

## مدل سازی عملکرد گندم با استفاده از شاخص های اقلیمی و خوشه بندی سلسله مراتبی

لاله پرویز<sup>۱\*</sup>، حسین بنیادی<sup>۲</sup>

۱. استادیار، دانشکده کشاورزی، دانشگاه شهید مدنی آذربایجان، تبریز

۲. دانشجوی دانشکده کشاورزی، دانشگاه شهید مدنی آذربایجان، تبریز

(تاریخ دریافت ۱۳۹۷/۰۹/۰۳؛ تاریخ تصویب ۱۳۹۷/۱۲/۲۷)

## چکیده

بررسی شاخص های اقلیمی با تابعیت داده های هواشناسی از عوامل مؤثر بر تصمیم گیری های مرتبط با برنامه ریزی اقلیمی، کشاورزی، مهندسی منابع آب و به کارگیری راهبردهای مدیریتی است. هدف مقاله حاضر، توسعه مدل سازی عملکرد محصول - شاخص های اقلیمی با تأکید بر ورودی های مدل براساس تحلیل خوشه ای است. استخراج اطلاعات از خوشه های مربوط به شاخص های اقلیمی با محاسبه میانگین هر خوشه انجام گرفت. شاخص های بررسی شده شامل یازده شاخص اقلیمی؛ شاخص های لانگ، دومارتن، کوپن ۱، کوپن ۲، کوپن ۳، آنگستروم، ایوانف، سیلیانینوف، بارش مؤثر، پوشش گیاهی و خشکی بود. مدل سازی عملکرد گندم - شاخص های اقلیمی در استان های گیلان، اصفهان، کرمانشاه و آذربایجان غربی با استفاده از رگرسیون ساده و شبکه عصبی مصنوعی انجام گرفت. اصلاح شاخص های اقلیمی با ساختار صحیح به افزایش دقت در شبیه سازی عملکرد محصول منجر شد، برای مثال شاخص سازش در استان کرمانشاه از شاخص دومارتن به کوپن ۲، ۱۲/۸۲ درصد افزایش داشت. خطای شبیه سازی عملکرد محصول (RMSE) با استفاده از شاخص های اقلیمی نسبت به استفاده مستقیم از داده های هواشناسی در کل استان ها، ۳۶/۶۶ درصد کاهش داشت. استفاده از تحلیل خوشه ای در تعیین ورودی های هر دو مدل استفاده شده کارایی مدل را افزایش داد (متوسط RMSE در کل استان ها با خوشه بندی و بدون خوشه بندی به ترتیب برابر با ۱/۱۵-۰/۷ و متوسط RRMSE با خوشه بندی و بدون خوشه بندی به ترتیب برابر با ۰/۲۹-۰/۵ بود). بنابراین، ترکیب شاخص های اقلیمی به عنوان ورودی های مدل با انجام تحلیل مناسب به بهبود عملکرد مدل سازی منجر می شود.

کلیدواژگان: تحلیل خوشه ای، ترکیب، توسعه مدل سازی، عملکرد گندم.

## مقدمه

پیش‌بینی عملکرد محصول برای انجام برنامه‌ریزی، مدل‌سازی و اجرای سیاست‌های مربوط به تهیه و توزیع غذا و تصمیم‌گیری صادرات و واردات مهم است [۱]. در کشاورزی عملکرد محصول به شدت تحت تأثیر فاکتورهای زیادی (مانند ژنتیک، خصوصیات خاک، آبیاری) است، اما آب و هوا یک فاکتور مهم غیرقابل کنترل است که بر رشد محصولات تأثیر می‌گذارد [۲]. تعیین و تشخیص فاکتورهای مؤثر بر رشد و عملکرد محصول، امکان استخراج استراتژی واکنش را نسبت به تغییرات اقلیمی آینده فراهم می‌سازد [۳]. در این میان، آگاهی در مورد ارتباط دقیق بین آب و هوا به عنوان فاکتور مؤثر و عملکرد محصول می‌تواند پیامدهای بالقوه آثار تغییرات اقلیمی را بر امنیت غذایی بیان کند [۴]. ارتباط عملکردی بین فاکتورهای آب و هوایی مانند بارندگی - دما و عملکرد محصول از موضوعات مورد توجه بوده است، به طوری که تحقیقات در این زمینه از سال ۱۹۹۰ آغاز شده است [۴]. استفاده مستقیم پارامترهای هواشناسی مانند بارندگی، دما در بسیاری از تحقیقات بررسی تأثیر عوامل آب و هوایی بر عملکرد محصول پذیرفته نشد [۵ و ۶]. به همین دلیل شاخص خشکی در مدل‌های عملکرد محصول - آب و هوا پیشنهاد شد [۶] و استدلال تحقیق به این صورت بود که آب و هوا شامل اجزای زیادی است و محدود کردن آن به یک فاکتور مشکل است، چرا که آن‌ها به هم مرتبطند [۶]. بسیاری از تحقیقات اهمیت اطلاعات آب و هوایی را در توصیف عملکرد محصول بیان کرده‌اند [۲ و ۳]. به این منظور، از سه روش مهم مدل‌سازی تأثیر عوامل آب و هوایی بر عملکرد محصول استفاده شده است که عبارت‌اند از: مدل‌های بیوفیزیک یا مدل شبیه‌سازی محصول، مدل‌های رگرسیون تجربی و مدل‌های تابعی<sup>۱</sup> که ساده‌شده یا ترکیب دو مدل یادشده هستند [۲]. مدل‌های رگرسیونی نسبت به مدل‌های شبیه‌سازی محصول کمتر روندگرا<sup>۲</sup> هستند و به اطلاعات کمکی مانند خصوصیات خاک کمتر نیاز دارند. واسنجی آنها با یکسری اطلاعات اولیه انجام می‌گیرد. همچنین، مدل‌های داده‌محوری هستند.

رولا و همکارانش تأثیر تغییرات داده‌های هواشناسی بر عملکرد محصول گندم، ذرت و سویا را در منطقه پامپاس آرژانتین با توجه به تغییرات اقلیمی بررسی کردند. نتایج پژوهش آنها بیان‌کننده افزایش دمای بیشینه و کمینه فصلی در دوره‌های ۲۰۱۵-۲۰۳۹ و ۲۰۷۵-۲۰۹۹ و نیز افزایش بارندگی فصلی طی دوره ۲۰۱۵-۲۰۳۹ و نیز افزایش معنادار بارندگی در دوره ۲۰۷۵-۲۰۹۹ بود. همچنین، کاهش عملکرد گندم، افزایش عملکرد ذرت و افزایش معنادار عملکرد سویا، از نتایج مدل‌سازی آثار اقلیمی بر عملکرد محصول بود [۷]. وانگ و همکارانش به مدل‌سازی عملکرد ذرت با تغییرات اقلیمی در چهار منطقه آمریکا پرداختند. نتایج پژوهش آنها بیان‌کننده کاهش ۸/۱، ۱۷/۵، ۱۵/۲ و ۹/۷ درصد میزان عملکرد محصول در مناطق بررسی شده بود [۸]. همچنین، عملکرد شاخص‌های اقلیمی کشاورزی برای توسعه پیش‌بینی عملکرد محصول ذرت در آمریکا ارزیابی شد. دو تخمینگر مهم عملکرد شامل دما و شاخص SPEI<sup>۳</sup> بودند. در ساحل شرقی، مدل ترکیبی براساس آب و هوا بیان‌کننده ۳۲ درصد تغییرات عملکرد (با همبستگی ۰/۵۷) و در مناطق حساس نسبت به تغییرات آب و هوایی مانند ویرجینیا این عدد به ۶۴ درصد (با همبستگی ۰/۸) رسید. نتیجه نهایی تحقیق شامل توسعه مدل‌سازی عملکرد محصول با استفاده از شاخص‌های اقلیمی کشاورزی در مقایسه با اطلاعات ساده و مستقیم آب و هوا بود [۲]. همچنین، پژوهشگرانی در هند کارایی شاخص خشکی برای اندازه‌گیری تأثیر آب و هوا بر عملکرد محصول را بررسی کردند. مدل‌های استفاده‌شده شامل توان سه از متغیرهای زمان، شاخص خشکی، دما و بارندگی طی سال‌های ۱۹۵۰-۲۰۰۹ بر عملکرد محصول برنج در هند بود. نتایج پژوهش یادشده نشان داد شاخص آنگستروم نسبت به سایر شاخص‌های استفاده‌شده عملکرد بهتر دارد (RMSE=169.88). نتیجه نهایی شامل کاهش وابستگی عملکرد محصول به بارندگی بوده است که دلیل آن می‌تواند به دلیل پیشرفت در آبیاری و یا سایر امکانات باشد [۴]. در بررسی دیگری، تأثیر شاخص‌ها و متغیرهای هواشناسی کشاورزی بر عملکرد محصول گندم در هفت شهر استان همدان مطالعه

1. Functional  
2. Process-oriented

3. Standardized Precipitation-Evapotranspiration Index

معیارهای ارزیابی در برخی استان های کشور انجام گرفت. مدل های به کار گرفته شده شامل رگرسیون ساده و شبکه عصبی مصنوعی بودند.

### مواد و روش ها

#### منطقه مطالعه شده

اطلاعات هواشناسی مورد نیاز در تعیین شاخص های اقلیمی مربوط به داده های ایستگاه های هواشناسی همدیدی (سینوپتیک) استان های اصفهان، گیلان، آذربایجان غربی و کرمانشاه بود. داده های استفاده شده برای تعیین شاخص ها طی دوره ۱۹۹۲-۲۰۱۶ است که پنج سال انتهایی (۲۰۱۲-۲۰۱۶) به عنوان دوره صحت سنجی و محصول منتخب نیز گندم در نظر گرفته شد. هدف از انتخاب استان های مطالعه شده، تنوع اقلیمی است که براساس شاخص دومارتن اقلیم استان های آذربایجان غربی، کرمانشاه نیمه خشک و استان گیلان بسیار مرطوب و استان اصفهان خشک محاسبه شد. کمترین و بیشترین میزان بارندگی سالانه به ترتیب مربوط به استان اصفهان با ۱۷۹/۵۳ میلی متر و استان گیلان با ۱۲۰۸/۵۴ میلی متر است. بیشترین و کمترین دمای سالانه به ترتیب مربوط به استان اصفهان با ۱۷/۲۴ درجه سانتی گراد و استان آذربایجان غربی با ۱۲/۱۴ درجه سانتی گراد است. بیشترین و کمترین عملکرد گندم طی دوره آماری مربوط به استان اصفهان و گیلان بوده است. درخور یادآوری است داده های عملکرد محصول از آمارنامه جهاد کشاورزی گردآوری شدند.

#### شاخص های استفاده شده

شاخص های اقلیمی توابعی از داده های هواشناسی هستند. تأکید بیشتر شاخص های اقلیمی بر دو داده دما و بارندگی است، زیرا سایر عناصر هواشناسی به گونه ای در این دو پارامتر متجلی می شوند و نیازی به وارد کردن آنها در بررسی های اقلیمی نیست [۱۱]. ابتدا با استفاده از داده های دما و بارندگی، شاخص ساده ای پیشنهاد شد [۱۲]. دومارتن شاخص لانگ را با اعمال تغییر در مخرج شاخص، اصلاح کرد [۱۳]. آنگستروم با بیان وابستگی شاخص با طول مدت و مقدار بارش، شاخص دومارتن را توسعه داد [۱۴]. همچنین، کوپن شاخص دومارتن را به

شد. شاخص های بررسی شده شامل بارش سالانه، بارش فصلی، بارش ماهیانه، بارش ده روزه، کمبود بارش مرجع، کمبود بارش گیاهی، درجه-روز رشد در مقیاس ماهانه، و درجه-روز رشد در مقیاس ده روزه بود. از بین شاخص های مطالعه شده، شاخص کمبود بارش گیاهی همبستگی قوی تری در مقایسه با دیگر شاخص ها از خود نشان داد (با میانگین ۰/۳۶). به طور کلی، شاخص کمبود بارش، عملکرد محصول گندم را با دقت بیشتری پیش بینی می کند [۹]. همچنین، شاخص های خشکسالی و پارامترهای هواشناسی براساس تحلیل مؤلفه های اصلی در ایستگاه های بجنورد، مشهد و بیرجند برای مدل سازی عملکرد گندم و جو دیم استفاده شدند. شاخص های خشکسالی بررسی شده شامل بارش استاندارد شده، شاخص شناسایی خشکسالی بود. براساس نتایج، رابطه عملکرد محصولات با شاخص های خشکسالی قوی تر از رابطه عملکرد با هر یک از پارامترهای هواشناسی بوده و در هیچ منطقه ای پارامترهای دما، بارندگی در رتبه اول قرار نگرفته است. همچنین، شاخص شناسایی خشکسالی با در بر داشتن پارامتر تبخیر - تعرق دقت زیادی در برآورد عملکرد محصول دارد [۱۰].

امنیت غذایی و خودکفایی در محصولات استراتژیک کشاورزی از مهم ترین دغدغه های سیاست گذاران استانی است. بنابراین، با توجه به اهمیت گندم به عنوان محصولی استراتژیک، بررسی تغییرات میزان عملکرد آن اهمیت چشمگیری دارد و با توجه به وابستگی شدید فعالیت های کشاورزی به عوامل آب و هوایی، مدل سازی اقلیمی عملکرد محصول ضروری به نظر می رسد. با توجه به تحقیقات انجام شده، انتخاب صحیح ورودی مدل در زمینه مدل سازی عملکرد محصول با تأثیرات آب و هوایی تأثیر زیادی بر افزایش دقت برآورد عملکرد محصول دارد. بنابراین، هدف اصلی تحقیق حاضر، افزایش دقت شبیه سازی عملکرد محصول گندم با ترکیب شاخص های اقلیمی به عنوان ورودی مدل است. به این منظور، ترکیب شاخص های اقلیمی با استفاده از تحلیل خوشه ای انجام گرفت که در تحقیق حاضر استخراج اطلاعات خوشه ها با پارامتر آماری میانگین انجام شد. مقایسه شبیه سازی عملکرد گندم با تحلیل خوشه ای و عملکرد گندم با شاخص های اقلیمی و بدون انجام خوشه بندی با استفاده از

یادشده گروه‌بندی مشاهدات براساس فاصله بین آنها انجام می‌گیرد. یعنی مشاهدات دارای فاصله کم در یک گروه قرار می‌گیرند. در تحلیل خوشه‌ای از هندسه اقلیدسی برای تعیین فاصله اعضا از یکدیگر استفاده می‌شود. براساس فاصله اقلیدسی بین نقاط مکانی یا زمانی، ماتریس فاصله‌ها ایجاد می‌شود که در نهایت گروه‌های مکانی و زمانی براساس فاصله‌های این ماتریس به دست می‌آید [۲۱ و ۲۲]. روش‌های مختلفی برای تعیین فاصله خوشه‌ها وجود دارد که روش Ward یکی از آنهاست. این روش از تحلیل واریانس برای ارزیابی فواصل بین خوشه‌ها استفاده می‌کند. در روش Ward از بین تمام خوشه‌ها زوجی با هم ترکیب می‌شوند که مجموع مربعات انحرافات اعضای خوشه ادغامی آنها از میانگین مربوطه کمترین باشد. در روش Ward برای هر خوشه یک تابع خطا محاسبه می‌شود. این تابع خطا شامل متوسط فاصله هر نقطه در یک خوشه به مرکز ثقل در خوشه است. روش Ward بیان‌کننده فاصله بین دو خوشه A و B است که میزان افزایش مجموع مربعات به هنگام ترکیب آنها را نشان می‌دهد. معادله حاکم بر فاصله به صورت رابطه ۱ تعریف می‌شود.

$$D_w = \sum_{i \in A \cup B} \overrightarrow{x_i - m_{A \cup B}}^2 - \sum_{i \in A} \overrightarrow{x_i - m_A}^2 - \sum_{i \in B} \overrightarrow{x_i - m_B}^2 \quad (1)$$

$$= \frac{n_A n_B}{n_A + n_B} \overrightarrow{m_A - m_B}^2$$

$\overrightarrow{m_j}$  مرکز خوشه؛  $j$ : ام؛  $n_j$ : تعداد نقاط درون آن؛  $\| \cdot \|$  معیار فاصله بین نقاط.

در تحقیق حاضر، برای استخراج اطلاعات هر خوشه از پارامتر آماری میانگین استفاده شد که به‌کارگیری عملگر میانگین در استخراج اطلاعات خوشه‌های بارندگی ماهانه نیز نتایج شبیه‌سازی را توسعه داد [۲۳]. مدل‌سازی بین عملکرد محصول و شاخص‌های اقلیمی با استفاده از رگرسیون خطی و شبکه عصبی مصنوعی (شبکه پیش‌خور با الگوریتم پس‌انتشار خطا) انجام گرفت. شبکه عصبی الگویی برای پردازش اطلاعات براساس تفکر مغز انسان از طریق سلول‌های عصبی است. پردازش داده‌ها به عهده پردازنده‌های کوچکی به نام نورون است که به صورت شبکه‌ای به هم پیوسته و موازی یکدیگر رفتار می‌کنند تا یک مسئله را حل کنند [۲۳]. در نهایت، مراحل انجام تحقیق در شکل ۱ آورده شده است.

سه صورت تغییر و اصلاح کرد که داده‌های هواشناسی در هر سه صورت یکسان ولی ساختار ریاضی متفاوتی دارند [۱۵]. فرم ریاضی شاخص‌های استفاده‌شده تحقیق در جدول ۱ آورده شده است.

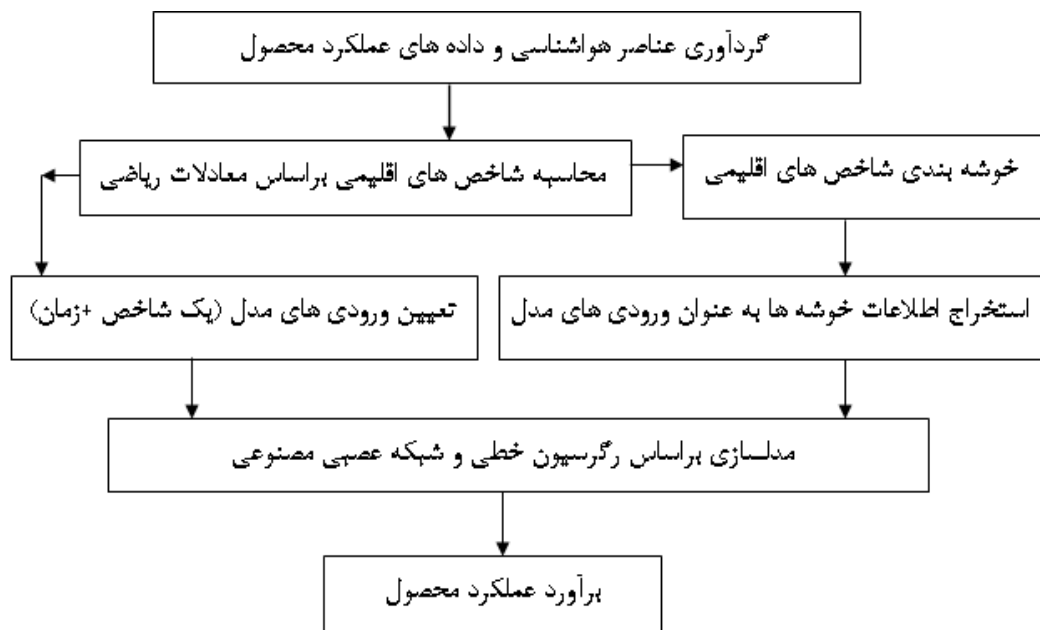
جدول ۱. معادلات ریاضی شاخص‌های استفاده‌شده

| Index                  | Equation                                         |
|------------------------|--------------------------------------------------|
| Lang index [12]        | $I = \frac{P}{\frac{1}{12} \sum T > 0}$          |
| De Martonne index [13] | $I = \frac{P}{T + 10}$                           |
| Koppen 1 index [15]    | $I = \frac{P}{5T + 120}$                         |
| Koppen 2 index [15]    | $I = \frac{P}{2T + 33}$                          |
| Koppen 3 index [15]    | $I = \frac{P}{T + 7}$                            |
| Angstrom index [14]    | $I = \frac{P}{1.07T}$                            |
| Ivanov index [16]      | $-E = 0.0018(2.5 + T^2)(100 - RH)$               |
| Selyaninov index [17]  | $I = \frac{P}{\frac{\sum E}{\sum P}}$            |
| PEI [18]               | $I = 115 \left( \frac{P}{T - 10} \right)^{1.11}$ |
| VCI [19]               | $I = \frac{P + RH}{T + 7}$                       |
| Aridity Index [20]     | $I = \frac{P}{ET_p}$                             |

$P$ : میزان بارندگی سالانه؛  $T$ : متوسط دما؛  $RH$ : متوسط رطوبت نسبی؛  $E$ : تبخیر ماهانه؛  $ET_p$ : تبخیر-تعرق پتانسیل.

### تحلیل خوشه‌ای

تحلیل خوشه‌ای از روش‌های کارآمد برای تخمین شباهت بین عناصر در یک گروه است. خوشه‌بندی در زیرمجموعه یادگیری بدون نظارت قرار دارد و فرایند خودکاری است که طی آن نمونه‌ها به دسته‌هایی که اعضای آن مشابه یکدیگر هستند، تقسیم شده که به این دسته‌ها خوشه می‌گویند. به‌طور ساده می‌توان گفت که فرایند تفکیک داده‌ها به زیرشاخه‌هایی با خصوصیات مشابه به نام خوشه انجام می‌شود. با استفاده از خوشه‌بندی مناسب، امکان جست‌وجوی ساده‌تر و دسترسی دقیق‌تری برای استفاده از اطلاعات امکان‌پذیر است. چندین الگوریتم برای خوشه‌بندی معرفی شده است که خوشه‌بندی سلسله‌مراتبی یکی از انواع آن است. در تحلیل خوشه‌ای مشاهدات هر خوشه بیشترین شباهت و مشاهدات خوشه‌های مختلف، کمترین شباهت را دارند. در روش



شکل ۱. روندنمای انجام تحقیق

ابتدا تغییرات دما، بارندگی و عملکرد محصول طی دوره آماری در شکل ۲ بررسی شد (در شکل ۲ «د» نماد دما، «ب» نماد بارندگی، «ع» نماد عملکرد محصول است).

بیشترین و کمترین مقدار عملکرد گندم طی دوره آماری مربوط به استان اصفهان و گیلان بوده است. بیشترین مقدار بارندگی و دما به ترتیب مربوط به استان گیلان و اصفهان بوده است. مقادیر شاخص های اقلیمی که براساس روابط جدول ۱ محاسبه شده اند، در شکل ۳ مشخص اند.

براساس شکل ۳، کمینه شاخص های ایوانف، سیلیانیونوف و خشکی در استان اصفهان و بیشینه در استان گیلان اتفاق افتاده است و با توجه به تغییرات داده های هواشناسی، تغییرات شاخص های اقلیمی قابل توجیه است. مدل سازی ابتدا مربوط به شاخص های اقلیمی و عملکرد محصول با استفاده از رگرسیون خطی بود. داده های ورودی در هر مدل شامل یکی از شاخص های بیان شده در جدول ۱ و پارامتر زمان بود. در کنار شاخص های بیان شده در مدل رگرسیون خطی، از دما-بارش-رطوبت نسبی به عنوان ورودی (پارامترهای هواشناسی به همراه عامل زمان) استفاده شد که در شکل ۴ مقایسه کارایی شاخص ها در برابر عملکرد محصول ترسیم شده است.

برای بررسی تأثیر نوع داده های ورودی در مدل های مختلف، معیارهای ارزیابی استفاده شد که عبارتند از: خطای ریشه متوسط مربعات ( $RMSE^1$ )، خطای ریشه متوسط مربعات نسبی ( $RRMSE^2$ ) و شاخص سازش<sup>۳</sup> که معادلات معیارها در روابط ۲-۴ آورده شده اند. مدلی با عملکرد دقیق مقادیر خطای کم و شاخص سازش زیاد است.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (O_i - S_i)^2} \quad (2)$$

$$RRMSE = \frac{\sqrt{\frac{1}{n} \sum_{i=1}^n (O_i - S_i)^2}}{\bar{O}} \quad (3)$$

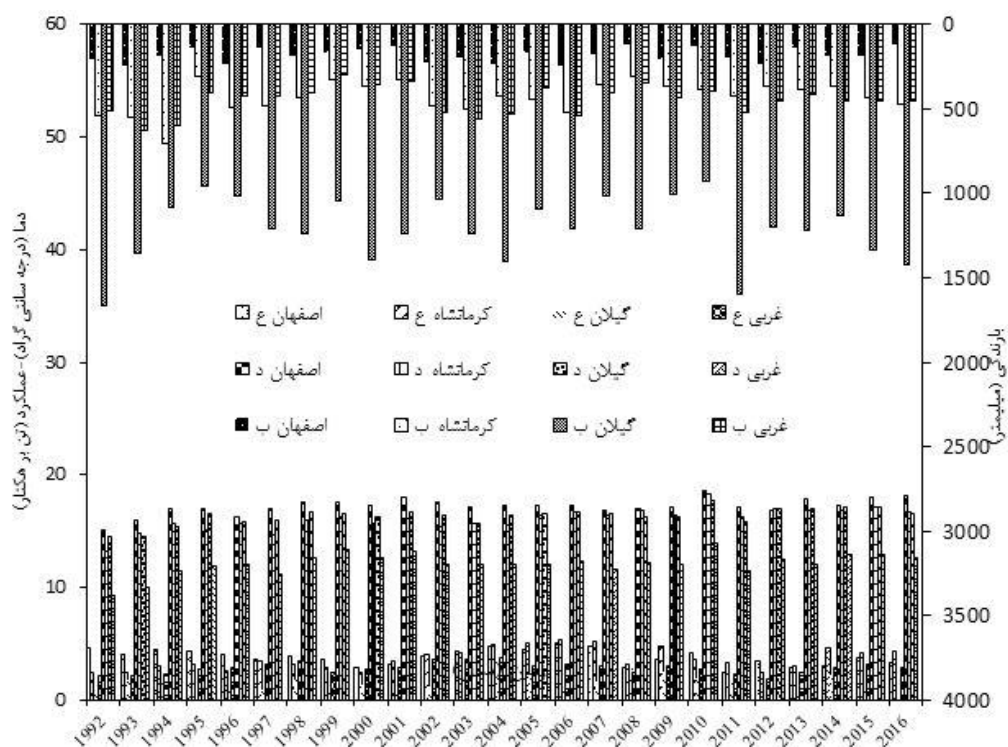
$$d = 1 - \frac{\sum_{i=1}^N (P_i - O_i)}{\sum_{i=1}^N (|P_i - \bar{O}| + |O_i - \bar{O}|)} \quad (4)$$

$O_i$ : مقادیر مشاهداتی،  $S_i$ : مقادیر شبیه سازی،  $n$ : تعداد کل داده ها،  $\bar{O}$ : متوسط مقادیر مشاهداتی.

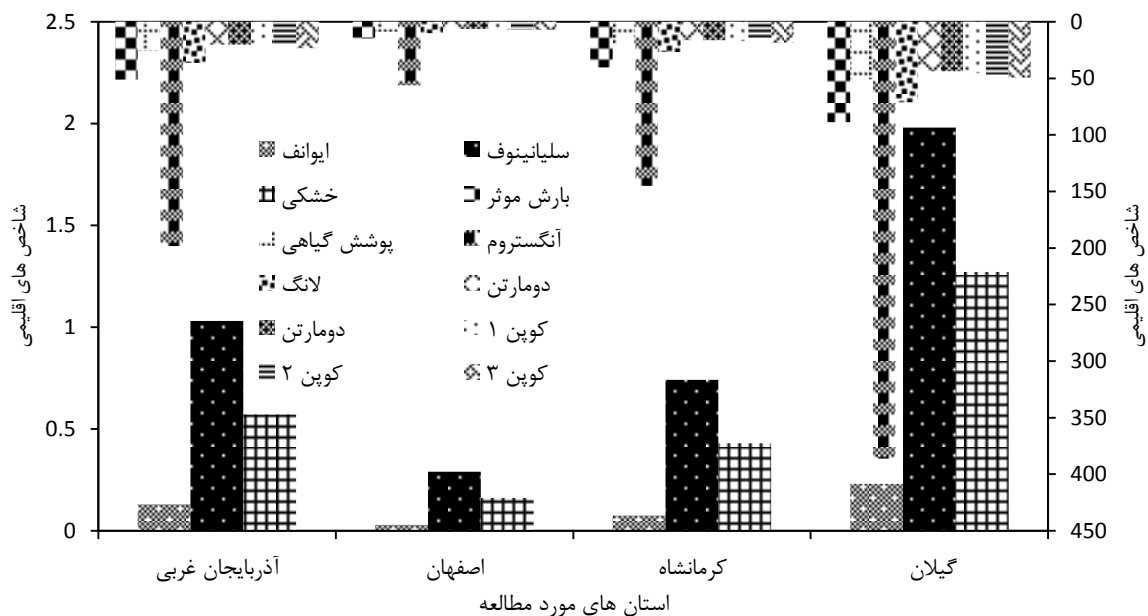
#### یافته ها

برای مدل سازی تأثیر عوامل آب و هوایی بر عملکرد گندم،

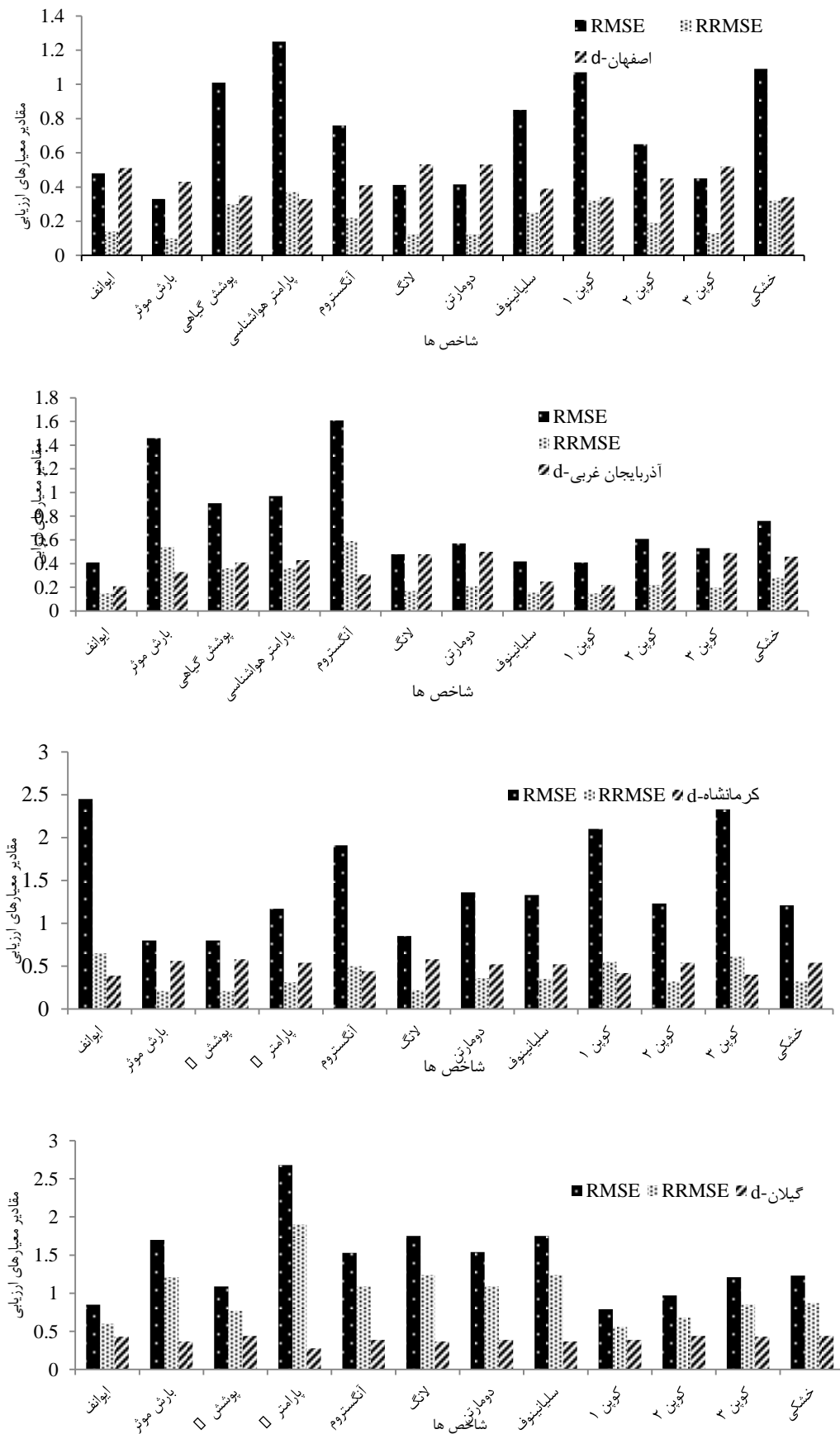
1. Root Mean Square Error
2. Relative Root Mean Square Error
3. Index of Agreement (d)



شکل ۲. نمودار دما، بارندگی و عملکرد محصول گندم طی دوره آماری



شکل ۳. مقادیر شاخص های اقلیمی طی دوره آماری



شکل ۴. مقایسه کارایی شاخص ها در برآورد عملکرد محصول گندم در استان های مختلف با روش رگرسیون خطی

سایر پارامترها را توجیه کنند. در ادامه، مقایسه‌ای بین مقادیر کمترین و بیشترین مقدار کمیته خطا در استان‌های مختلف انجام گرفت. کمترین مقدار کمیته خطا مربوط به استان اصفهان و بیشترین مقدار آن مربوط به استان گیلان بود. در حالت دیگر، به جای استفاده از یک شاخص اقلیمی در رگرسیون خطی، از ترکیب دوتایی شاخص‌های اقلیمی استفاده شد. انتخاب شاخص‌های اقلیمی در ترکیب دوتایی براساس نتایج بخش قبل یعنی کمیته معیار خطا در برآورد عملکرد محصول هر استان بود. نتایج برآورد عملکرد محصول در جدول ۲ آورده شده است.

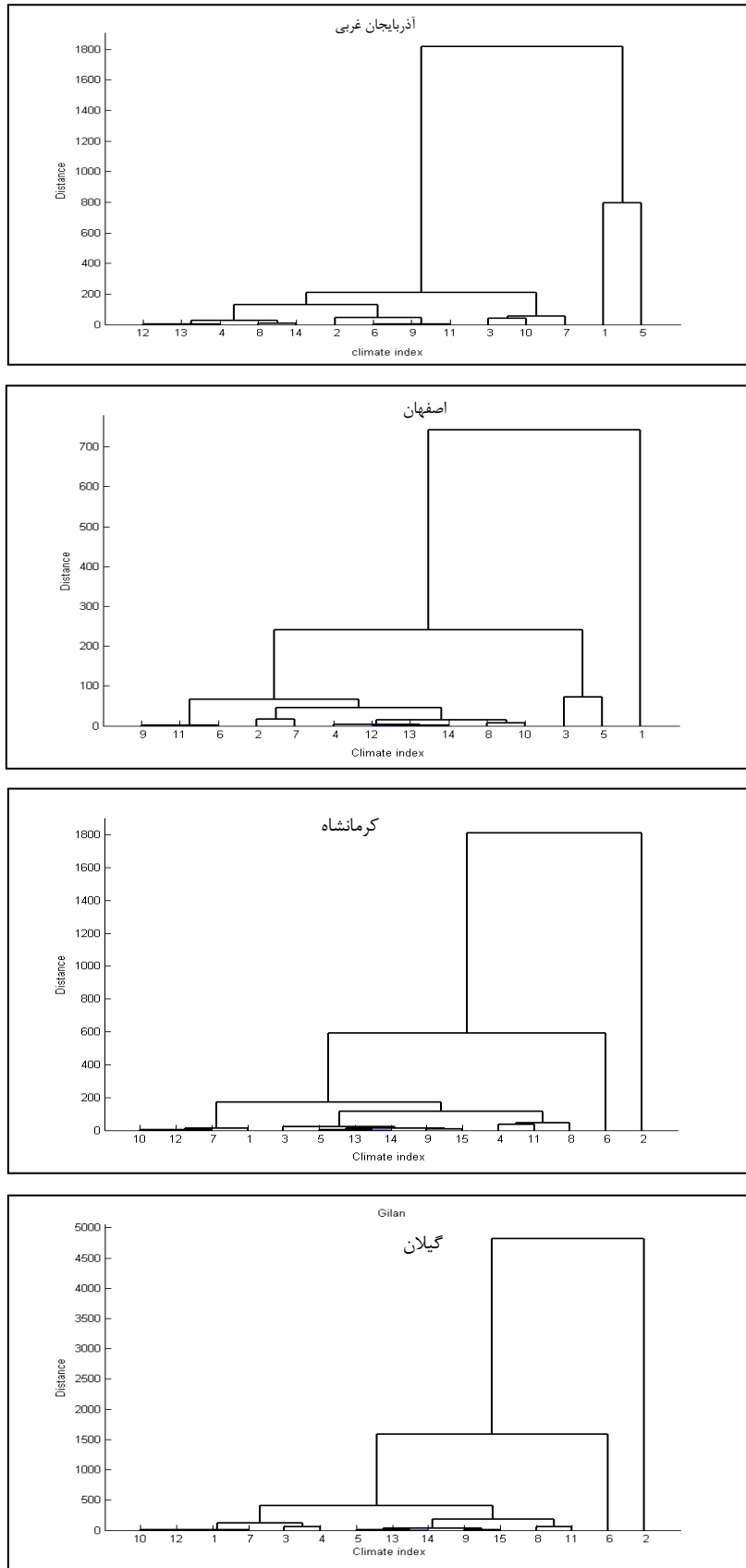
با توجه به جدول ۲، مقادیر خطا در حالت ترکیبی نسبت به حالت استفاده یک شاخص اقلیمی افزایش داشته است، بنابراین سعی در استفاده از روش دیگری برای ترکیب شاخص‌های اقلیمی شد. به این منظور، از تحلیل خوشه‌ای با روش Ward برای تعیین فاصله خوشه‌ها و فاصله اقلیدی استفاده شد. دلیل کاربرد روش Ward به دلیل راه‌حل‌های تفسیرپذیرتر و بامعناتری نسبت به سایر روش‌هاست [۲۶]. نتایج تحلیل خوشه‌ای شاخص‌ها و پارامترهای اقلیمی در شکل ۵ آورده شده است. درخور یادآوری است که در کنار شاخص‌های اقلیمی از داده‌های هواشناسی دما، بارش و رطوبت نسبی استفاده شد (شماره شاخص‌های اقلیمی در نمودار درختی: بارندگی سالانه=۱، متوسط دما=۲، رطوبت نسبی=۳، دومارتن=۴، آنگستروم=۵، ایوانف=۶، بارش مؤثر=۷، پوشش گیاهی=۸، خشکی=۹، لانگ=۱۰، سیلیانینوف=۱۱، کوپن=۱۲، کوپن ۲=۱۳، کوپن ۳=۱۴). در شکل ۵ محور افقی شاخص‌ها، داده‌های اقلیمی و محور قائم فاصله محاسبه‌شده روش خوشه‌بندی است.

براساس مقادیر معیارها، مقادیر بیشینه d و کمیته RMSE-RRMSE در استان‌های مختلف به این صورت است: در استان کرمانشاه شاخص‌های بارش مؤثر، پوشش گیاهی و لانگ در استان گیلان شاخص‌های کوپن ۱، ایوانف و کوپن ۲ در استان اصفهان شاخص بارش مؤثر، آنگستروم، کوپن ۱ و در استان آذربایجان غربی شاخص‌های ایوانف، سیلیانینوف، شاخص کوپن ۱ و لانگ. با توجه به مقادیر معیارها، خطای مربوط به داده هواشناسی مقادیر زیادی دارد به طوری که مقادیر RMSE در بیشتر استان‌ها از یک بیشتر است. بنابراین، استفاده مستقیم از داده‌ها در تخمین عملکرد محصول دقت کافی ندارد. در تحقیق دیگری نیز به این مورد اشاره شده است [۴]. برخی تحقیقات نیز به این اشاره داشتند که عملکرد محصولات با شاخص‌ها نسبت به داده‌های هواشناسی رابطه قوی‌تری دارند [۲-۱۰]. همچنین، شاخص‌های به‌دست‌آمده از داده‌های هواشناسی می‌توانند ارتباط بین آب و هوا و رشد محصول را به خوبی نشان دهند و در نهایت تصمیم‌گیری در کشاورزی را آسان کنند [۲۴ و ۲۵]. مقادیر معیارها در استان‌ها بیان می‌کند که شاخص‌های اصلاح‌شده نسبت به شاخص دومارتن عملکرد بهتری دارند. به طور کلی، ترکیبات مناسب دما و بارش در تمام استان‌ها دقت زیادی دارند. داده‌های هواشناسی مربوط به شاخص‌های اقلیمی با خطای کم، داده‌های دما و بارش هستند، ولی نوع شاخص‌های دارای خطای کم در هر استان فرق می‌کند و این اتفاق نشان می‌دهد علاوه بر پارامترهای دخیل در شاخص‌های اقلیمی، نوع فرمول‌بندی نیز اهمیت زیادی دارد. به‌طور کلی، با توجه به شاخص‌ها با کمیته خطا، داده‌های مؤثر دما و بارندگی هستند، چرا که این دو پارامتر به خوبی می‌توانند تغییرات

جدول ۲. مقادیر RMSE در مقایسه عملکرد شاخص‌های ترکیبی در برآورد عملکرد محصول با رگرسیون خطی

|                | Koppen 2 | Ivanov     | Koppen 1 | Koppen 2-Ivanov   | -Koppen 1 Ivanov |                     |             |
|----------------|----------|------------|----------|-------------------|------------------|---------------------|-------------|
| گیلان          | ۰/۹۷     | ۰/۸۵       | ۰/۷۹     | ۱/۶۴              | ۱/۴۳             |                     |             |
|                | Koppen 1 | Selyaninov | Lang     | Ivanov            | Lang-Koppen 1    | Selyaninov-Koppen 1 | Ivanov-Lang |
| آذربایجان غربی | ۰/۴۱۲    | ۰/۴۱۸      | ۰/۵۳     | ۰/۴۱              | ۰/۶۲             | ۱/۰۱                | ۱/۲۱        |
|                | Lang     | PEI        | VCI      | VCI-Lang          | PEI-Lang         | VCI-PEI             |             |
| کرمانشاه       | ۰/۸۵     | ۰/۸        | ۰/۸۱     | ۱/۵۶              | ۱/۱۳             | ۱/۸۷                |             |
|                | Koppen 1 | Angstrom   | PEI      | Angstrom-Koppen 1 | PEI-Koppen 1     | PEI-Angstrom        |             |
| اصفهان         | ۰/۴۵     | ۰/۴۱       | ۰/۳۴     | ۰/۴               | ۱/۲۱             | ۰/۷۲                |             |

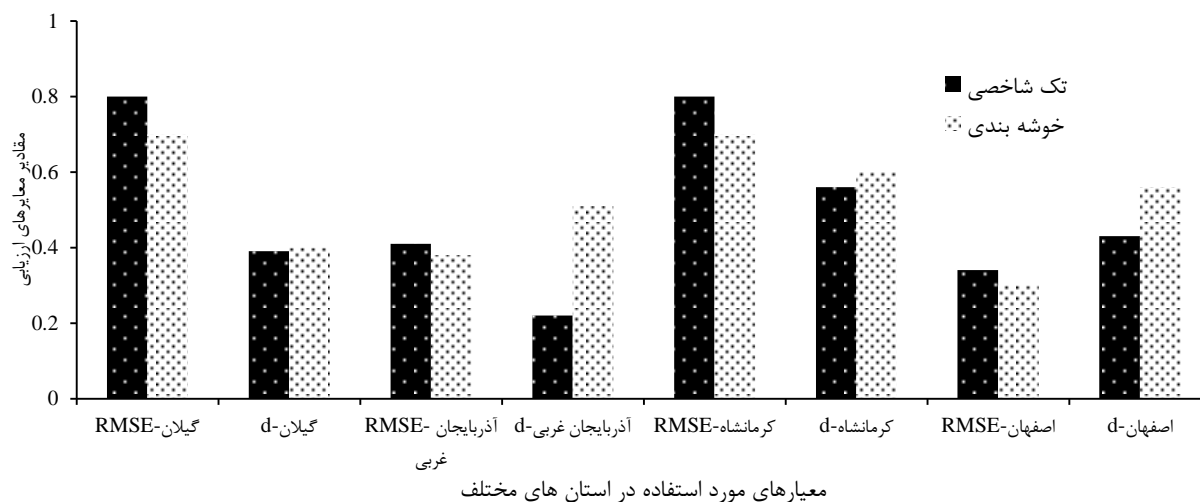




شکل ۵. نمودار درختی خوشه بندی شاخص های اقلیمی و داده های هواشناسی

سیلیانینوف، کوپن ۱، کوپن ۲، کوپن ۳، دومارتن، پوشش گیاهی. استان کرمانشاه شامل سه خوشه؛ خوشه ۱: متوسط بارندگی، خوشه ۲: آنگستروم خوشه ۳: متوسط دما، رطوبت نسبی، ایوانف، دومارتن، لانگ، پوشش گیاهی، بارش مؤثر، کوپن ۱، کوپن ۲، کوپن ۳. بعد از تعیین شاخص‌ها و داده‌های هر خوشه، اقدام به استخراج اطلاعات از هر خوشه شد، به طوری که از پارامتر آماری میانگین استفاده شد. بنابراین، در این حالت داده‌های ورودی مدل رگرسیون خطی شامل زمان و مقادیر داده‌های استخراجی از هر خوشه بود. مقایسه عملکرد محصول در حالت استفاده فقط از شاخص اقلیمی و خوشه‌بندی شاخص‌های اقلیمی در شکل ۶ آورده شده است.

بر اساس نمودار درختی ترسیم‌شده، نتایج خوشه‌بندی در استان‌ها به این صورت است: استان اصفهان شامل چهار خوشه؛ خوشه ۱: بارندگی سالانه خوشه ۲: رطوبت نسبی، آنگستروم خوشه ۳: ایوانف، سیلیانینوف، خشکی خوشه ۴: لانگ، پوشش گیاهی، کوپن ۱، کوپن ۲، کوپن ۳، دومارتن، متوسط دما، بارش مؤثر. استان گیلان شامل سه خوشه؛ خوشه ۱: بارندگی سالانه خوشه ۲: آنگستروم خوشه ۳: دما، رطوبت نسبی، ایوانف، بارش مؤثر، پوشش گیاهی، لانگ، دومارتن، سیلیانینوف، کوپن ۱، کوپن ۲، کوپن ۳، خشکی. آذربایجان غربی شامل سه خوشه؛ خوشه ۱: بارندگی سالانه، آنگستروم خوشه ۲: رطوبت نسبی، بارش مؤثر، لانگ خوشه ۳: متوسط دما، ایوانف، خشکی،



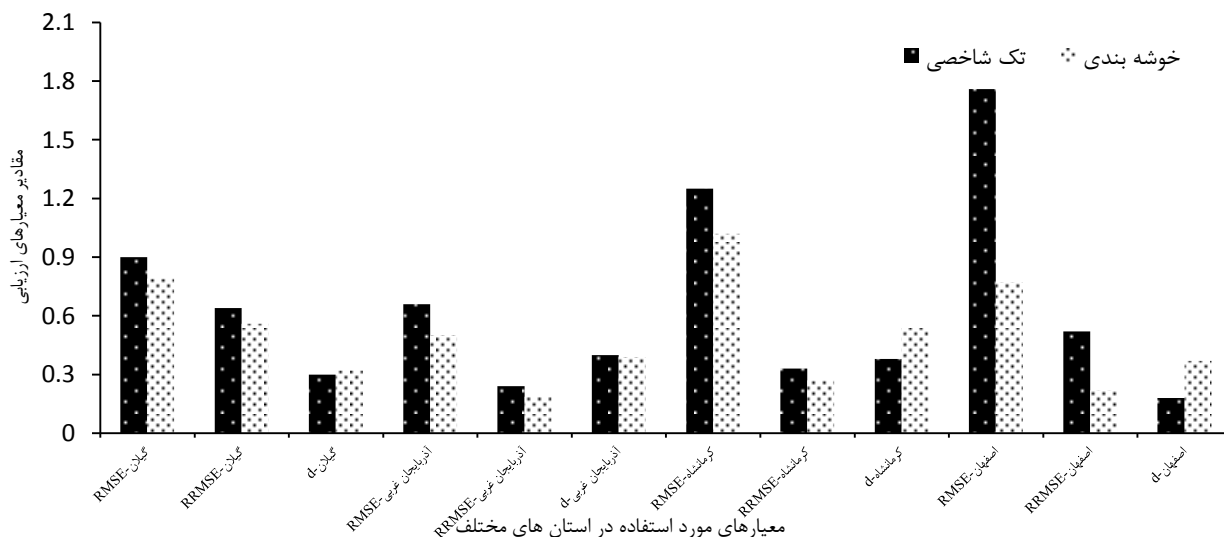
شکل ۶. ارزیابی عملکرد تحلیل خوشه‌ای و استفاده از تک‌شاخص اقلیمی در تخمین عملکرد محصول با رگرسیون خطی

کمینه مقدار خطا مربوط به استان اصفهان بود. در حالت دیگر نوع مدل‌سازی حاکم بر سیستم را تغییر داده و از مدل‌سازی شبکه عصبی مصنوعی استفاده شد. نوع شبکه استفاده شده، شبکه عصبی پیش‌خور پس‌انتشار خطا با سه لایه بود. گام نخست مدل‌سازی شبکه عصبی مصنوعی، شامل استانداردسازی داده‌های ورودی و خروجی است. همچنین، تحلیل حساسیت شبکه عصبی شامل تعداد نورون‌ها و توابع تحریک بررسی شد. تعداد نورون‌های بررسی شده ۲ تا ۷ بود و توابع تحریک بررسی شده شامل لگاریتم زیگموئیدی، خطی و تانژانت زیگموئیدی بودند. تحلیل حساسیت اهمیت زیادی در مدل‌سازی دارد، برای مثال با تغییر توابع تحریک تانژانت زیگموئیدی - لگاریتم

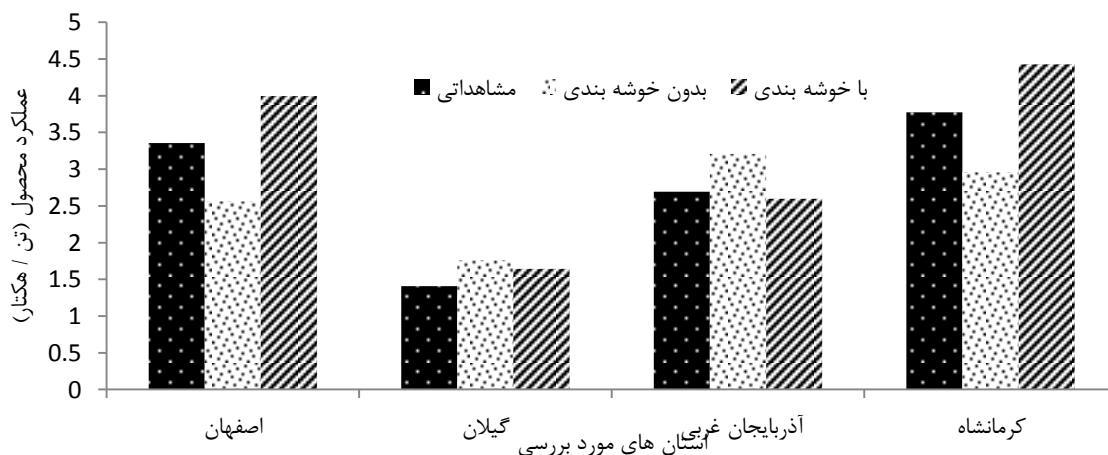
با توجه به شکل ۶ و با استفاده از خوشه‌بندی شاخص‌های اقلیمی، میزان خطا کاهش و مقدار شاخص‌سازی افزایش داشته است. بنابراین، تحلیل خوشه‌ای می‌تواند کارکرد مؤثری در برآورد عملکرد محصول داشته باشد. میزان درصد کاهش خطا در استان گیلان ۱۲/۵، آذربایجان غربی ۷/۳۱، کرمانشاه ۱۲/۵، اصفهان ۱۱/۷۶ بود. مسئله مهم دیگری که در مراحل خوشه‌بندی باید به آن دقت کرد، نوع خوشه‌های انتخابی است، چرا که در صورت انتخاب نادرست خوشه‌ها، مقدار خطا افزایش خواهد یافت. برای نمونه، در خوشه‌بندی استان اصفهان از تغییر چهار خوشه به پنج خوشه و تفکیک خوشه آخر، مقدار خطا ۵۶/۱۸ درصد افزایش خواهد یافت. در این حالت نیز

میزان درصد کاهش متوسط خطا (RMSE و RRSME) در استان گیلان ۱۲/۳۶، آذربایجان غربی ۲۳/۱۶، کرمانشاه ۱۸/۲۹، اصفهان ۵۶۱/۹۷ بود. افزایش کارایی عملکرد با استفاده از تحلیل خوشه‌های در بیشتر تحقیقات مشاهده شده است از جمله تحقیق در خوشه‌بندی بارندگی فصلی [۲۳] و در خوشه‌بندی زمانی سیلاب و تحلیل آثار آن بر شاخص‌های اقلیمی [۲۷]. در این بخش نیز کمینه مقدار خطا مربوط به استان اصفهان بود. در ادامه، مقایسه‌ای بین مقادیر عملکرد محصول مشاهداتی و تخمینی با شبکه عصبی مصنوعی در دو حالت خوشه‌بندی و بدون خوشه‌بندی انجام گرفت که در شکل ۸ نشان داده شده است.

زیگموئیدی (لایه میانی - خروجی) به تانژانت زیگموئیدی - تانژانت زیگموئیدی (لایه میانی - خروجی) میزان کاهش RMSE، ۷۳/۶ درصد بود. در شبکه عصبی نیز در حالت نخست ورودی‌های سیستم شامل زمان و یک شاخص اقلیمی که براساس نتایج شکل ۴ دارای کمترین مقدار خطا در تخمین عملکرد محصول انتخاب شد و در حالت دوم ورودی‌های سیستم شامل زمان و مقادیر استخراجی خوشه‌بندی بودند که تأثیر خوشه‌بندی شاخص‌های اقلیمی در شکل ۷ آورده شده است. در این حالت نیز با استفاده از خوشه‌بندی شاخص‌های اقلیمی دقت مدل‌سازی افزایش یافته است، به طوری که



شکل ۷. ارزیابی عملکرد خوشه‌بندی و استفاده از یک شاخص اقلیمی در تخمین عملکرد محصول با شبکه عصبی مصنوعی



شکل ۸. مقایسه تأثیر خوشه‌بندی بر عملکرد محصول طی دوره صحت‌سنجی

تمامی شاخص‌ها استفاده می‌شود و از سوی دیگر، تحلیل خوشه‌ای امکان استخراج اطلاعات از شاخص‌ها و ترکیب شاخص‌ها را با روش صحیحی امکان‌پذیر می‌سازد. در خوشه‌بندی با روش Ward از تحلیل واریانس در این مورد استفاده می‌شود. این روش در ماتریس مجاورت جست‌وجو می‌کند و دو نفری که کوچک‌ترین مقدار فاصله را داشته باشند، خوشه‌بندی می‌کنند. در واقع از خطای درون خوشه‌ها استفاده می‌کند تا بهترین خوشه‌بندی را به وجود آورد. نوع ترکیب داده‌های ورودی در مدل‌سازی اهمیت زیادی دارد، زیرا در حالت ترکیب شاخص‌ها با مقدار خطای کم (بدون خوشه‌بندی) بهبودی در نتایج مشاهده نشد. به‌طور کلی، در بررسی حالت‌های مختلف تخمین عملکرد گندم، کمینه خطا مربوط به استان اصفهان بود و با توجه به این مسئله شاید بتوان گفت که تأثیر نوع اقلیم نیز در کارایی مدل‌ها تأثیر زیادی دارد و اگر در مطالعات عملکرد محصول، پهنه‌بندی اقلیمی دخالت داشته باشد، به افزایش دقت مدل‌سازی منجر خواهد شد. براساس نتایج به‌دست‌آمده، دو عامل اصلی خوشه‌بندی و پهنه‌بندی اقلیمی نتایج را تحت تأثیر قرار می‌دهد. بنابراین، تحقیق درباره توسعه مسائل مطرح‌شده مانند خوشه‌بندی دومرحله‌ای و میانگین K برای افزایش کارایی مدل‌سازی عملکرد محصول - شاخص‌های اقلیمی پیشنهاد می‌شود.

#### منابع

- [1]. Biswas B, Dhaliwal L.K, Singh S.P, Sandhu S.K. Forecasting wheat production using ARIMA model in Punjab. *International Journal of Agricultural Sciences*.2014; 10(1): 158-161.
- [2]. Mathieu JA, Aires F. Assessment of the agro-climatic indices to improve crop yield forecasting. *Agricultural and Forest Meteorology*.2018; 253-254:15-30.
- [3]. Mousavi-Baygi M, Bannayan M, Ashraf B, Asadi Oskuei E. Assessment of climatic indices limiting rainfed wheat yield. *Ecological Indicators*.2016; 62: 298-305. [Persian]
- [4]. Paltasingh KR, Goyaria P, Mishra R.K. Measuring Weather Impact on Crop Yield Using Aridity Index: Evidence from Odisha. *Agricultural Economics Research Review*.2012; 25(2): 205-216.
- [5]. Shaw LH. The effect of weather on agricultural output: A look at methodology. *Journal of Farm Economics*.1964; 46(1): 218-230.

با توجه به شکل ۸، با استفاده از خوشه‌بندی مقادیر شبیه‌سازی به مقادیر مشاهداتی نزدیک می‌شوند. در مورد بیش‌برآورد و کم‌برآوردی هماهنگی جامعی بین استان‌ها وجود نداشت. در استان آذربایجان غربی دارای کم‌برآورد و در استان‌های کرمانشاه، اصفهان و گیلان دارای بیش‌برآورد است. درصد کاهش عملکرد محصول اصفهان در حالت بدون خوشه‌بندی نسبت به مقادیر مشاهداتی ۲۳/۵۹ درصد افزایش عملکرد محصول اصفهان در حالت خوشه‌بندی ۱۹/۰۳ درصد افزایش عملکرد محصول گیلان در حالت بدون خوشه‌بندی ۲۴/۹۸ درصد افزایش عملکرد محصول گیلان در حالت با خوشه‌بندی ۱۶/۴ درصد افزایش عملکرد محصول آذربایجان غربی در حالت بدون خوشه‌بندی ۱۹/۲۲ درصد کاهش عملکرد محصول آذربایجان غربی در حالت با خوشه‌بندی ۳/۵۵ درصد کاهش عملکرد محصول کرمانشاه در حالت بدون خوشه‌بندی ۲۱/۵ درصد افزایش عملکرد محصول کرمانشاه در حالت خوشه‌بندی ۱۷/۴ بود. در حالت خوشه‌بندی تغییرات کاهش و افزایش عملکرد محصول نسبت به مقادیر مشاهداتی کاهش می‌یابد.

#### بحث و نتیجه‌گیری

اقلیم هم به‌عنوان منبع و هم به‌عنوان محدودیت در کشاورزی مد نظر است، بنابراین باید درک درستی از تأثیر آن بر عملکرد محصول وجود داشته باشد. به این منظور، اقدام به مدل‌سازی شاخص‌های اقلیمی و عملکرد محصول گندم با تأکید بر ترکیب شاخص‌های اقلیمی و داده‌های هواشناسی در قالب خوشه‌بندی شده است. ارزیابی عملکرد شاخص‌های اقلیمی توسعه‌یافته بیان می‌کند که علاوه بر داده‌های دخیل در شاخص‌های اقلیمی، نوع فرمول‌بندی نیز اهمیت زیادی دارد. ادغام داده‌های هواشناسی در قالب شاخص‌های اقلیمی با ساختار صحیح نسبت به داده‌های هواشناسی، دقت مدل‌سازی را افزایش داد. انتخاب ورودی‌های مدل براساس تحلیل خوشه‌ای نسبت به استفاده مستقیم شاخص‌های اقلیمی عملکرد رگرسیون خطی و شبکه عصبی مصنوعی را افزایش داد (متوسط RMSE در کل استان‌ها با خوشه‌بندی = ۰/۷، بدون خوشه‌بندی = ۱/۱۵ - متوسط RRMSE با خوشه‌بندی = ۰/۲۹، بدون خوشه‌بندی = ۰/۵). دلیل این مسئله مربوط به ترکیب شاخص‌های اقلیمی است، چرا که از یک سو از

- [6].Oury B. Allowing for weather in crop production model building. *Journal of Farm Economics*.1965; 47(2): 270-283.
- [7].Rolla AL, Nuñez MN, Guevara ER, Meira SG, Rodriguez GR, Zárata MI. Climate impacts on crop yields in Central Argentina. *Adaptation strategies*. *Agricultural Systems*.2018; 60: 44-59.
- [8].Wang R, Bowling L.C, Cherkauer K.A. Estimation of the effects of climate variability on crop yield in the Midwest USA. *Agricultural and Forest Meteorology*.2016; 216: 141-156.
- [9].Sabziparvar AA, Torkaman M, Maryanaji Z. Investigating the effect of agroclimatic indices and variables on optimum wheat performance (Case study: Hamedan Province). *Journal of Water and Soil*.2013; 26(6): 1554-1567. [Persian]
- [10].Mosaedi A, Mohammadi Moghaddam S, Ghabaei Sough M. Modeling rain-fed wheat and barley based on meteorological features and drought indices. *Journal of Water and Soil*.2015; 29(3): 730-749. [Persian]
- [11].Alizadeh A, Kamali G, Mousavi F, Mousavi Bayeghi M. *Weather and climate*. Ferdowsi University of Mashhad;2012. [Persian]
- [12].Lang R. *Verwitterung und Bodenbildung alsEiführung die Bodenkunde*, Stuttgart;1920.
- [13]. De Martonne E. Une nouvelle fonction climatologique: l'Indice d'Aridite, *La Meteorologie*, October, Paris (Cited by Oury, 1965);1926.
- [14].Angstrom A A coefficient of humidity of general applicability. *Geografiska Annaler*,1936; 18: 245-254.
- [15].Köppen W. *Das geographische System der Klimate*. *Handbuch der Klimatologie*, v.1, Part C (W. Köppen & R. Geiger, eds.). Gerbrüder Bornträger, Berlin,1936: p. C1-C44, (Cited by Oury, 1965).
- [16].Ivanov NN. Landscape-climatic zones of the earth surface. *Proceedings , All Soviet Geographical Congress*, Vol. 1. Publ i cat ion of the Academy of Science of USSR, Leningrad.1948.
- [17]. Selianinov GT. On agricultural climate valuation. *Proceedings of Agricultural Meteorology*.1928; 20: 165-177(in Russian, cited by Meshcherskaya and Blazhevich,1997).
- [18].Thorntwaite CW. An approach toward a rational classification of climate. *Geogr Rev*.1948; 38:55-94.
- [19]. Gorbanzadeh Kharazi H, Cheleh Mal Dezfulnejad M. Innovation a new climate classification method based on hydrometeorology. 2014; 2(2):97-108.
- [20].Food and Agriculture Organization (FAO). *Arid Zone Forestry: A Guide for Field Technicians*; FAO: Rome, Italy, 1989.
- [21]. Arman N, Salajegheh A, Feiznia S, Ahmadi H, Ghodousi J, and Kiani Rad A. Determination of homogeneous watersheds for estimating erosion and sediment using different cluster analysis methods (Case study: Northern Alborz Range). *Journal of Range and Watershed Management*.2016; 69(2): 261-273. [Persian]
- [22].Khosravi M, Armesh M. Climatic Regionalization of Markazi Province: application of factor and cluster analysis. *Geography and Environmental Planning*. 2012; 46(2): 87-104. [Persian]
- [23].Wang HR, Wang C, Lin X, Kang J. An improved ARIMA model for precipitation simulations. *Nonlinear Processes in Geophysics*.2014; 21: 1159-1168.
- [24].Caubel J, de Cortázar-Atauri I, Launay M, de Noblet-Ducoudré N, Huard F, Bertuzzi P, Graux A. Broadening the scope for ecoclimatic indicators to assess crop climate suitability according to ecophysiological, technical and quality criteria. *Agricultural and Forest Meteorology*.2015; 207: 94-106.
- [25].Lepage M, Bourgeois G, Bélanger G. Indices agrométéorologiques pour l'aide à la 491eference dans un 491eferen de climat variable et en 491eference. Centre de 491eference en agriculture et agroalimentaire du Québec (CRAAQ). Québec, QC, Canada.2012.
- [26].Murtagh F, Legendre P. Ward's hierarchical agglomerative clustering method: which algorithms implement ward's criterion? *Journal of Classification*.2014; 31: 274-295.
- [27].Gu X, Zhang Q, Singh VP, Chen YD, Shi P. Temporal clustering of floods and impacts of climate indices in the Tarim River basin, China. *Global and Planetary Change*.2016; 147: 12-24.