید. جدول ۲ ماتریس همبستگی را نشان می‌دهد.

جدول ۲ ماتریس همبستگی

دمای حداقل ماه ژانویه (°C)	دمای حداقل ماه فوریه (°C)	دمای حداقل ماه مارس (°C)	دمای حداقل ماه آوریل (°C)	
۱	۰/۴	۰/۳	۰/۲	دمای حداقل ماه ژانویه (°C)
۰/۴	۱	۰/۲	۰/۱	دمای حداقل ماه فوریه (°C)
۰/۳	۰/۲	۱	۰/۳	دمای حداقل ماه مارس (°C)
۰/۲	۰/۱	۰/۳	۱	دمای حداقل ماه آوریل (°C)
۱/۹	۱/۷	۱/۸	۱/۶	مجموع ستونی ضرایب

تعیین مقادیر ویژه و برآوردهای ویژه دارای شش مرحله است که به ترتیب عبارتند از:
مرحله اول- تعیین مجموع ستونی ضرایب: در جدول ۲ مجموع ستونی ضرایب همبستگی محاسبه شده‌اند. اگر این مجموع ضرایب را با U_1 نشان دهیم آنگاه خواهیم داشت:

$$U_1 = 1/9, 1/7, 1/8, 1/6$$

مرحله دوم- نرمال کردن مقادیر بردار U_1 : در این مرحله مجذور هریک از مقادیر بردار U_1 محاسبه گردیده و سپس با هم جمع می‌شوند آنگاه جذر آن محاسبه می‌شود. بدین ترتیب خواهیم داشت:

$$(3/61 + 2/89 + 3/24 + 2/56) = 12/3$$

$$\sqrt{12/3} = 3/51$$

سپس برای نرمال کردن هریک از عناصر بردار U_1 بر عدد $3/51$ تقسیم می‌شود. مقادیر نرمال شده با بردار V_1 نشان داده می‌شود. این اولین بردار آزمایشی است:

$$V_1 = (0/54, 0/48, 0/51, 0/46)$$

مرحله سوم- در این مرحله عناصر بردار V_1 به ترتیب در اولین، دومین، سومین و چهارمین ستون ماتریس همبستگی (جدول ۲) ضرب شده و در آخر مجموع ستونی محاسبه می‌شود. بردار جدید با U_2 نشان داده می‌شود و نظیر مرحله قبل مقادیر این بردار نیز نرمال می‌شوند. بنابراین U_2 چنین خواهد بود:

$$U_2 = (0/97, 0/85, 0/90, 0/77)$$

سپس مجموع مجذور بردار U_2 تعیین گردیده و جذر آن گرفته می‌شود تا مقدار ویژه ماتریس به دست آید:

$$U_2 = 0/97^2 + 0/85^2 + 0/90^2 + 0/77^2 = 3/06$$

$$\sqrt{3/06} = 1/75$$

مرحله چهارم- هر یک از مقادیر U_2 بر عدد $1/75$ (مقدار ویژه ماتریس) تقسیم می‌شوند تا مقادیر بردار نرمال شوند. بردار حاصله (V_2) به شکل زیر خواهد بود:

$$V_2 = (0/55, 0/49, 0/51, 0/44)$$

بردار V_2 بردار ویژه نامیده می‌شود و $1/75$ مقدار ویژه این ماتریس می‌باشد.

مرحله پنجم- پس از تعیین مقادیر ماتریس همبستگی و بردار ویژه، ضرایب عامل تعیین می‌شوند. در این مرحله عناصر بردار ویژه یعنی V_2 را در جذر مقدار ویژه ضرب کرده و بدین ترتیب اولین مؤلفه اصلی را به دست می‌آوریم:

$$\sqrt{1/75} = 1/32 = \text{جذر مقدار ویژه} \quad 0/55 \times 1/32 = 0/726 = 0/74$$

$$0/49 \times 1/32 = 0/646 = 0/65 \quad 0/51 \times 1/32 = 0/673 = 0/67$$

$$0/44 \times 1/32 = 0/580 = 0/57$$

ارتباطی که بین ضرایب عامل و مقدار ویژه وجود دارد، آن است که مجموع مجذور ضرایب عامل با مقدار ویژه برابر است (جدول ۳).

جدول ۳ ضرایب اولین مؤلفه اصلی برای هریک از متغیرها

متغیرها	ضرایب عامل	مجذور مؤلفه‌ها
دمای حداقل ماه ژانویه (°C)	0/74	0/5476
دمای حداقل ماه فوریه (°C)	0/65	0/4225
دمای حداقل ماه مارس (°C)	0/67	0/4489
دمای حداقل ماه آوریل (°C)	0/57	0/3249
مقدار ویژه	-	1/7439

بدین ترتیب مهمترین مؤلفه اصلی با استفاده از بردار ویژه و مقدار ویژه تعیین می‌گردد. جدول ۴ اولین فاکتور یا مؤلفه اصلی را که با روش فوق تعیین شده است نشان می‌دهد.

جدول ۴ تعیین مهمترین مؤلفه اصلی

متغیرها	ضرایب عامل
دمای حداقل ماه ژانویه (°C)	0/74
دمای حداقل ماه فوریه (°C)	0/65
دمای حداقل ماه مارس (°C)	0/67
دمای حداقل ماه آوریل (°C)	0/57

به طوری که ملاحظه می‌شود متغیرهای دماهای حداقل ماههای ژانویه، فوریه و مارس دارای ضرایب بیش از 0/6 می‌باشند بنابراین به عنوان متغیرهای مهم و مجموعه آنها به عنوان مؤلفه اصلی یا اولین فاکتور در نظر گرفته می‌شود (کلاین، 1994). چنانکه ملاحظه می‌شود، این متغیرها نسبت به آخرین متغیر (دمای حداقل ماه آوریل) که دارای ضریب 0/57 می‌باشد سهم بیشتری در تبیین کل تغییرات داشته‌اند. لازم به ذکر است که متغیرهای فوق کاملاً فرضی بوده و فقط به منظور نشان دادن فرآیند محاسبات به کار رفته‌اند. داده‌های آماری مورد مطالعه شامل هشت متغیر در کلیه ماههای سالهای 1993 تا 1995 می‌باشند که مربوط به 41 ایستگاه هواشناسی سینوپتیک کشور هستند.

مرحله ششم - تشکیل "ماتریس باقیمانده"^{۱۸} برای به دست آوردن دومین مؤلفه اصلی: دومین

مؤلفه اصلی از ماتریس باقیمانده ماتریس اولیه بوده و از جدول ۲ استخراج می‌شود. سایر محاسبات دقیقاً نظیر مراحل ذکر شده‌ای است که برای مؤلفه اول انجام گردید. ماتریس باقیمانده‌ها با توجه به ضرایب عامل متغیرها به شرح زیر محاسبه می‌شوند:

ابتدا حاصل ضرب ضرایب عامل محاسبه شده برای هریک از متغیرها (جدول ۳) در خود همان ضرایب عامل بدست می‌آید. به‌طور مثال ضریب عامل متغیر دمای حداقل ژانویه یکبار در خود، یکبار در متغیر دمای حداقل فوریه، یکبار در دمای حداقل مارس و در نهایت در دمای حداقل آوریل و الی آخر ضرب می‌شود. پس از آنکه همه ضرایب در یکدیگر ضرب شدند، جدولی نظیر جدول ۵ تنظیم می‌شود. به این ماتریس، ماتریس متقاطع گفته می‌شود (Mardia et al, 1979).

جدول ۵ ماتریس متقاطع حاصلضرب ضرایب عامل متغیرها با یکدیگر

دمای حداقل ماه ژانویه (°C)	دمای حداقل ماه فوریه (°C)	دمای حداقل ماه مارس (°C)	دمای حداقل ماه آوریل (°C)
۰/۵۵	۰/۴۸	۰/۵۰	۰/۴۲
۰/۴۸	۰/۴۳	۰/۴۴	۰/۳۷
۰/۵۰	۰/۴۴	۰/۴۵	۰/۳۸
۰/۴۲	۰/۳۷	۰/۳۸	۰/۳۲

پس از تنظیم جدول ۵، مقادیر این جدول از ماتریس اولیه یعنی مقادیر جدول ۲ کم می‌شوند. اعداد به دست آمده، ماتریس باقیمانده خواهند بود. در نهایت با تعیین مقدار ویژه و بردار ویژه بر اساس این ماتریس، محاسبه دومین مؤلفه انجام می‌گیرد. جدول ۶ ماتریس باقیمانده را نشان می‌دهد:

جدول ۶ اولین ماتریس باقیمانده

دمای حداقل ماه ژانویه (°C)	دمای حداقل ماه فوریه (°C)	دمای حداقل ماه مارس (°C)	دمای حداقل ماه آوریل (°C)
۰/۴۶	-۰/۸۰	-۰/۲۰	-۰/۲۲
-۰/۸۰	-۰/۵۷	-۰/۲۴	-۰/۲۷
-۰/۲۰	-۰/۲۴	۰/۵۵	-۰/۰۸
-۰/۲۲	-۰/۲۷	-۰/۰۸	۰/۶۸

بدین ترتیب روشن می‌شود که چرا اولین مؤلفه به‌خوبی می‌تواند بیشترین سهم را در تبیین واریانس داده‌ها به خود اختصاص دهد. سایر مؤلفه‌ها و عوامل نیز به همین شکل از ماتریس‌های باقیمانده استخراج می‌شوند تا اینکه آخرین مؤلفه، که کمترین و ناچیزترین سهم را در بین متغیرها

دارد، مشخص شود. ممکن است ضرایب آخرین مؤلفه‌ها به صفر هم برسد (کلاین، ۱۹۹۴). پس از تعیین مؤلفه‌ها و میزان درصد تغییراتی که هر یک از آنها تبیین می‌کنند می‌توان با توجه به موضوع مورد مطالعه به تحلیل و تفسیر متغیرها پرداخت.

پیش از انجام آنالیز تحلیل عاملی چند آزمون ویژه جهت بررسی ماتریس انجام داده شده است که در اینجا به اختصار شرح داده می‌شوند:

آزمون بارلت^{۱۹}: هدف از این آزمون آن است که ماتریس همبستگی نباید ماتریس واحد باشد. منظور از ماتریس واحد، ماتریسی است که کلیه ازنه‌های آن صفر و عناصر قطری آن یک باشد (ماتر، ۱۹۷۶). این آزمون در نرم افزار SPSS قابل اجرا می‌باشد. در مورد این آزمون باید توجه داشت که اگر سطح معنادار کمتر از ۰/۰۵ باشد، در این صورت ماتریس همبستگی، ماتریس واحدی نخواهد بود و از این نظر برای تحلیل عاملی ماتریس مناسبی است (SPSS, 1999).

شاخص کایزر- مایر- اولکین^{۲۰} KMO: این شاخص به منظور مقایسه مقادیر همبستگی مشاهده شده با مقادیر ضرایب همبستگی جزئی (همبستگی زوج- زوج متغیرها) به کار می‌رود. هر چقدر مقدار این شاخص کمتر باشد، معلوم می‌شود که متغیرها برای تحلیل عامل مناسب نیستند. مقادیر ۹۰ (بالا ترین رقم)، ۸۰ و ۷۰ به ترتیب خیلی خوب، خوب و متوسط (قابل قبول) ارزیابی می‌شوند و چنانچه این شاخص زیر ۵۰ باشد غیر قابل قبول است (SPSS, 1999). در این مطالعه شاخص کایزر- مایر ۸۷ به دست آمده است (KMO = 87) که نشان می‌دهد متغیرهای مورد مطالعه از همبستگی بالایی برخوردار هستند و لذا ماتریس همبستگی برای این تحلیل مناسب تشخیص داده می‌شود.

تجزیه و تحلیل داده‌ها

پس از وارد کردن داده‌های خام و انجام مراحل ذکر شده، داده‌های مورد نظر آنالیز گردیده‌اند. جدول ۷ درصد تغییرات تبیین شده در مؤلفه‌ها را نشان می‌دهد. به طوری که ملاحظه می‌شود اولین، دومین و سومین مؤلفه‌ها به ترتیب ۴۰/۱۹٪، ۲۳/۳۷٪ و ۷/۱۳٪ تغییرات را از کل واریانس به خود اختصاص داده‌اند. اما نکته مهم این است که در یابیم کدام متغیرها دارای همبستگی بالایی با ضرایب عامل در اولین مؤلفه بوده‌اند. بدین منظور به جدول ضرایب عامل مراجعه می‌شود.

19. Bartlett's test if sphericity.

20. Kaiser-Mayer-Olkin.

جدول ۷ مؤلفه‌های اصلی و درصد واریانس تبیین شده از کل واریانس

مؤلفه‌ها	مجموع مقادیر ویژه	درصد واریانس تبیین شده	درصد حجمی
۱	۱۱۷/۸۰۳	۴۰/۹۰۴	۴۰/۹۰۴
۲	۶۷/۳۱۸	۲۳/۳۷۴	۶۴/۲۷۸
۳	۲۰/۵۳۹	۷/۱۳۱	۷۱/۴۱۰
۴	۱۵/۳۹۱	۵/۳۴۴	۷۶/۷۵۴
۵	۱۲/۱۰۶	۴/۲۰۳	۸۰/۹۵۷
۶	۷/۲۳۴	۲/۵۱۲	۸۳/۴۶۹
۷	۵/۲۰۸	۱/۸۰۸	۸۵/۲۷۷
۸	۴/۹۴۸	۱/۷۱۸	۸۶/۹۹۵
۹	۴/۳۴۳	۱/۵۰۸	۸۸/۵۰۳
۱۰	۳/۸۸۰	۱/۳۴۷	۸۹/۸۵۱
۱۱	۳/۱۴۰	۱/۰۹۰	۹۰/۹۴۱
۱۲	۲/۷۳۸	۰/۹۵۱	۹۱/۸۹۱
۱۳	۲/۴۱۲	۰/۸۳۸	۹۲/۷۲۹
۱۴	۲/۲۴۲	۰/۷۷۸	۹۳/۵۰۷
۱۵	۲/۰۰۸	۰/۶۹۷	۹۴/۲۰۵
۱۶	۱/۸۷۴	۰/۶۴۱	۹۴/۸۴۶
۱۷	۱/۷۲۶	۰/۵۹۹	۹۵/۴۴۵
۱۸	۱/۵۷۷	۰/۵۴۸	۹۵/۹۹۳
۱۹	۱/۳۹۶	۰/۴۸۵	۹۶/۴۷۸
۲۰	۱/۱۵۸	۰/۴۰۲	۹۶/۸۸۰
۲۱	۱/۰۹۴	۰/۳۸۰	۷۹/۲۵۹

جدول ۸ همبستگی ضرایب عامل را با مؤلفه‌های بدست آمده نشان می‌دهد. نظر به اینکه جدول اصلی بسیار طولانی می‌باشد، لذا در اینجا به صورت خلاصه فقط ابتدای جدول اصلی انتخاب گردیده و نتایج آن نشان داده شده است. بدیهی است جدول اصلی حاوی کلیه ضرایب همبستگی مربوط به هشت متغیر و برای کلیه ماههای سال از تعداد ۴۱ ایستگاه می‌باشد. با توجه به جدول ضرایب عامل، متغیرهایی که دارای ضرایب بالای ۰/۶ تا ۰/۹ بوده‌اند با علامت (*) مشخص شده‌اند. مثلا مقادیر نم نسبی ایستگاه آبادان در اولین مؤلفه دارای بالاترین ضرایب عامل می‌باشند. نتایج مربوط به سایر ایستگاهها نشان می‌دهند که نم نسبی بالاترین تغییرات را نسبت به کل واریانس تبیین نموده است (Eigen Value = 117.803، Percent of Variance = 40.904%). سایر ضرایب (کمتر از ۰/۵) کاملا ناچیز بوده و قابل صرفنظر کردن می‌باشند. به طوری که ملاحظه می‌شود مقادیر نم نسبی در مولفه دوم سهم بسیار ناچیزی در تبیین تغییرات دارند. برعکس دماهای حداقل و حداکثر بالاترین ضرایب را به خود اختصاص داده‌اند.

با توجه به اینکه نتایج مربوط به سایر ایستگاهها نیز بسیار مشابه ایستگاه آبادان می باشند، از اینرو دماهای حداقل و حداکثر بعنوان دومین متغیر اصلی که سهم بالایی از تغییرات را توجیه می کنند در نظر گرفته شده اند ($\text{Eigen Value} = 67.318$ ، $\text{Percent of Variance} = 23.374\%$).

نتایج بررسی ضرایب عامل سایر مؤلفه ها نشان می دهند که متغیرهای دیگر به طور متفاوت در بقیه مؤلفه ها بارگذاری شده اند و مقادیر مجموع بارش ماهانه و تعداد روزهای با بارش بیشتر از ۱۰ میلیمتر در مؤلفه سوم بیشتر از سایر متغیرها همبستگی نشان داده است ($\text{Eigen Value} = 20.539$ ، $\text{Percent of Variance} = 7.131\%$). مؤلفه های چهارم و پنجم نیز به ترتیب ۵/۳۴ و ۴/۲ درصد از کل تغییرات را تبیین نموده اند.

جدول ۸ همبستگی ضرایب عامل و مؤلفه ها (ایستگاه آبادان)

مؤلفه ها یا فاکتورها			سال	متغیرها
۳	۲	۱		
۰/۰۰۵	۰/۱۶۲ *	۰/۰۰۵	۱۹۹۳	دمای حداقل ماه ژانویه (°C)
۰/۰۰۲	۰/۱۴۱ *	۰/۱۳۷	۱۹۹۴	دمای حداقل ماه ژانویه (°C)
۰/۰۲۵	۰/۱۵۹ *	۰/۱۵۰	۱۹۹۵	دمای حداقل ماه ژانویه (°C)
۰/۲۶۱	۰/۱۸۷۶ *	۰/۱۳۸	۱۹۹۳	دمای حداکثر ماه ژانویه (°C)
۰/۳۰۳	۰/۵۸۹۳ *	۰/۲۰۹	۱۹۹۴	دمای حداکثر ماه ژانویه (°C)
۰/۲۶۵	۰/۷۵۲ *	۰/۲۸۸	۱۹۹۵	دمای حداکثر ماه ژانویه (°C)
۰/۵۲۶	۰/۵۲۹	۰/۷۶۸ *	۱۹۹۳	نم نسبی اول ژانویه ۹۳ (%)
۰/۲۶۳	۰/۴۶۱	۰/۶۹۰ *	۱۹۹۴	نم نسبی اول ژانویه ۹۴ (%)
۰/۴۴۲	-۰/۱۳۹	۰/۷۷۰ *	۱۹۹۵	نم نسبی اول ژانویه ۹۵ (%)
۰/۰۰۶	-۰/۲۸۸	۰/۸۹۰ *	۱۹۹۳	نم نسبی دوم ژانویه ۹۳ (%)
۰/۱۵۴	۰/۱۰۵	۰/۷۳۹ *	۱۹۹۴	نم نسبی دوم ژانویه ۹۴ (%)
۰/۱۱۵	-۰/۰۰۳	۰/۸۲۱ *	۱۹۹۵	نم نسبی دوم ژانویه ۹۵ (%)
۰/۰۰۳	-۰/۱۴۳	۰/۷۵۴ *	۱۹۹۳	نم نسبی سوم ژانویه ۹۳ (%)
۰/۱۰۷	۰/۱۶۷	۰/۹۲۰ *	۱۹۹۴	نم نسبی سوم ژانویه ۹۴ (%)
۰/۵۱۷	۰/۲۵۳	۰/۷۳۲ *	۱۹۹۵	نم نسبی سوم ژانویه ۹۵ (%)
۰/۷۶۳ *	۰/۳۲۰	۰/۴۳۴	۱۹۹۳	بارش ماهانه ژانویه ۹۳ (mm)
۰/۶۳۹	۰/۱۵۹	۰/۵۹۹	۱۹۹۴	بارش ماهانه ژانویه ۹۴ (mm)
۰/۸۹۹ *	۰/۱۳۴	۰/۴۲۱	۱۹۹۵	بارش ماهانه ژانویه ۹۵ (mm)

بحث یافته ها و نتیجه گیری نهایی

از آنجایی که مطالعه حاضر صرفاً در جهت تکمیل نتایج تحقیق قبلی مؤلفان در خصوص طبقه بندی اقلیمی با استفاده از روش های آماری چند متغیره بوده است، بنابراین لازم است پیش از بحث در مورد یافته های این تحقیق، ابتدا به طور خلاصه به نتایج حاصل از روش تحلیل خوشه ای و طبقه بندی اقلیمی که با استفاده از داده های ۴۱ ایستگاه هواشناسی و از سال ۱۹۶۶ تا ۱۹۹۵ (دوره ۳۰ ساله)

انجام گردیده است اشاره شود. داده‌های آماری ماهانه هشت متغیر اقلیمی قبلا معرفی شده‌اند. نتایج تحقیقی که بر اساس روش تحلیل خوشه‌ای برای طبقه‌بندی اقلیمی ایران انجام پذیرفت، منجر به تعیین پنج منطقه مشخص اقلیمی به اسامی معتدل خزری، سرد کوهستانی، گرم نیمه خشک، گرم خشک و گرم ساحلی گردید. این طبقه‌بندی بر مبنای متغیرهای اقلیمی هشتگانه و با استفاده از روشهای آماری چند متغیره انجام گرفت ولی در عین حال پرسشی نیز بدین صورت مطرح شد که آیا روش تحلیلی بکار رفته توانسته است کلیه متغیرها را در این طبقه‌بندی بگنجانند؟ و اگر جواب مثبت است کدام متغیرها بیش از بقیه در طبقه‌بندی تأثیرگذار بوده‌اند؟ یافته‌های تحقیق حاضر پاسخ این سؤالات را به‌خوبی فراهم نموده است. از میان متغیرهای مورد مطالعه، نم نسبی که بیشترین درصد واریانس کل تغییرات را به‌خود اختصاص داده است، به‌عنوان اولین متغیر زمینه‌ای و مهمترین مؤلفه تعیین گردید. این یافته‌گویای آن است که نم نسبی نقش مهمی در تعیین مرزبندی نواحی اقلیمی تعیین شده داشته است. درجه حرارت دومین متغیر مهم است که در این طبقه‌بندی مؤثر بوده است. سایر متغیرها اگرچه درصدهای کمتری را به‌خود اختصاص داده‌اند اما کماکان تأثیر آنها به‌خوبی در ضرایب عامل منعکس شده است. به‌طوری‌که مؤلفه سوم دربرگیرنده ضریب همبستگی بالایی با مجموعه‌ای از سایر متغیرها می‌باشد. این یافته کاملاً مبنای علمی داشته و متکی بر اصول و قوانین هواشناسی و اقلیم‌شناسی می‌باشد. در این باره لوتنگنز و تربوک^{۲۱} (۱۹۹۸) چنین اشاره می‌کنند که نم نسبی در رابطه با مقدار بخار آب موجود در هوا بوده و تغییر در نم نسبی به‌طور طبیعی با ایجاد تغییر در مقدار رطوبت یا دمای هوا صورت می‌گیرد. مقدار رطوبت مورد نیاز برای رسیدن به شرایط اشباع و ایجاد بارندگی نیز به درجه حرارت هوا وابسته است. از سوی دیگر نامبردگان اشاره می‌کنند که نم نسبی نشان‌دهنده فاصله مقدار رطوبت هوا تا شرایط اشباع است که خود مبین ارتباط نم نسبی با بارندگی می‌باشد. لذا کاملاً واضح است که چرا نم نسبی حداکثر واریانس را به خود اختصاص داده است. در حقیقت چنین نتیجه می‌شود که نم نسبی دارای حداکثر همبستگی با سایر متغیرها بوده و شدت این همبستگی از درجه حرارت نیز بیشتر بوده است. لازم به ذکر است که طبقه‌بندی انجام شده، یک طبقه‌بندی کلی بوده و به‌عنوان یک رهیافت اولیه برای انجام مطالعات مفصل‌تر، به‌خصوص مطالعاتی که جنبه مقایسه بین اقلیم نواحی مختلف را دارند، مطرح است. همچنان‌که هدف اصلی از طبقه‌بندی مذکور، بررسی و پیش‌بینی تغییرات دما و بارش در ایران بوده است.

نهایتاً یافته این تحقیق‌گویای آن است که طبقه‌بندی انجام شده با توجه به هدف اصلی مطالعه،

(تعیین نواحی ممتاز اقلیمی ایران) اقدامی موفق بوده و بر مبنای اصول علمی استوار می‌باشد. به‌طور مثال ملاحظه می‌شود که در این طبقه‌بندی منطقه گرم نیمه خشک که دربرگیرنده ایستگاههای حاشیه کویر شامل: سمنان، تهران، سبزوار، کاشان، شیراز، یزد، زاهدان، کرمان، بیرجند، اصفهان و طبس و همه از لحاظ مقدار نم نسبی، دما و بارش دارای شرایط یکسانی هستند، به‌خوبی از سایر مناطق متمایز گردیده است. همچنین منطقه معتدل خزری نیز که شامل ایستگاههای بابلسر، گرگان، انزلی، رشت و رامسر می‌باشد، با این روش به‌طور بسیار مشخصی از سایر مناطق مجزا شده است. لازم به ذکر است که نتایج حاصل از تطبیق این طبقه‌بندی اقلیمی با سایر طبقه‌بندی‌ها به همراه نقشه‌های طبقه‌بندی اقلیمی در پژوهش قبلی مؤلفان آورده شده است (ترابی و همکاران، ۱۳۸۱). همچنین این مطالعه، کاربرد روش تحلیل عاملی و اهمیت آن را در بررسیهای مفصل به‌خوبی نشان می‌دهد. در مطالعاتی که نیاز به تحلیل و کنکاش بیشتر در مورد روابط بین متغیرها وجود داشته باشد این روشها یافته‌های مناسبی را در اختیار پژوهشگر قرار می‌دهند به‌نحوی که می‌توان از آنها برای طراحی پیش فرض‌های اولیه در تحقیقات بعدی استفاده نمود.

منابع و مآخذ

۱. ترابی سیما، سعید جهانخوش، بهلول علیجانی و خلیل شفیمی (۱۳۸۰): طبقه‌بندی اقلیمی ایران: کاربرد روش چند متغیره، مجله پژوهشهای جغرافیایی، مؤسسه جغرافیای دانشگاه تهران، شماره ۳۹.
۲. مقدم محمد، سیدابولقاسم محمدی شوطی و مصطفی آقایی سربزه، ترجمه انتشارات پیش‌تاز علم، (۱۳۷۳): آشنایی با روش‌های آماری چند متغیره، تبریز.
3. Drosowsky, W (1993): An analysis of Australian seasonal rainfall anomalies: 1950-1987, Int. J. of Climatology, Vol.13, 1-30.
4. Ebdon David (1977): **Statistics in Geography: A Practical Approach**, Basil Blackwell.
5. Esteban-Para, M.J. and Rodrigo, F.S (1995): Temperature trends and change points in the northern Spanish Plateau during the last 100 years, Int. J. of Climatology, Vol.15, 1031-1042.
6. Jensen John, R (1986): **Introductory Digital Imagery Processing**, Printice-Hall, Englewood Cliffs, N.J.O. 7632.
7. Johnson A., Richard and Wichem Dean, (1988): **Applied Multivariate Statistical Analysis**, Prentice Hall International Inc.
8. Kelly P.M. and Jones, P.D (1999): **Spatial patterns of variability in the global surface air temperature data set.**, J.of Geophysical Research, Vol. 104, No.D20, 24237-24256.
9. Kline Paul (1994): **An Easy Guide to Factor Analysis**, Rutledge, London and New York.
10. Kripalani, R.H., S. Inamdar and N. A. Sontakke (1998): **Rainfall variability over Bangladesh and Nepal: Comparison and connections with features over India**, Int. J. of climatology, Vol. 16, 689-703.
11. Lutgens Frederick K. and Edward J. Tarbuck (1998): **The Atmosphere**, Prentice-Hall International, Inc.
12. Mardia, K.V., Kent J.T. and Mibby J.M (1979): **Multivariate Analysis**. Academic press, London, 521pp.
13. Mather Paul M (1976): **Computational Methods of Multivariate Analysis in Physical Geography**, John Wiley & Sons.
14. Mather Paul and Brandt TSO (2001): **Classification Methods for Remotely Sensed Data**, Taylor and Francies Ltd.
15. Mestas-Nunez A. M (2000): **Orthogonality properties of rotated empirical methods**, Int. J. of Climatology, 20:1509-1516.
16. Mulaik, S.A (1972): **The Foundations of Factor Analysis**, New York, Mc Graw-Hill.
17. Ogallo L (1989): **The spatial and temporal patterns of east African seasonal rainfall derived from principal component analysis**, Int. J. of Climatology, Vol.9, 145-167.
18. Shaw Gareth and Wheeler Dennis (1985): **Statistical Techniques in Geographical Analysis**, John Wiley & Sons Ltd.
19. Small, J. and Witherick, M (1990): **Modern Geographical Dictionary**, Prentice Hall Inc.
20. SPSS (1999): **SPSS for Windows Advanced Statistics Release 9.0** SPSS Inc.
21. Strahler A. and Arthur Strahler (1994): **Introducing Physical Geography**, John Wiley & Sons Inc., New York, 575 PP.
22. Terjung, W.H (1996): **Physiologic Climates of the Coterminous United States**, Am. Asso. Geog. Annals, 60, 466-492.
23. Willmott C.J (1978): **P-mode Principal Component Analysis Grouping and Precipitation Regions in California**, Bioclimatology, Springer Verlag, Vol.26, No.4.
24. Whetton, P.H (1988): **A Synoptic Climatological Analysis of Rainfall Variability in Southeastern Australia**, J. of Climatology, Vol.8, 155-177.