

## الگوهای جدید برآورد رطوبت نسبی روزانه در ناحیه کوهستانی ایران

محبوبه فرزندی<sup>۱\*</sup>، سیدحسین ثنایی نژاد<sup>۲</sup>، بیژن قهرمان<sup>۳</sup>، حجت رضایی پزند<sup>۴</sup>

۱- کارشناسی ارشد هواشناسی کشاورزی، دانشگاه فردوسی مشهد

۲- دکتری، دانشیار دانشگاه فردوسی مشهد

۳- دکتری، استاد دانشگاه فردوسی مشهد

۴- مربی، کارشناسی ارشد، دانشگاه آزاد اسلامی واحد مشهد

تاریخ وصول مقاله: ۱۳۹۰/۰۸/۱۱، تاریخ پذیرش مقاله: ۱۳۹۱/۰۴/۰۲

### چکیده

روش مرسوم برآورد میانگین رطوبت نسبی روزانه معدل‌گیری از رطوبت‌نسبی ساعت‌های استاندارد (۰۶:۳۰، ۱۲:۳۰، ۱۸:۳۰) است. منحنی رطوبت‌نسبی روزانه چولگی دارد و متقارن نیست. بنابراین استفاده از ضرایب مساوی برای برآورد متوسط روزانه دقیق نیست. همچنین فقط رطوبت نسبی نیمی از شبانه‌روز در این روش‌ها در نظر گرفته می‌شود. افزون بر این منحنی‌های رطوبت نسبی روزانه به شرایط اقلیمی (خوشه-بندی) و ماه نیز وابسته است. ۱۴۹ ایستگاه همدید ایران مطالعه شد. ابتدا خوشه‌بندی ایران برای تعیین ناحیه‌های همگن و هم‌نوا با نه عامل هوا و اقلیم‌شناسی در دسترس (دما، بارش، تبخیر، رطوبت نسبی، دامنه تغییرات دمای سالانه، ارتفاع ایستگاه و سه نمایه اقلیم‌بندی دومارتن، ایوانف و تورنت‌وایت) انجام شد. روش افرازی میانه محور با سه خوشه (نوار ساحلی، کوهستانی و بیابانی- نیمه‌بیابانی) به حجم‌های ۴۲، ۴۸، ۶۰ به دست آمد. این مقاله ناحیه کوهستانی را انتخاب و الگوهایی جدید برای برآورد رطوبت نسبی روزانه ارائه داده است. چون حجم داده‌های خوشه کوهستانی زیاد است ناچار به استفاده از طرح نمونه‌گیری سیستماتیک دوری شد. الگوهای مختلف رگرسیونی خطی و غیرخطی (خطی‌پذیر) پس از غربال و آماده‌سازی بر داده‌های ساعتی برآزش و سیزده الگوی برتر در ماه‌های مختلف و سالانه انتخاب شدند. مقادیر بالای ضریب تعیین تصحیح شده (Adj-R<sup>2</sup>) و آماره F قدرت بالای الگوها را نشان می‌دهد. کمترین ضریب تعیین این ۱۳ الگو (۰/۹۸۹) مربوط به ماه ژانویه و بیشترین آن (۰/۹۹۷) مربوط به ماه سپتامبر است. واسنجی و مقایسه الگوهای ارائه شده در این مقاله با الگوهای مرسوم قدیمی نشان از دقت بیشتر الگوهای ارائه شده در این پژوهش دارد. معیار مقایسه میانگین مربع‌خطا است. این معیار نشان داد که الگوهای ارائه شده در این مقاله خطای کمتری دارند به طوری که MSE ماه‌های مختلف تقریباً نصف الگوی قدیمی (رابطه ۱) و بسیار کمتر از آرگون (رابطه ۲) است. متغیرهای مورد استفاده در الگوهای رگرسیونی رطوبت نسبی روزانه عبارت‌انداز: رطوبت نسبی ساعتی، معدل واقعی رطوبت‌نسبی روزانه. معدل رطوبت نسبی روز قبل، دمای حداکثر، حداقل و میانگین روزانه و بارش روزانه.

**واژگان کلیدی:** الگوی رگرسیونی، متوسط رطوبت‌نسبی روزانه، خوشه‌بندی، مناطق کوهستانی، نمونه‌گیری سیستماتیک.

## مقدمه

رطوبت نسبی یکی از متغیرهای مهم جو است که در مطالعات و پژوهش‌های کشاورزی، سدسازی، هیدرولوژی، شهرسازی، تعیین تبخیر از سطح آزاد آب و تبخیر و تعرق و... به کار می‌رود. اغلب متوسط روزانه، هفتگی، ماهانه و سالانه رطوبت نسبی مورد نیاز است. معدل روزانه اساس و پایه سایر فاصله‌های زمانی فوق است. بنابراین افزایش دقت در برآورد روزانه سبب افزایش دقت در سایر فواصل زمانی است. روش مرسوم و قدیمی برای برآورد میانگین رطوبت نسبی روزانه (رابطه ۱) معدل‌گیری از رطوبت نسبی سه ساعت استاندارد (۶:۳۰، ۱۲:۳۰، ۱۸:۳۰) است که به عنوان یک دستورالعمل استاندارد جهانی در سازمان هواشناسی ایران و وزارت نیرو استفاده می‌شود (آرشیو وزارت نیرو و سالنامه‌های آماری، سازمان هواشناسی کشور، ۲۰۰۶).

$$\overline{RH} = 0.33RH_{6:30} + 0.33RH_{12:30} + 0.33RH_{18:30} \quad (1)$$

نقطه ضعف‌هایی در رابطه (۱) وجود دارد که سبب بزرگی خطای محاسبه توسط این رابطه می‌شود. منحنی رطوبت نسبی روزانه چوله است و رفتار غیرخطی دارد. یانو نشان داد که تابع بتا برازش خوبی بر منحنی رطوبت نسبی روزانه دارد (Yao, 1974). لذا استفاده از رابطه استاندارد شده (۱) خطا آمیز است و استفاده از ضرایب مساوی برای محاسبه متوسط روزانه دقیق نیست. هم چنین رابطه (۱) فقط رطوبت نسبی نیمی از شبانه‌روز را در نظر می‌گیرد. افزون بر این منحنی‌های رطوبت نسبی روزانه به شرایط اقلیمی و به ماه‌های مختلف سال نیز وابسته است. بنابراین نیاز به خوشه‌بندی اقلیمی است تا روابط مناسب در هر اقلیم با توجه به داده‌های مربوط ارائه شود. ناحیه بندی یا تفکیک زمانی - مکانی به واحدهای نسبتاً مستقل براساس یک یا چند معیار مفروض اقلیم‌شناسی و جغرافیایی همواره در کانون توجه اقلیم‌شناسان بوده است. امروزه اقلیم شناسان واقعیت‌های موجود در پهنه‌های اقلیمی را با استفاده از روش‌های نوین آماری با حداقل خطا به کار می‌برند (Theodoridis & Koutroumbas, 2009) خوشه به مجموعه‌ای از داده‌ها گفته می‌شود که از نظر ویژگی‌های مورد بررسی به هم شبیه باشند.

محققین مختلف روش‌هایی را برای محاسبه رطوبت نسبی به طور مستقیم و غیر مستقیم برای کاهش خطای رابطه (۱) ارائه داده‌اند. این موارد هیچ کدام خوشه‌بندی را رعایت نکردند. چند نمونه به شرح زیر است. سازمان تحلیل اقلیم‌های فضایی در دانشگاه ایالتی آرگون رابطه (۲) را برای به عنوان جانشین رابطه (۱) برای محاسبه متوسط رطوبت نسبی روزانه پیشنهاد کرد که برای ماه‌های مختلف کمی اریب است (با مقدار واقعی اختلاف دارد، Corvallis, 2008). این رابطه نشان می‌دهد که متوسط روزانه فقط به رطوبت نسبی ساعت ۳ و ۱۵ گریونچ (صبح و بعد از ظهر) وابسته است. همچنین مشاهده می‌شود که اثر رطوبت نسبی صبح تقریباً ۲ برابر رطوبت نسبی بعد از ظهر است.

$$\overline{RH} = 0.65RH_{morning} + 0.35RH_{afternoon} \quad (2)$$

ادواردلینارس دو رابطه (۳) و (۴) را برای تخمین رطوبت نسبی به کمک دمای نقطه شبنم و دمای هوا ارائه داد (Linarce, 1991).

$$RH = 100 - 4(T - T_d) \quad (3)$$

$$T_d = 10.9 + 0.63T - 0.53Rd - 0.35Ra - 0.0023H \quad (4)$$

Td میانگین ماهانه دما، T میانگین ماهانه دما، Ra دامنه تغییرات سالانه دما، Rd دامنه تغییرات روزانه دما و H ارتفاع ایستگاه (متر) است. دماها بر حسب درجه سلسیوس است. هولمبو و همکاران (۱۹۴۵) رابطه بین دما و لگاریتم رطوبت نسبی را مطابق رابطه (۵) بررسی کردند، (Holomboe, 1945) c و d دو پارامتر ثابت‌اند که به دمای نقطه شبنم و دمای هوا وابسته هستند.

$$\log RH = c - dT \quad (5)$$

کورت و واکو اظهار داشتند که میانگین رطوبت نسبی روزانه که از معدل قرائت‌های صبح و بعد از ظهر به دست می‌آید دقیق نیست و بیشتر از معدل واقعی ۲۴ ساعته است. همچنین آن‌ها متوجه شدند که این معدل نیز بیشتر از معدل کمینه و بیشینه رطوبت نسبی روزانه است. آن‌ها شهر شیکاگو را محل پژوهش خود قرار دادند و بیان کردند که میانگین روزانه به ماه مورد نظر نیز بستگی دارد (Court and Waco, 1956) دی میانگین رطوبت نسبی ماهانه را

وابسته است. هدف مقاله حاضر ارائه روابط جدیدی برای برآورد رطوبت نسبی روزانه از روی رطوبت ساعات مختلف روز، همچنین اضافه کردن دما و بارش روزانه در این روابط است. افزون بر این اقلیم و ماه را نیز در روابط دخالت داده و روابطی مناسب برای ماه‌های مختلف در اقلیم (خوشه) کوهستانی ارائه شده است. ایران با روش افزایشی میانه محور (با ۹ متغیر) خوشه‌بندی و به سه خوشه تفکیک شده است. مقاله حاضر روابط را برای خوشه کوهستانی در نظر گرفته است. روش نمونه‌گیری سیستماتیک برای انتخاب نمونه مناسب استفاده شده است. مقایسه نتایج حاصل از روابط ارائه شده در این مقاله با رابطه (۱) نیز توسط میانگین مربعات خطا (MSE) انجام که نشان از دقت روابط ارائه شده است.

## مواد و روش‌ها

### منطقه مورد مطالعه و داده‌ها

کشور ایران با وسعت ۱۶۴۸۰۰۰ کیلومتر مربع در نیم کره شمالی، در قاره آسیا و در قسمت غربی فلات ایران واقع شده است. این کشور بین دو نصف النهار  $44^{\circ}$  و  $64^{\circ}$  شرقی و دو مدار  $25^{\circ}$  و  $40^{\circ}$  شمالی قرار دارد. حدود ۹۰ درصد خاک ایران در محدوده فلات ایران واقع شده و کشوری کوهستانی محسوب می‌شود. رشته کوه‌های البرز و زاگرس مهمترین تاثیر را در آب و هوایی ایران دارند. این دو به ترتیب از شمال غربی به شمال شرقی و از شمال غربی به جنوب شرقی کشیده شده‌اند (بختیاری، ۱۳۸۳). ایران هم اکنون ۲۸۹ ایستگاه همدید دارد که آمار آن‌ها دقت بالایی دارند. آن‌ها به همین دلیل انتخاب شده‌اند. تقریباً نیمی از ایستگاه‌ها دارای آمار ناقص یا بسیار کوتاه مدت‌اند. حدود ۱۴۹ ایستگاه دارای آمار کامل با طول متوسط ۲۵ سال است. ایستگاه‌ها در سراسر ایران پراکنده‌اند. آمار ساعتی رطوبت نسبی، دمای حداقل، حداکثر و متوسط روزانه، بارش روزانه و رطوبت نسبی روزانه این ایستگاه‌ها در اختیار است (پایگاه اطلاع رسانی سازمان هواشناسی کشور).

محاسبه و نتیجه گرفتند که این میانگین به ماه، فصل و مختصات جغرافیایی وابسته است و خطا آمیز است. این خطا در برخی از نقاط و در برخی مناطق منفی است. آن‌ها جداول و نمودارهایی برای تصحیح محاسبه میانگین رطوبت نسبی روزانه در امریکا ارائه دادند (Day, 1917). بلانک میانگین رطوبت نسبی ایستگاهی را با رطوبت نسبی ۲۴ ساعته در چند نقطه امریکا مقایسه کرد. او اظهار داشت که میانگین حاصل از روش‌هایی که فقط بعضی از ساعات روز را در نظر می‌گیرد با معدل واقعی ۲۴ ساعته تفاوت دارد. این تفاوت به محل بررسی (موقعیت جغرافیایی) بستگی دارد (Blanc, 1961).

تعیین نواحی اقلیمی در دو سده گذشته عمدتاً مروهون الکساندر فون همبولت (۱۸۱۷)، ولادیمیر کوپن (۱۸۴۶-۱۹۴۰)، کارلوس لینه (۱۷۳۵)، لورک هوارد (۱۸۱۵)، تورنت وایت (۱۹۳۱) و پنمن (۱۹۰۹-۱۹۸۴) بوده است. این روش‌ها به دلیل استفاده از معیارها و آستانه‌های از پیش تعیین شده به عنوان روش‌های سنتی قلمداد می‌شوند (مسعودیان و عطایی، ۱۳۸۴). گوسن کشورهای حاشیه مدیترانه را با روش‌های تحلیل خوشه‌ای و مولفه‌های اصلی طبقه‌بندی کرد. او آمار ۹۰ ایستگاه را به کاربرد و حاشیه مدیترانه را به پنج ناحیه تقسیم کرد. این پنج ناحیه توسط بردارهای ویژه معنی‌دار انتخاب شدند (Goossens, 1985). جکسن و ویناند (۱۹۹۵) پهنه‌بندی بارش ایستگاه‌های مدیترانه‌ای را براساس ۳۴ متغیر اقلیمی جغرافیایی با روش خوشه‌بندی ساده انجام و سپس با سه روش وارد، پیوند متوسط و لنس- ویلیامز مقایسه نموده است (Jackson & Weinand, 1995). فریوان و کادی اوغلو (۲۰۰۸) کشور اردن را بر پایه متغیرهای بارش، رطوبت نسبی، حداقل و حداکثر دمای ۱۶ ایستگاه با طول دوره آماری ۳۰ سال به ۳ ناحیه بارش همگن تقسیم نمودند. روش آن‌ها براساس رسم منحنی‌های هم ضریب تغییرات و هم ضریب چولگی بنا شده است. آن‌ها براساس این ترسیمات ناحیه‌ها را به دست آوردند (Freiwan, & Kadioglu, 2008).

پژوهش‌های قبل نشان می‌دهد که برآورد رطوبت نسبی از روی سه ساعت استاندارد دقیق نیست و به ماه و اقلیم نیز

## خوشه‌بندی افزای میانه محور

بررسی شوند. روش گشتاوری خطی یک از کارآمدترین روش‌های تعیین همگنی و هم‌نویایی است (Hosking and Wallis, 1997).

## عرض سیلهوت

عرض سیلهوت اندازه‌ای برای یک بردار ویژگی است که مقدار مشابهت را با سایر بردارهای خوشه مربوط به خود و بردارهای سایر خوشه‌ها مقایسه می‌کند. عرض سیلهوت  $s(i)$  برای بردار ویژگی  $i$ ام در خوشه  $K$ ام مطابق رابطه (۷) است.

$$s(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}} \quad (7)$$

$a(i)$  متوسط فاصله بردار ویژگی  $i$ ام نسبت به سایر بردارهای ویژگی در خوشه  $k$ ام است.  $b(i)$  کمترین متوسط فاصله از بردار ویژگی  $i$ ام نسبت به سایر بردارهای ویژگی در خوشه‌های دیگر است. عرض سیلهوت بین  $-1$  و  $1$  است. اگر  $s(i)$  نزدیک به یک باشد، نشان دهنده خوشه بندی موفق است. اگر  $s(i)$  نزدیک به  $-1$  باشد، خوشه‌بندی بد و رد می‌شود. اگر  $s(i)$  نزدیک به صفر باشد، یعنی بردار ویژگی  $i$ ام به‌طور مساوی بین دو خوشه قرار دارد (روسو، ۱۹۸۷).

## آزمون‌های همگنی براساس روش گشتاورهای خطی

داده‌های هر خوشه باید از نظر همگنی و هم‌نویایی بررسی شوند. روش گشتاوری خطی یکی از کارآمدترین روش‌های تعیین همگنی و هم‌نویایی است. گشتاورهای خطی  $7$  به صورت ترکیب خطی از گشتاورهای وزن‌دار احتمالی  $8$  تعریف می‌شود (Hosking and Wallis, 1997). کاربرد گشتاورهای خطی آسان‌تر از گشتاورهای وزن دار احتمالی است. اگر امید ریاضی یک متغیر تصادفی متناهی باشد، آنگاه تمام گشتاورهای خطی آن وجود دارد و می‌توان تابع چگالی این متغیر را به طور کامل برحسب گشتاورهای خطی‌اش نوشت. پارامترهای موقعیت، مقیاس و شکل یک تابع توزیع با نسبت‌های گشتاوری خطی تفسیر

روش خوشه‌بندی افزای میانه محور<sup>۱</sup> در آب و هواشناسی توصیه شده است (Rao and Srinivas, 2008). این خوشه بندی بر روی ماتریس عدم تشابه مجموعه داده‌ها با تعریف یک ماتریس عدم تشابه عمل می‌کند. این روش استوارتر از سایر روش‌هاست. زیرا مجموع عدم تشابهات را کمینه می‌کند. در حالی که برخی از روش‌ها مانند روش میانگین محور فاصله اقلیدسی را کمینه می‌کند. همچنین عرض سیلهوت و نمودار سیلهوت را برای انتخاب تعداد اعضای بهینه خوشه نیز در بر می‌گیرد. برای هر خوشه یک شیئی یا مبنا (میانه) وجود دارد. این روش ابتدا  $k$  شیئی را به نام Medoids محاسبه می‌کند (برای  $k$  خوشه). هر میانه به صورت یک شیئی از یک خوشه تعریف می‌شود که متوسط عدم تشابه آن نسبت به سایر اشیا خوشه کمینه است. این نوع نمایش را در خوشه بندی (مرکز نوعی<sup>۲</sup>) می‌نامند. پس از یافتن مجموعه میانه‌ها هر عضو مجموعه داده‌ها به نزدیک ترین میانه خود مطابق رابطه (۶) نسبت داده می‌شود.

$$d(i, mv_i) \leq d(i, mw_i); \forall w = 1, \dots, k \quad (6)$$

شیئی  $i$  در خوشه  $v_i$  قرار می‌گیرد هنگامیکه میانه  $mv_i$  از هر میانه دیگر مانند  $mw$  به این خوشه نزدیک تر باشد.  $k$  باید به گونه‌ای باشد که تابع هدف  $(P = \sum d(i, mv_i))$  کمینه شود (Elgar, 2008 & Everitt, et.al, 2011). قطر خوشه بزرگ‌ترین عدم تشابه بین اشیا داخل این خوشه و تفکیک خوشه کوچک‌ترین عدم تشابه بین دو شیئی داخل و خارج خوشه است. هم چنین معدل فاصله تمام اشیا خوشه را از میانه، متوسط فاصله می‌نامند.

## اعتبارسنجی خوشه‌بندی

قبول یا رد خوشه‌بندی با آزمون‌های آماری انجام می‌شود. این آزمون‌ها متنوع‌اند. مهمترین آنها: ضریب همبستگی کوفنتیک<sup>۳</sup>، عرض سیلهوت<sup>۴</sup>، ناهمگنی<sup>۵</sup> و ناهم‌نویایی<sup>۶</sup> ناحیه‌ای گشتاوری خطی است (Romesburg, 2004). داده‌های هر خوشه باید از نظر همگنی و هم‌نویایی

1. Partitioning around medoids clustering  
3. Cophenetic correlation coefficient  
5. Heterogeneity  
7. Linear Moments

2. Centrotypes  
4. Silhouette width  
6. Discordancy  
8. Probability Weighted Moments

آورد.  $F_{3,n-4}$  چندک توزیع فیشر با ۳ و  $(n-4)$  درجه آزادی واقع در بالای  $100\alpha/n$  این توزیع است، (Hosking and Wallis, 1997).

### اندازه ناهمگنی<sup>۱</sup>

اندازه ناهمگنی تخمین درجه ناهمگنی در یک گروه از ایستگاه‌ها و ارزیابی همگن بودن ناحیه است. این معیار، تغییرات بین ایستگاهی گشتاور خطی نمونه را برای گروهی از ایستگاه‌ها با مقدار مورد انتظار برای یک ناحیه همگن مقایسه می‌کند. تمامی ایستگاه‌های واقع در یک ناحیه همگن جامعه، گشتاور خطی یکسانی دارند. در حالی که گشتاورهای خطی نمونه آن‌ها می‌توانند متفاوت باشند. اندازه ناهمگنی با شبیه‌سازی منطقه‌ای به کمک توزیع کاپا

چهار پارامتری انجام، آماره  $H_1 = \frac{V_1 - \mu_{v_1}}{\sigma_{v_1}}$  محاسبه و

ناهمگنی بر حسب آماره  $H_1$  آزمون می‌شود.  $\mu_{v_1}$  و  $\sigma_{v_1}$  به ترتیب میانگین و انحراف معیار داده‌های شبیه‌سازی شده است (هاسکینگ، ۱۹۹۷).

الف- اگر  $H_1 < 1$  باشد، ناحیه همگن تلقی می‌شود.  
ب- اگر  $1 \leq H_1 < 2$  باشد، ناحیه مشکوک به ناهمگنی است.

پ- اگر  $H_1 > 2$  باشد: قطعاً ناهمگن است.

### متوسط واقعی رطوبت نسبی روزانه

رفتار رطوبت نسبی در طول شبانه روز یک تابع غیر خطی و نامتقارن مانند  $f(t)$  با دامنه  $[a, b]$  است (توتونیان، ۱۳۷۱). مقدار دقیق این متوسط از انتگرال سطح زیرمنحنی مطابق رابطه (۱۱) به دست می‌آید.

$$T_{ave} = \frac{1}{b-a} \int_a^b f(t) dt \quad (11)$$

حل انتگرال فوق برای محاسبه رطوبت نسبی متوسط روزانه به روش عددی (قائده سیمسون) محاسبه شده است. داده‌های در دسترس سه ساعته هستند. بنابراین تقریب سیمسون انتگرال فوق مطابق رابطه (۱۲) است (حبیبی،

می‌شود (رضایی پزند، ۱۳۸۰). امروزه محققان رشته‌های هواشناسی و هیدرولوژی برای تایید ناحیه‌بندی و تحلیل فراوانی از روش گشتاورهای خطی بهره می‌برند. گشتاورهای خطی به داده‌های پرت و حجم اندک نمونه حساسیت کمی دارند. در حالی که گشتاورهای معمولی چنین نیستند. تخمین پارامترهای توزیع احتمالی منتخب، به ویژه برای نمونه‌های کوچک با این روش قابل اطمینان‌تر از روش‌های گشتاوری معمولی است. برآوردگرهای گشتاورهای خطی کاملاً نارایب هستند (هاسکینگ و والیس، ۱۹۹۷).

### اندازه ناهمنوایی

نسبت‌های گشتاورهای خطی نمونه هر ایستگاه (ضریب تغییرات LCV، چولگی LCS و کشیدگی LCK) یک نقطه در یک فضای سه بعدی در نظر گرفته می‌شود. یک گروه از ایستگاه‌ها ابری از این نقاط را تولید می‌کنند. نشانه ناهمنوایی هر نقطه (ایستگاه) میزان دوری آن از مرکز ابر است. دوری در اینجا مفهوم همبستگی بین نسبت‌های گشتاورهای خطی نمونه‌ای را تداعی می‌کند. اندازه ناهمنوایی یک ایستگاه به صورت روابط (۸) تا (۱۰) محاسبه می‌شود.

$$\bar{U} = N^{-1} \sum_{i=1}^N U_i \quad (8)$$

$$S = (N-1)^{-1} \sum_{i=1}^N (U_i - \bar{U})(U_i - \bar{U})^T \quad (9)$$

$$D_i = \frac{1}{3} (U_i - \bar{U})^T S^{-1} (U_i - \bar{U}) \quad (10)$$

$N$  تعداد ایستگاه‌های گروه،  $U_i$  بردار نسبت‌های گشتاوری خطی ایستگاه نام،  $\bar{U}$  معدل بردارهای  $U_i$ ،  $S^{-1}$  وارون ماتریس کوواریانس ایستگاه‌ها و  $D_i$  اندازه ناهمنوایی ایستگاه نام است. اگر  $D_i$  بزرگ باشد، آنگاه ایستگاه نام ناهمنوایی تلقی می‌شود. تعریف مقدار بزرگ بودن بستگی به تعداد ایستگاه‌های موجود در گروه دارد. مقدار حداکثر  $D_i$  تابعی از  $\alpha$  و  $n$  است. می‌توان آن را در سطح  $\alpha$  از آماره  $\max(D_i) = \frac{(n-1)F_{3,n-4}}{n-4+3F_{3,n-4}}$  به دست

نیز باید انجام شود (نیرومند، ۱۳۸۴ و رضایی پزند و بزرگ نیا، ۱۳۸۰).

### نرم افزارهای مورد استفاده

نرم افزار "R.12.1" برای خوشه‌بندی استفاده شده است (www.r-project.org). نقشه خوشه‌بندی ایران با نرم افزار "ArcView" رسم و ارائه شده است. نرم افزار SPSS.18 برای الگوبندی رطوبت نسبی روزانه استفاده شده است.

### تجزیه و تحلیل و نتایج

- خوشه‌بندی: خوشه‌بندی ۱۴۹ ایستگاه همدید ایران با شش متغیر دما، بارش، تبخیر، رطوبت نسبی و دامنه تغییرات دمای سالانه، ارتفاع ایستگاه و سه نمایه اقلیم بندی دومارتن، ایوانف و تورنت وایت با چند روش انجام شده است. انتخاب این متغیرها به علت در دسترس بودن و رابطه آنها با رطوبت نسبی است. نتایج روش‌ها نشان می‌دهد که خوشه‌بندی به روش‌افزایی میانه محور نسبت به سایر روش‌ها برتری دارد. آزمون‌های مناسب مانند عرض سیلهوت، همگنی و هم‌نوایی گشتاورهای خطی نشان از برتری این روش دارد. خوشه‌بندی با روش میانه محور منجر به تقسیم ایران به سه ناحیه به نام مناطق ساحلی (دریای مازندران با بارش زیاد، خلیج فارس و دریای عمان با بارش کم)، نواحی کوهستانی (البرز، زاگرس، کپه‌داغ و نواحی مجاور آنها) و نواحی بیابانی و نیمه بیابانی (نواحی داخلی کشور ایران) تقسیم شده‌اند (جدول ۱). ارتفاع ایستگاه بهترین معیار برای تشخیص خوشه‌هاست. عرض سیلهوت و آماره‌های ناهمگنی (H1) برای هر سه خوشه در جدول (۲) آمده است که نشان از همگنی سه خوشه دارد. آماره ناهم‌نوایی (Di) را برای این خوشه نشان نیز محاسبه شد. چهار ایستگاه در سه خوشه ناهم‌نوایی تشخیص داده شد. ایستگاه تبریز در خوشه کوهستانی ناهم‌نوایی با اندازه ناهم‌نوایی آن ۴/۴۶۸ است که از این خوشه حذف شد (شکل ۱). خوشه کوهستانی برای تحلیل انتخاب و الگوهای ماهانه بر داده‌های آن برآزش داده شد. موارد تحلیل خوشه کوهستانی به شرح زیر است.

۱۳۸۷).  $RH_i = f(i)$ . رطوبت نسبی ساعت‌آم و عدد ۲۴ طول شبانروز (دامنه تابع) است.

(۱۲)

$$RH_{ave} \cong \frac{1}{24} \{RH_0 + 4(RH_3 + RH_9 + RH_{15} + RH_{21}) + 2(RH_6 + RH_{12} + RH_{18}) + RH_{24}\}$$

### الگوهای رگرسیون چندمتغیری

تحلیل رگرسیونی یک روش آماری برای پیش‌بینی مقادیر یک یا چند متغیر پاسخ (وابسته) از مجموعه‌ای از مقادیر متغیرهای پیش‌بینی کننده (مستقل) است. رگرسیون خطی برآورد و نتایج استنباط پارامترهای  $\beta = [\beta_1, \beta_2, \dots, \beta_p]^T$  را در الگوی (۱۳) فراهم می‌کند.

$$Y_n = \beta_1 x_{n1} + \beta_2 x_{n2} + \dots + \beta_p x_{np} + Z_n = (x_{n1}, \dots, x_{np}) \beta + Z_n \quad (13)$$

$Z_n$  در این الگو متغیر تصادفی،  $Y_n$  نمایش پاسخ در حالت  $n$  با  $n = 1, 2, \dots, N$  است که از دو قسمت قطعی و تصادفی (خطا) تشکیل می‌شود. مولفه قطعی (یعنی  $\beta(x_{n1}, \dots, x_{np})$ ) به پارامتر  $\beta$  و متغیرهای مستقل (پیش‌گو)  $x_{np}$ ،  $p = 1, 2, \dots, P$ ، بستگی دارد. مولفه تصادفی یا خطا، اختلالی است که انحراف پاسخ را در آن حالت بیان می‌کند. بالانویس T نشان دهنده ترانواده ماتریس است. صورت ماتریسی رگرسیون خطی چند متغیره مطابق رابطه (۱۴) است که معمولاً از نماد ماتریسی استفاده می‌شود.

$$Y = X\beta + Z \quad (14)$$

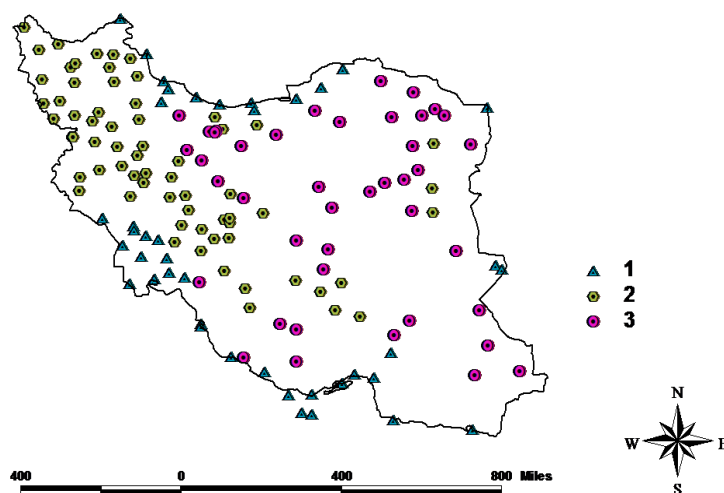
فرض‌های پایه زیر برای برآزش الگوی رگرسیونی پذیرفته می‌شود. معدل خطاها صفر است ( $E[Z] = 0$ )، واریانس خطاها ثابت است ( $Var(Z) = \sigma^2 I$ ). خطاها دو به دو مستقل‌اند. متغیر پاسخ به صورت نرمال توزیع شده است. بنابراین باید برقراری این فرض‌ها را در هنگام تحلیل داده‌ها آزمون کرد. افزون بر این آزمون همخطی و داده پرت

جدول شماره ۱- اطلاعات کلی سه خوشه انتخابی ایران

تفکیک	قطر	متوسط فاصله	بیشترین فاصله	حجم	
۱۸۴/۸۶	۲۰۷۳/۱	۵۰۳/۴۷	۱۷۴۶/۱	۴۱	خوشه اول
۷۲/۹۸	۱۶۲۶/۲	۳۲۸/۵۶	۱۲۵۵/۶۳	۶۰	خوشه دوم
۷۲/۹۸	۹۴۷/۳۶	۲۹۱/۴	۶۲۸/۶۶	۴۸	خوشه سوم
خوشه اول	خوشه دوم	خوشه سوم	شماره خوشه		
۰/۳۸	۰/۳۶	۰/۴۰	میانگین عرض سیلهوت هر خوشه		
۰/۳۸			میانگین عرض سیلهوت همه مجموعه داده‌ها		

جدول شماره ۲- معیارهای همگنی هاسکینگ-والیس برای سه خوشه پذیرفته شده

$V_1$	$H_1$	شماره خوشه
۰/۰۶۸	-۰/۰۹۲	خوشه اول
۰/۰۷۰	-۰/۱۳۳	خوشه دوم
۰/۰۶۰	۰/۰۴۲	خوشه سوم



شکل شماره ۱- نمایش نموداری نقطه‌ای خوشه بندی ایران به سه ناحیه. خوشه کوهستانی با شش ضلعی مشخص شده

سایر موارد در غربال داده‌ها انجام شده‌است. این عمل منجر به حذف تعدادی از داده‌ها شد. تعداد داده‌های مشکوک حذف شده در خوشه کوهستانی حدود ۱۵٪ کل داده‌هاست. تعداد ۱۸۴۸۷ روز- داده برای تحلیل الگوی سالانه به دست آمد. داده‌های قابل قبول برای هر ماه تفکیک شدند. تقریباً حدود ۱۵۵۰ داده-روز برای هر ماه در تحلیل شرکت داده شدند.

- تحلیل رگرسیون: رگرسیون خطی چند متغیره برای الگوبندی انتخاب شده است. هدف به دست آوردن رابطه

- نمونه‌گیری: روش سیستماتیک دوری در این تحقیق با توجه به حجم نمونه و حجم جامعه به کمک بسته‌های نرم‌افزاری R صورت گرفته است. برای اینکار ناچار بوده‌ایم که بسته‌های نرم‌افزاری را نیز تا نبل به هدف دلخواه تغییر دهیم. تعداد ۵ ایستگاه به روش نمونه‌گیری سیستماتیک دوری برای خوشه کوهستانی انتخاب و تحلیل رگرسیون روی داده‌های ساعتی این پنج ایستگاه انجام شده است.

- آماده‌سازی و غربال داده‌ها: اولین گام برای انجام هر تحلیل است. بررسی صحت، دقت، داده‌های مشکوک و

ندارد. معیار، باقی مانده های حذف شده استیودنتایز و دامنه این باقی مانده ها بازه (۲/۲۴، ۲/۰۹) است. فاصله کوک در دامنه مناسبی قرار دارد. مقادیر اهرم گون نیز نشان از رفتار خوب باقی مانده هاست. شکل (۳) نمودار مستطیلی و چندک- چندک باقی مانده هاست که نشان از پیروی باقی مانده ها از توزیع نرمال دارد. شکل (۴) نمودار پراکنش باقی مانده های حذف شده استیودنتایز در مقابل متغیر پاسخ است. این نمودار نشان می دهد که پراکنندگی نقاط در یک نوار مستطیلی افقی قرار دارد. بنابراین تثبیت واریانس و نیاز به تبدیل در الگو احساس نمی شود. الگوی نهائی رطوبت نسبی روزانه برای ماه سپتامبر (جدول ۴) مطابق رابطه (۱۵) است.

$$RH_T = -10.54 + 0.29RH_3 + 0.26RH_9 + 0.34RH_{15} - 0.042T_{max} + 5.03Ln(RH_T) \quad (15)$$

موارد زیر از این الگو استنباط می شود.

- ۱- نسبت ضریب  $RH_{15}$  به ضرایب  $RH_3$  و  $RH_9$  به ترتیب برابر  $1/4$  و  $1/84$  است که اهمیت بیشتر  $RH_{15}$  را در الگو نشان می دهد. این اهمیت در کلیه الگوهای ماهانه و سالانه برقرار است.
- ۲- الگوی ارائه شده توسط دانشگاه آرگون (رابطه ۲) تاکید بر اهمیت بیشتر رطوبت نسبی صبح (ساعت ۳) دارد که این مطلب با نتایج رابطه (۱۵) در تناقض است.
- ۳- الگوهای رابطه (۱) و (۲) بدون عرض از مبدا هستند. یعنی رطوبت نسبی روزانه می تواند صفر شود. در حالی که الگوی (۱۵) این مطلب را قویا رد می کند. این موضوع از نظر فیزیکی نیز تایید می شود. زیرا هیچگاه رطوبت نسبی روزانه در مناطق کوهستانی صفر نمی شود. نتیجه نهایی ارائه ۱۳ الگوی برآورد رطوبت نسبی روزانه برای خوشه است که ۱۲ الگوی ماهانه و یک الگوی سالانه در هر خوشه پیشنهاد شده است.

- ۴- **واسنجی و مقایسه الگوها:** الگوهای مرسوم قدیمی (رابطه ۱)، الگوی آرگون (رابطه ۲) و الگوهای ارائه شده در این مقاله مقایسه و واسنجی شدند. داده های جدول (۶) این مقایسه عملی را برای کلیه ماه ها و الگوی سالانه نشان می دهد. میانگین مربع خطا (MSE)

رگرسیون بین متوسط واقعی رطوبت نسبی روزانه ( $RH_T$ ) به عنوان متغیر پاسخ و عوامل دیگر به عنوان متغیرهای پیشگوست. این متغیرها به ترتیب عبارت اند از رطوبت نسبی ساعات ۳، ۹، ۱۵ گرینویچ و متوسط رطوبت نسبی روز قبل که به ترتیب با نمادهای  $RH_9$ ،  $RH_3$ ،  $RH_{15}$  و  $RH_T$  نشان داده شده اند. همچنین چهار متغیر دمای حداقل، حداکثر، متوسط روزانه (سلسیوس) و بارش روزانه به میلی متر (با نمادهای  $T_{min}$ ،  $T_{max}$ ،  $T_{mean}$  و RAIN) نیز در این تحلیل استفاده شده اند. متوسط رطوبت نسبی واقعی روزانه با انتگرال گیری سیمسون و از روی داده های ساعتی مطابق رابطه (۴) به درصد حساب شده است.

الگوهای مختلف رگرسیونی خطی و غیر خطی (خطی پذیر) با نرم افزار Spss.18 به روش گام به گام برداده های ماهانه و سالانه برازش داده شد. نتایج نهایی سیزده الگوی برتر سالانه و ماهانه به همراه آماره های مهم در جدول (۴) آمده است.

داده های ماه سپتامبر برای شرح الگوی برتر به عنوان نمونه ارائه می شوند. نتایج تحلیل داده های این ماه در جداول (۵) تا (۸) و نمودارهای (۲) تا (۴) ارائه شده اند. شکل (۲) نمودار ماتریسی پراکنش متغیرهای ورودی در الگو و رفتار آن ها را نسبت به هم نشان می دهد. ضرایب تعیین و اصلاح شده (Adj-R2) در جدول (۵) آمده است ضریب تعیین اصلاح شده مربوط به الگوی پذیرفته شده مساوی  $0.997$  است که نشان دهنده قدرت خوب الگوست. زیرا فقط  $0.003$  تغییرات متغیر پاسخ توسط متغیرهای پیشگو بیان نمی شود. مقدار کم انحراف استاندارد برآورد نیز مؤید قوی بودن الگوست. مقادیر موجود در داخل جدول تحلیل واریانس (جدول ۶) نشان دهنده قدرت الگوست. ضرایب الگو در جدول (۷) به همراه خطای استاندارد آنها و مقادیر احتمال مربوط آمده است که نشان از حضور قوی متغیرهای پیشگوی شرکت کننده در الگو دارد. آماره های VIF (عامل تورم واریانس) عدم هم خطی متغیرهای پیشگو را نشان می دهند، زیرا کمتر از ۱۰ هستند. برقراری فرض های پایه در خصوص باقی مانده ها باید بررسی شوند. جدول (۸) رفتار باقی مانده ها را نشان می دهد. داده پرت در باقی مانده ها وجود



که MSE ماه‌های مختلف تقریباً نصف الگوی قدیمی (رابطه ۱) و بسیار کمتر از آرگون (رابطه ۲) است. جدول (۹) نشان می‌دهد که الگوهای جدول (۴) خطای کمتری از الگوهای (۱) و (۲) دارند. مثلاً در مقیاس سالانه MSE الگوهای (۱)، (۲) و الگوی سالانه در مقاله حاضر به ترتیب ۴/۱، ۲۲/۷ و ۲/۴ است که کمترین خطا مربوط به الگوی سالانه پیشنهادی است.

معیار کارایی الگوهاست. مقادیر MSE1، MSE2 و MSE3 به ترتیب میانگین مربع خطا برای الگوی قدیمی (رابطه ۱)، الگوی آرگون (رابطه ۲) و الگوهای مقاله حاضر (الگوهای جدول ۴) است. مقایسه و واسنجی این ۱۳ الگو با الگوهای مرسوم قدیمی (روابط ۱ و ۲) با معیار مجموع کمترین مربعات خطا (MSE) انجام شد. این مقایسه و واسنجی نشان داد که این ۱۳ الگو خطای کمتری ارائه می‌کنند (جدول ۹)، به طوری

جدول شماره ۳- نمونه‌ای از غربال داده‌ها برای رطوبت نسبی ساعات مختلف شبانه روز داده‌های مشکوک زمینه تیره دارند.

ماه	روز	RH0	RH3	RH6	RH9	RH12	RH15	RH18	RH21	RH24	Rain	Tmin	Tmax	Tmean	RH <sub>T</sub>	RH <sub>Y</sub>
3	20	68	15	18	15	73	86	100	100	94	1	6.4	26.6	16.5	58.7	72.9
4	4	95	92	90	84	86	20	93	95	98	0	8.6	12.5	10.5	78.8	85.5
8	23	87	7	74	57	64	77	88	92	87	0	23.2	33	28.1	64.6	71.5
7	18	87	81	76	59	45	6	85	83	87	0	22.6	30.8	26.7	62.8	72.5

جدول شماره ۴- آماره‌های اصلی الگوهای نهایی ماهانه و سالانه متوسط رطوبت نسبی روزانه در اقلیم کوهستانی

ماه	الگوی نهایی متوسط رطوبت نسبی روزانه	AdjR2	Std .Err	Durbin-Watson	VIF	F	P- valu
Jan	$RH_T = 6.09 + 0.29RH_3 + 0.22RH_9 + 0.41RH_{15} - 0.02T_{max} - 0.01RH_Y$	0.989	1.45	1.878	<3.7	3.16E4	.000
Feb	$RH_T = -3.97 + 0.3RH_3 + 0.22RH_9 + 0.4RH_{15} - 1.89T_{max} - 1.8Ln(T_{min}) + 0.285Ln(RH_Y) - 0.16Ln(RAIN)$	0.992	1.51	1.844	<4.7	2.55E4	.000
Mar	$RH_T = 12.17 + 0.3RH_3 + 0.23RH_9 + 0.39RH_{15} - 0.11T_{max} - 2.8Log(RH_Y) + 2.37Ln(T_{min}) - 0.29Log(RAIN)$	0.995	1.55	1.832	<4.3	4.119E4	.000
Apr	$RH_T = -6.66 + 0.27RH_3 + 0.24RH_9 + 0.4RH_{15} + 0.05T_{max} + 3.8Ln(RH_Y)$	0.993	1.60	1.856	<3.3	4.31E4	.000
May	$RH_T = -11.5 + 0.26RH_3 + 0.27RH_9 + 0.38RH_{15} - 5.07Ln(RH_Y)$	0.994	1.60	1.774	<5.3	6.39E4	.000
Jun	$RH_T = -15.3 + 0.27RH_3 + 0.28RH_9 + 0.36RH_{15} - 0.05T_{max} + 4.67Ln(RH_Y) + 1.56Log(T_{min})$	0.995	1.55	1.776	<9.2	5.59E4	.000
Jul	$RH_T = -13.41 + 0.27RH_3 + 0.31RH_9 + 0.33RH_{15} - 5.54Ln(RH_Y)$	0.995	1.51	1.738	<8.6	8.39E4	.000
Aug	$RH_T = -15.98 + 0.27RH_3 + 0.29RH_9 + 0.32RH_{15} + 6.48Ln(RH_Y)$	0.996	1.53	1.758	<8.8	9.23E4	.000

ادامه جدول شماره ۴- آماره های اصلی الگوهای نهایی ماهانه و سالانه متوسط رطوبت نسبی روزانه در اقلیم کوهستانی

ماه	الگوی نهایی متوسط رطوبت نسبی روزانه	AdjR2	Std .Err	Durbin-Watson	VIF	F	P- valu
Sep	$RH_T = -10.5 + 0.29RH_3 + 0.26RH_9 + 0.34RH_{15} - 0.04T_{max} + 5.03Ln(RH_Y)$	0.997	1.45	1.821	<9.7	9.77E4	.000
Oct	$RH_T = -6.34 + 0.28RH_3 + 0.23RH_9 + 0.38RH_{15} - 0.09T_{max} + 3.95Log(RH_Y)$	0.996	1.51	1.646	<6.6	7.77E4	.000
Nov	$RH_T = -4.12 + 0.28RH_3 + 0.24RH_9 + 0.4RH_{15} + 2.85Log(RH_Y)$	0.993	1.51	1.823	<3.4	5.39E4	.000
Dec	$RH_T = 2.15 + 0.28RH_3 + 0.24RH_9 + 0.4RH_{15} + 1.11Log(RH_Y)$	0.991	1.44	1.911	<3.1	4.26E4	.000
Annual	$RH_T = 10.94 + 0.29RH_3 + 0.24RH_9 + 0.37RH_{15} - 0.019T_{min} + 4.7Ln(RH_Y) + 0.26\sqrt{RAIN} - 0.044RAIN$	0.995	1.54	1.784	<5.9	5.64E5	.000

جدول شماره ۵- ضریب تعیین الگوی ماه سپتامبر

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	.969	.938	.938	6.25015	
2	.992	.985	.985	3.11355	
3	.998	.995	.995	1.72438	
4	.998	.997	.997	1.46241	
5	.998	.997	.997	1.45658	1.821

جدول شماره ۶- تجزیه واریانس الگوهای مختلف برای ماه سپتامبر

Model	Sum of Squares	df	Mean Square	F	Sig.	
1	Regression	975350.500	1	975350.500	2.497E4	.000 <sup>a</sup>
	Residual	64378.093	1648	39.064		
	Total	1039728.593	1649			
2	Regression	1023762.306	2	511881.153	5.280E4	.000 <sup>b</sup>
	Residual	15966.287	1647	9.694		
	Total	1039728.593	1649			
3	Regression	1034834.217	3	344944.739	1.160E5	.000 <sup>c</sup>
	Residual	4894.376	1646	2.973		
	Total	1039728.593	1649			
4	Regression	1036210.539	4	259052.635	1.211E5	.000 <sup>d</sup>
	Residual	3518.055	1645	2.139		
	Total	1039728.593	1649			
5	Regression	1036240.654	5	207248.131	9.768E4	.000 <sup>e</sup>
	Residual	3487.939	1644	2.122		
	Total	1039728.593	1649			

جدول شماره ۷- جدول ضرایب الگو، آماره t و اندازه احتمال الگوی ماه سپتامبر

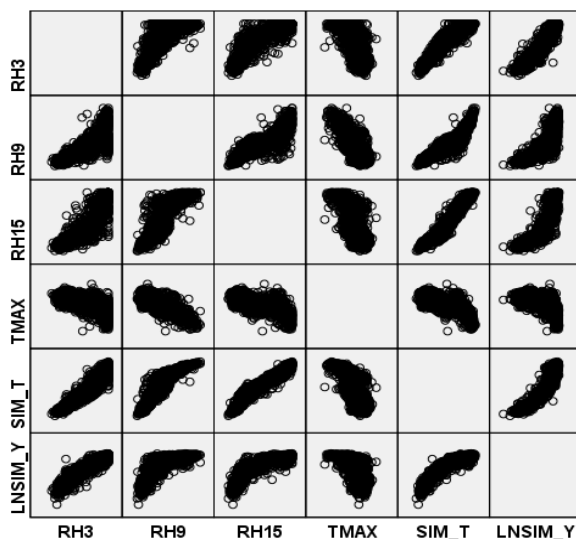
Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	14.223	.277		51.427	.000		
	RH15	.810	.005	.969	158.012	.000	1.000	1.000
2	(Constant)	.349	.240		1.453	.146		
	RH15	.481	.005	.575	90.673	.000	.232	4.319
	RH3	.448	.006	.448	70.668	.000	.232	4.319
3	(Constant)	.416	.133		3.135	.002		
	RH15	.354	.004	.423	97.901	.000	.154	6.514
	RH3	.381	.004	.382	103.904	.000	.211	4.731
	RH9	.273	.004	.237	61.021	.000	.190	5.274
4	(Constant)	-11.767	.493		-23.854	.000		
	RH15	.343	.003	.410	111.141	.000	.151	6.627
	RH3	.293	.005	.294	62.802	.000	.094	10.639
	RH9	.264	.004	.229	69.034	.000	.188	5.328
	LNRHY	4.947	.195	.113	25.368	.000	.104	9.598
5	(Constant)	-10.535	.590		-17.846	.000		
	RH15	.343	.003	.409	111.124	.000	.150	6.653
	RH3	.291	.005	.292	62.163	.000	.093	10.790
	RH9	.257	.004	.222	60.533	.000	.151	6.617
	LNRHY	5.028	.195	.115	25.729	.000	.103	9.715
	TMAX	-.042	.011	-.009	-3.768	.000	.360	2.777

جدول شماره ۸- مشخصات باقیمانده‌های الگوی ماه سپتامبر

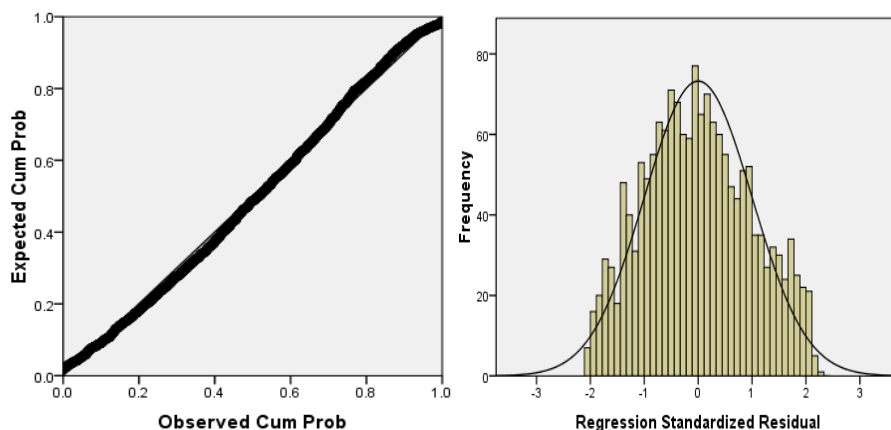
	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	5.6199	100.7188	50.5358	25.06802	1650
Std. Predicted Value	-1.792	2.002	.000	1.000	1650
Standard Error of Predicted Value	.040	.235	.084	.024	1650
Adjusted Predicted Value	5.6075	100.7308	50.5356	25.06827	1650
Residual	-3.02810	3.25271	.00000	1.45437	1650
Std. Residual	-2.079	2.233	.000	.998	1650
Stud. Residual	-2.085	2.235	.000	1.000	1650
Deleted Residual	-3.04479	3.25895	.00025	1.45976	1650
Stud. Deleted Residual	-2.087	2.238	.000	1.001	1650
Mahal. Distance	.249	42.016	4.997	3.967	1650
Cook's Distance	.000	.011	.001	.001	1650
Centered Leverage Value	.000	.025	.003	.002	1650

جدول شماره ۹- مقایسه معدل رطوبت نسبی با معادلات (۱) و (۲) و الگوهای جدول (۴) با معیار MSE

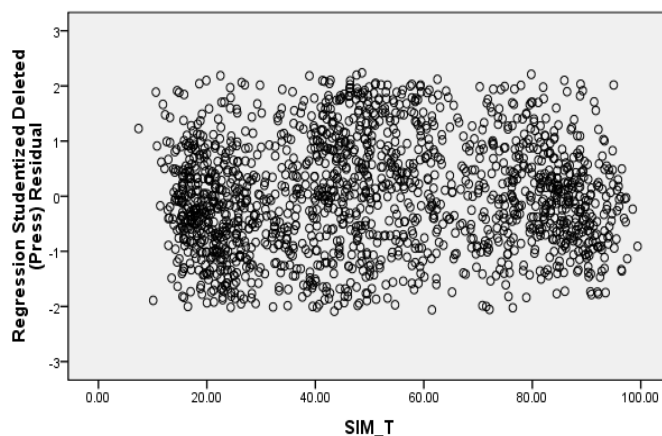
ماه	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Annual
MSE <sub>۱</sub>	۴/۵	۴/۸	۴/۲	۵/۱	۴/۲	۳/۶	۳/۶	۳/۹	۳/۷	۳/۹	۴/۳	۴/۴	۴/۱
MSE <sub>۲</sub>	۲۴/۷	۲۷	۳۰/۶	۳۴/۲	۲۴/۵	۱۷/۸	۱۶/۲	۲۱/۴	۱۵/۹	۲۱/۴	۲۶/۳	۲۳/۸	۲۲/۷
MSE <sub>۳</sub>	۲/۱	۲/۳	۲/۵	۲/۶	۲/۵	۲/۴	۲/۳	۲/۳	۲/۱	۲/۳	۲/۳	۲/۱	۲/۴



شکل شماره ۲- نمودار پراکنش ماتریسی رفتار متغیرهای ورودی الگو در برابر یکدیگر برای ماه سپتامبر. SIM-Y و SIM-T به ترتیب رطوبت نسبی واقعی روزانه و روز قبل است.



شکل شماره ۳- نمودار مستطیلی (راست) و چندکی (چپ) باقی مانده‌های استاندارد شده ماه سپتامبر



شکل شماره ۴- نمودار پراکنش باقی‌مانده‌های استاندارد شده شده ماه سپتامبر. SIM-T رطوبت نسبی واقعی روزانه است

### نتیجه‌گیری

خوشه‌بندی ۱۴۹ ایستگاه همدید ایران با شش متغیر دما، بارش، تبخیر، رطوبت نسبی و دامنه تغییرات دمای سالانه، ارتفاع ایستگاه و سه نمایه اقلیم‌بندی دومارتن، ایوانف و تورنت وایت با روش افزایی میانه محور انجام شده است. آزمون‌های مناسب مانند عرض سیلهوت، همگنی و هم‌نوایی گشتاورهای خطی برای قبول خوشه بندی به کار رفته است. ایران به سه خوشه تقسیم شد (شکل ۱). اطلاعات کلی سه خوشه در جدول (۱) آمده است. آماره‌های ناهمگنی (HI) برای هر سه خوشه در جدول (۲) آمده است که نشان از همگنی سه خوشه دارد.

مقاله حاضر خوشه کوهستانی (خوشه دوم) را انتخاب و تحلیل کرده است. شکل (۱) پراکنندگی این خوشه را نشان می‌دهد. میانگین عرض سیلهوت خوشه دوم ۰/۳۶ است. آماره ناهم‌نوایی (Di) برای خوشه دوم نیز محاسبه شد. این آماره نشان داد که ایستگاه تبریز با اندازه ناهم‌نوایی ۴/۴۶۸ یک ایستگاه ناهم‌نواست و حذف شد. غربال نشان داد که ۱۵٪ از داده‌های خوشه کوهستانی مشکوک و نادرست‌اند (جدول ۳). پنج ایستگاه همدید اردبیل، ارومیه، شهرضا، لردگان و قروه در خوشه دوم پس از محاسبه حجم نمونه مورد نیاز به روش نمونه‌گیری سیستماتیک دوری به منظور الگوبندی (تحلیل رگرسیونی) انتخاب شد. الگوها در مقیاس ماه و سال تبیین بر داده‌ها برازش داده شد (جدول ۴). نرم‌افزار Spss.18 به روش گام به گام الگوها را تبیین

نمود. ضرایب تعیین و اصلاح شده (Adj-R2) در جدول (۴) آمده است. مقادیر بالای ضریب تعیین تصحیح شده (Adj-R2) و آماره F در جداول تحلیل واریانس قدرت بالای الگوها را نشان می‌دهد. کمترین ضریب تعیین این ۱۳ الگو (۰/۹۸۹) مربوط به ماه ژانویه و بیشترین آن (۰/۹۹۷) مربوط به ماه سپتامبر است. کمترین مقدار آماره F (۲۵۰۰) مربوط به ماه فوریه و بیشترین آن (۵۶۴۰۰۰) مربوط به الگوی سالانه است. مقدار کم انحراف استاندارد آماره F نیز مؤید قوی بودن الگوهاست. ضرایب‌الگو به‌همراه خطای استاندارد آنها و مقادیر احتمال مربوط نشان از حضور قوی متغیرهای پیشگوی شرکت کننده در الگو دارد. آماره‌های VIF عدم هم‌خطی متغیرهای پیشگو را نشان می‌دهند. مقدار آن برای همه الگوها کمتر ۹/۷ است (جدول ۴). تبدیلاتی روی بعضی از متغیرهای پیشگو در بعضی از ماه‌ها قدرت الگو را افزایش داد. جدول (۴) نشان می‌دهد رطوبت نسبی روزانه در اکثر ماه‌ها به لگاریتم رطوبت نسبی روز قبل (RHY) بستگی دارد. مقایسه و واسنجی الگوهای ارائه شده در این مقاله با الگوهای مرسوم قدیمی و آرگون (رابطه ۱ و ۲) با معیار مقایسه MSE نشان از دقت بیشتر الگوهای ارائه شده در این مقاله دارد. به طوری که MSE الگوهای سیزده‌گانه این تحقیق تقریباً نصف الگوی قدیمی (رابطه ۱) و بسیار کمتر از آرگون (رابطه ۲) است. همچنین الگوهای جدول (۴) وجود عرض از مبدا را در الگو تایید می‌کند که از نظر فیزیکی نیز معقول است. زیرا رطوبت نسبی صفر در

11. Everitt, B. S., Landau S., Leese, M., Stahl, D., 2011, Cluster Analysis, 5th Edition, John Wiley and Sons.
12. Corvallis, Oregon, USA, 2008, United States Average Monthly or Annual Relative Humidity, (SCAS/OSU).
13. Court, A. And Waco, D. 1956. Means and midranges of relative humidity. Monthly weather review vol. 93, 8: 517-522.
14. Day, P. C. 1917, Relative Humidity and Vapor Pressures over the United States, including a Discussion of Data from Recording University of Chicago Press, 375 pp. (ref. on page 242).
15. Freiwan, M. and Kadioglu, M. 2008. Spatial and temporal analysis of climatological data in Jourdon, International Journal of Climatology, 28, 521- 535.
16. Goossens, C. 1985. Principal component analysis of Mediterranean rainfall, International Journal of Climatology, 5:379-388.
17. Holomboe, J., Forsythe, G. E., and Gustin, W. 1945. Dynamic Meteorology, John Wiley & Sons, Inc., New York.
18. Hosking, J. R. M, Wallis, J.R., 1997. Regional frequency analysis: An approach based on L-moments. Cambridge University Press, New York.
19. Jackson, I. and Weinand, H. 1995. Classification of tropical rainfall stations: A comparison of clustering techniques. International Journal of Climatology, 15: 985-994.
20. Linarie, E., 1991, Climate Data and Resources: A Reference and Guide.
21. Rao, A. R. and Srinivas, V.V. 2008. Regionalization of watersheds: An approach based on cluster analysis, Springer Science.
22. Romesburg, H., C. 2004. Cluster Analysis for Researchers. Lifetime Learning Publications, Belmont, C. A.
23. Theodoridis, S., Koutroumbas, K., 2009, Pattern Recognition, 4nd edition, Academic Press, Elsevier Pub.
24. Yao, A. Y. M. 1974. A Statistical Model for Relative Humidity, Environmental Data Service, NOAA, Washington, D.C. 20235: 14-23.

مناطق کوهستانی عملاً امکان‌پذیر نیست. افزون بر این رطوبت نسبی ساعت ۱۵ بیشترین تاثیر را در برآورد رطوبت نسبی روزانه دارد. این موارد با الگوهای مرسوم و آرگون مغایرت دارد.

الگوهای جدول (۴) با توجه به موارد ذکر شده برای مناطق کوهستانی ایران پیشنهاد می‌شود. خطاهای برآورد رطوبت نسبی را می‌توان با به کارگیری این الگوها به جای الگوهای قدیمی و مرسوم کاهش داد.

### منابع

1. Bakhtiyari. S., 2004, Geographical Atlas of the provinces in Iran, Geographical & Cartographic Institute.
2. www.irimo.ir
3. Toutounian, F., 1992, Numerical Methods for computer Science Engineering and Mathematics. Khorasan Pub, pp736.
4. Habiby, M. Rezaee-Pazhand H., and Farzandi M., 2008, Mean Daily Temperature Pattern for Arid and Semi-arid Zones in Iran, Iran-Water Resources Research, Volume 4, No. 1, Spring 2008 (IR-WRR).
5. Rezaee-Pazhand, H., Bozorgnia, A., 2002, Nonlinear Regression Analysis with application, Mashhad University Pub, pp400.
6. Meteorological yearbook, 2006, Iran Meteorological ORG.
7. Masoodian, S. A. and Ataei, H., 2005, A Cluster Analysis of Precipitation Seasons of Iran, Research Journal of Esfahan University.
8. Niroumand, H., 2005, Regression Analysis by Example. Mashhad University Pub, pp384.
9. Blanc, M. L. 1961, A Comparison of Methods for Computing Daily Mean Values of 66 Dry Bulb Temperatures, Dew Point, and Relative Humidity, Monthly Weather Review, vol. 89, 10: 401-410.
10. Elgar, E., 2008, Handbook of Research on Cluster Theory, Edited by Charlie Karlsson.