

پیش‌بینی بلندمدت بارش بر پایه الگوهای پیوند دور اقلیمی، مطالعه موردی: حوضه آبریز اهرچای

قاسم کیان‌پیشه^۲

الهام کاکایی لدفانی^۳

علی مریدی^۱

آزاده احمدی^۱

پذیرش ۹۲/۳/۱۳

(دریافت ۹۱/۴/۳۰)

چکیده

هدف از این پژوهش، پیش‌بینی بارش فصلی حوضه آبریز اهرچای در شمال غربی ایران بود که با استفاده از الگوهای پیوند دور اقلیمی شامل اطلاعات متوسط فشار هوا و دما در سطح دریاها در طول دوره آماری ۱۹۶۵ تا ۲۰۰۵ به‌دست آمد. مدل‌های پیش‌بینی برای دو فصل تر (از دسامبر تا می) و فصل خشک (از ژوئن تا نوامبر) توسعه داده شدند. از این رو، پس از دریافت اطلاعات الگوهای پیوند دور اقلیمی شناسایی شده بر اقلیم شمال غرب کشور، از روش همبستگی و روش آزمون گاما برای انتخاب بهترین متغیرهای پیش‌بینی‌کننده و بهترین ترکیب آنها استفاده شد. در نهایت با استفاده از مدل ماشین بردار پشتیبان، بارش فصلی پیش‌بینی و نتایج آن با مدل رگرسیون چند متغیره مقایسه شد. نتایج نشان‌دهنده نقش مؤثر مدل آزمون گاما در تعیین متغیرهای ورودی و ترکیب آنها بود. همچنین عملکرد مدل ماشین بردار پشتیبان با مدل مبنای رگرسیون چند متغیره به‌عنوان یک مدل مبنای مقایسه شد.

واژه‌های کلیدی: پیش‌بینی بارش، الگوهای پیوند دور اقلیمی، ماشین بردار پشتیبان، آزمون گاما

Long-Lead Rainfall Prediction Based on Climate Patterns of Tele-Connection, A Case Study: Aharchay Basin

A. Ahmadi¹

A. Moridi²

E. Kakaei Lafdani³

Gh. Kianpisheh⁴

(Received July 20, 2012 Accepted June 3, 2013)

Abstract

This study aims to develop a seasonal rainfall prediction model for the Aharchay Basin, northwest of Iran. The model is based on climate patterns of tele-connection including sea level pressure (SLP) and sea surface temperature (SST) over the period from 1965 to 2005. The models cover both wet (from December to May) and dry (from June to November) seasons. For this purpose, the climatic patterns affecting the climate of the northwest of Iran were initially determined. In the second stage of the study, the correlation coefficient analysis and the Gamma Test (GT) technique were used to select the best predictors and to determine the best combination of the variables. The results revealed that the gamma test model outperformed the other model in determining the required input variables and their best combination. The seasonal rainfall in the basin was also predicted using the Support Vector Machines (SVM) and the results thus obtained were compared with those of the multivariate linear regression model as a benchmark to show the performance of the SVM model in rainfall prediction.

Keywords: Rainfall Prediction, Climate Patterns of Tele-Connection, Support Vector Machines, Gamma Test.

1. Assist. Prof., Dept. of Civil Eng., Isfahan University of Tech., Isfahan (Corresponding Author) (+98 31) 33912210 aahmadi@cc.iut.ac.ir
2. Assist. Prof., Dept. of Water Eng., Shahid Abbaspour College of Tech., Shahid Beheshti University, Tehran
3. PhD Student of Watershed Management Eng., Faculty of Natural Resources, Tarbiat Modares University, Noor, Iran
4. Grad. MSc Student, Dept. of Civil Eng., Isfahan University of Tech., Isfahan

- ۱- استادیار، دانشکده مهندسی عمران، دانشگاه صنعتی اصفهان (نویسنده مسئول) aahmadi@cc.iut.ac.ir (۰۳۱) ۳۳۹۱۲۲۱۰
- ۲- استادیار، دانشکده مهندسی آب، پردیس فنی مهندسی شهید عباسپور، دانشگاه شهید بهشتی، تهران
- ۳- دانشجوی دکتری علوم و مهندسی آبخیزداری، دانشکده منابع طبیعی، دانشگاه تربیت مدرس، نور، ایران
- ۴- دانش‌آموخته کارشناسی ارشد، دانشکده مهندسی عمران، دانشگاه صنعتی اصفهان

و نمایه بارش استاندارد شده، برای پیش‌بینی پدیده خشکسالی هواشناسی با استفاده از روش ماشین بردار پشتیبان استفاده کرده‌اند [۸].

یکی از مراحل مهم در تدوین مدل پیش‌بینی، تعیین ورودی‌های مناسب مدل است. ازدیاد تعداد متغیرهای ورودی سبب پیچیدگی ساختار شبکه و ناپایداری آن می‌شود و دقت آن را در مراحل آزمون مدل کاهش می‌دهد [۹]. به این منظور از روش‌هایی مانند تحلیل مولفه‌های اصلی^۷، خوشه‌بندی فازی و آزمون گاما استفاده می‌شود [۱۰-۱۴]. نوری و همکاران در سال ۲۰۱۰ و ۲۰۱۱ به بررسی نقش پیش‌پردازش پارامترهای ورودی به روش‌های تحلیل مولفه‌های اصلی، آزمون گاما^۸ و موجک بر عملکرد مدل ماشین بردار پشتیبان برای پیش‌بینی جریان رودخانه پرداخته‌اند. نتایج نشان دهنده نقش مؤثر پیش‌پردازش در تعیین متغیرهای ورودی بوده است [۱۵ و ۱۶]. دستورانی و همکاران در سال ۲۰۱۱ به منظور مدل‌سازی بارش-رواناب در حوضه آبخیز سد زاینده رود با استفاده از شبکه‌های عصبی مصنوعی و سیستم استنتاج عصبی-فازی تطبیقی، از الگوریتم ژنتیک و آزمون گاما برای تعیین بهترین ترکیب ورودی از پارامترها استفاده کرده‌اند [۱۷].

در این پژوهش از برخی روش‌های داده‌کاوی بر اساس الگوهای پیوند دور اقلیمی به منظور پیش‌بینی بارش حوضه آبریز اهرچای در استان آذربایجان شرقی استفاده شد. در این راستا ابتدا متغیرهای تأثیرگذار با استفاده از روش همبستگی و روش نوین آزمون گاما مشخص شد. در گام بعدی، مدل‌های پیش‌بینی بارش شش ماهه بر اساس اطلاعات پیوند دور اقلیمی شش ماه قبل و با استفاده از مدل‌های رگرسیون و ماشین بردار پشتیبان تدوین شد و نتایج دو مدل در پیش‌بینی بارش مقایسه شد.

۲- مواد و روش‌ها

۲-۱- منطقه مورد پژوهش و داده‌ها

در بخش شمال غرب کشور، حوضه آبریز اهرچای به دلیل پوشش دادن بخش قابل توجهی از حوضه آبریز ارس، از اهمیت قابل توجهی برخوردار است. میزان حداقل و حداکثر بارش حوضه ۲۳۰ و ۴۲۳ میلی‌متر و متوسط آن ۳۲۰ میلی‌متر است [۱۸]. خسارت‌های ناشی از خشکسالی در این حوضه، علی‌رغم میزان بارش نسبتاً مناسب، بسیار محسوس است. گرچه تفاوت زیادی در بارش در سال‌های اخیر در حوضه آبریز اهرچای وجود نداشته است و حوضه آبریز دوره‌هایی با بارش مشابه را تجربه کرده است، اما

پیش‌بینی بارش در ماه‌ها و یا حتی فصل‌های آینده، می‌تواند ابزار مناسبی در تخمین شرایط آبی به منظور مدیریت منابع آب باشد. از جمله فواید مدل پیش‌بینی با افق‌های زمانی بلند مدت، این است که تصمیم‌گیری‌های مبتنی بر ذخیره و آزادسازی از مخازن آب، پویاتر خواهد بود و انعطاف در برخورد با شرایط خاص بیشتر شده و منجر به کسب سود بیشتر در بهره‌برداری از منابع آب می‌شود. همچنین با مشخص بودن این پیش‌بینی‌ها می‌توان در صورت نیاز، برنامه‌های اضطراری برای مدیریت بهینه تأمین آب در دوره‌های خشکسالی و مقابله با سیلاب‌های احتمالی ارائه داد، که اثرات آن بر روی سلامت عمومی، فعالیت‌های تجاری و اقتصادی و محیط زیست کاملاً محسوس است.

در دهه‌های اخیر، به منظور توسعه مدل‌های پیش‌بینی بارش در مقیاس بلند مدت (ماهانه و فصلی)، تلاش‌های بسیاری صورت گرفته است. در این رابطه، مدل‌های مختلفی برای ایجاد رابطه بین الگوهای پیوند دور اقلیمی، از جمله متوسط فشار در سطح دریا^۱ و دما در سطح دریا^۲ (پیش‌بینی کننده)، و متغیرهای محلی (پیش‌بینی شونده) مانند دما و بارندگی، توسعه یافته‌اند. همچنین در سال‌های اخیر بسیاری از روش‌های غیرخطی مانند شبکه‌های عصبی مصنوعی^۳، ماشین‌های بردار پشتیبان^۴ و K نزدیک‌ترین همسایگی به دلیل توانایی بالای آنها در شبیه‌سازی سیستم‌های غیرخطی و پیچیده، به عنوان مدل پیش‌بینی به کار برده شده‌اند [۱، ۲، ۳]. ناگش کومار و همکاران در سال ۲۰۰۷ برای پیش‌بینی بارش، از الگوهای پیوند دور اقلیمی، شامل فشار و دما در سطح آزاد دریاها، و روش شبکه عصبی مصنوعی به همراه الگوریتم ژنتیک^۵، برای به دست آوردن پارامترهای بهینه مدل شبکه عصبی مصنوعی استفاده کرده‌اند [۴]. اموگبای در سال ۲۰۱۰ به پیش‌بینی بارندگی شمالی نیجریه بر اساس داده‌های دمای سطح دریا^۶ بر اساس روش رگرسیون پرداخته است [۵]. کارآموز در سال‌های ۲۰۰۴ و ۲۰۰۹، الگوهای پیوند دور اقلیمی مؤثر بر ایران را، بر اساس مشاهده تغییرات بازه‌ای و استفاده از روش‌های آماری از جمله تحلیل عاملی و همبستگی شناسایی نموده و از روش‌های داده‌کاوی، مانند قوانین فازی و شبکه عصبی مصنوعی به عنوان مدل‌های پیش‌بینی بارش استفاده کرده است [۶ و ۷]. نیکبخت شهبازی و همکاران در سال ۲۰۱۱، از متغیرهای هواشناسی شامل دمای هوا، ارتفاع ژئوپتانسیل

¹ Mean Sea Level Pressure (SLP)

² Mean Sea Level Temperature (SST)

³ Artificial Neural Network (ANN)

⁴ Support Vector Machine (SVM)

⁵ Genetic Algorithm (GA)

⁶ Sea Surface Temperature (SST)

⁷ Principal Component Analysis (PCA)

⁸ Gamma Test (GT)

اطلاعات ۲/۵ درجه طول و عرض جغرافیایی است. در این پژوهش، داده‌های بارش برای سه ایستگاه باران‌سنجی کسائق، ورزقان و کاسین از بانک اطلاعاتی سازمان هواشناسی کشور و وزارت نیرو استخراج شد و میانگین بارش حوضه‌ای به‌دست آمد. طول دوره آماری از سال ۱۹۶۵ تا ۲۰۰۵ و برای دو فصل تر در ماه‌های آذر تا اردیبهشت (دسامبر تا می) و فصل خشک در ماه‌های خرداد تا آبان (ژوئن تا نوامبر) در نظر گرفته شد. عملکرد مدل‌های پیش‌بینی با استفاده از معیارهای آماری خطای انحراف متوسط^۳، خطای ریشه متوسط مربعات^۴، خطای قدر مطلق میانگین^۵ و همبستگی بین مقادیر پیش‌بینی و واقعی^۶ (R^2) مورد بررسی قرار گرفت [۱۹].

۲-۲- مدل‌های مورد استفاده در پیش‌بینی بارش

در این بخش، شرح مختصری از برخی مدل‌های استفاده شده برای مدل‌سازی ارائه شده است. همچنین مراحل انجام کار در شکل ۲ نشان داده شده است.

³ Mean Bias Error (MBE)

⁴ Root Mean Square Error (RSME)

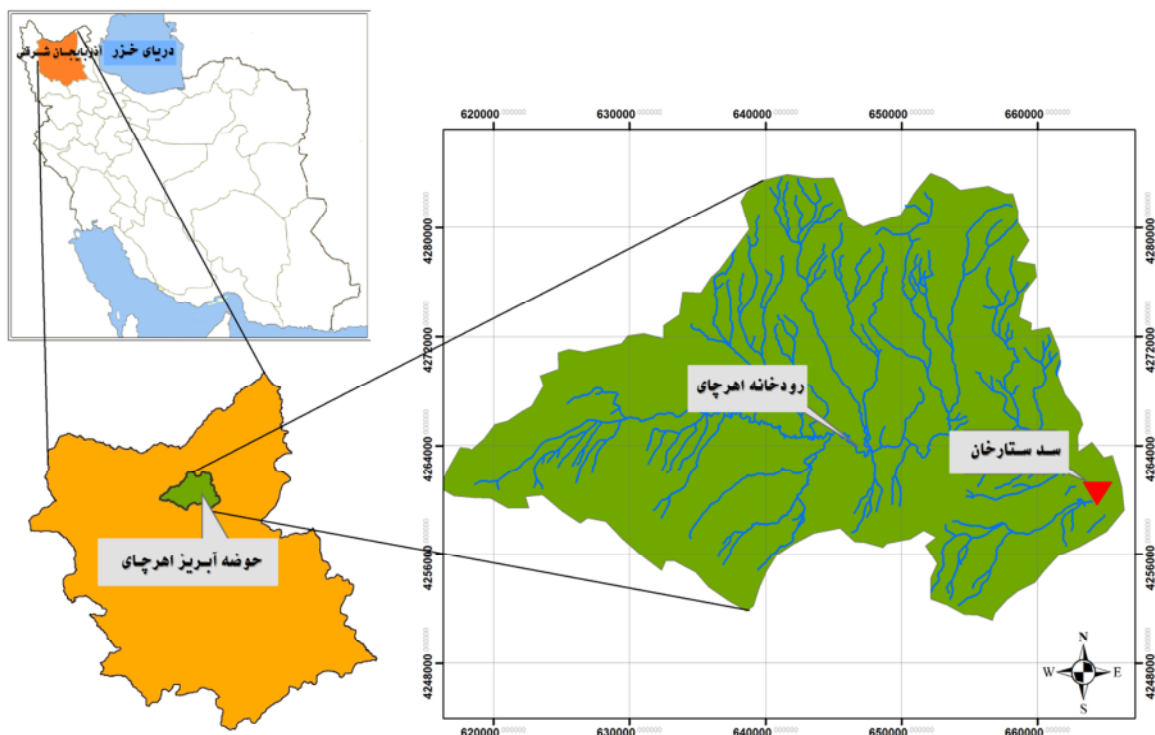
⁵ Mean Absolute Error (MAE)

⁶ R Squared

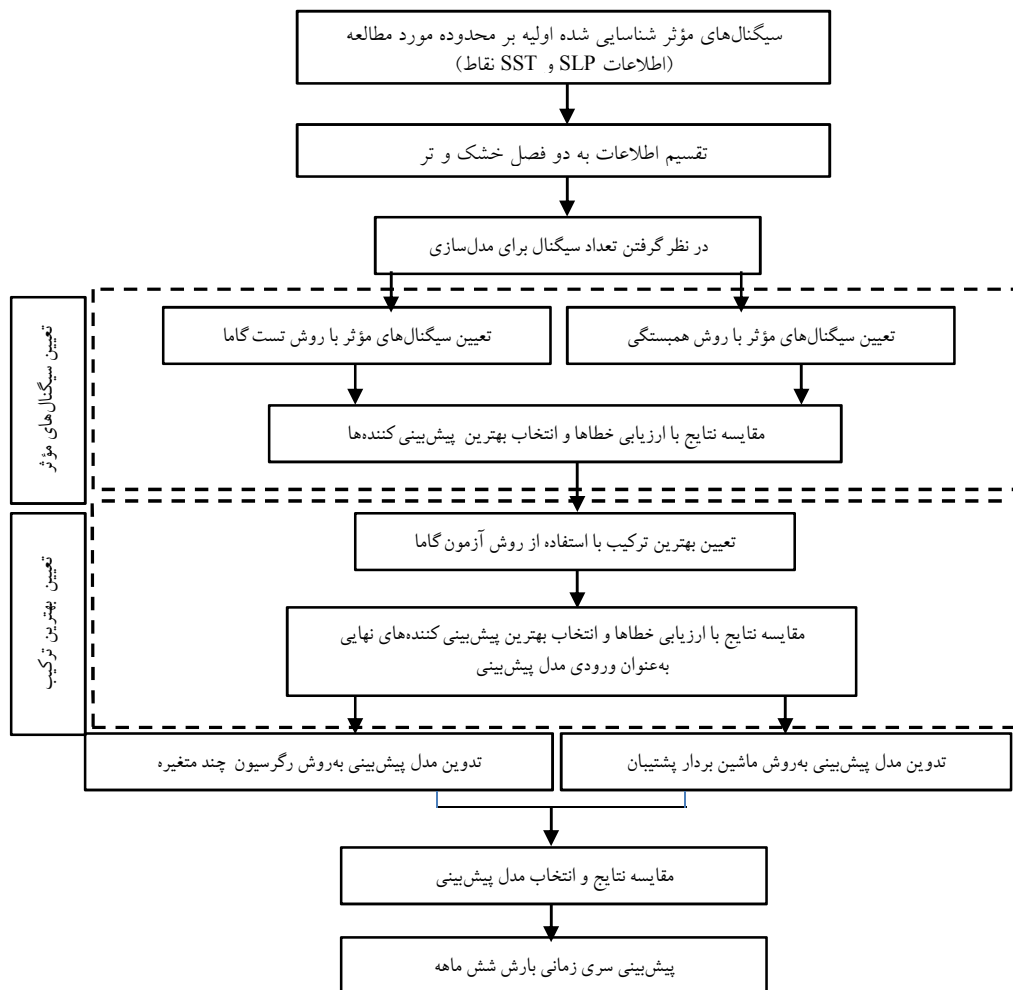
جریان ورودی به مخزن سد تا حدود زیادی کاهش یافته و مخزن سد در بیشتر اوقات سال نسبتاً خالی است. مهم‌ترین منبع ذخیره آب سطحی موجود، در حوضه سد ستارخان است و میزان آسیب‌پذیری حوضه نسبت به کم‌آبی بالا است. با وجود احداث شبکه‌های آبیاری در پایین‌دست سد، تدوین مدل شبیه‌سازی بارش حوضه آبریز و شناسایی دوره‌های خشک و تر آبی برای تدوین سیاست‌های مخزن سد ضروری است. در شکل ۱ موقعیت حوضه آبریز رودخانه مورد پژوهش و سد ستارخان نشان داده شده است. پژوهش‌ها در دهه‌های اخیر نشان می‌دهد که تغییرات آب و هوایی جهان و همچنین چرخه عمومی جو، تحت تاثیر الگوهای پیوند دور اقلیمی است که در سطح آب‌های آزاد جهان اتفاق می‌افتد. الگوهای پیوند دور اقلیمی مزبور، معمولاً با استفاده از شاخص‌هایی که بر اساس فشار هوا و دمای سطح آب اقیانوس‌ها محاسبه می‌شوند، کمی می‌شوند. از این داده‌ها به‌منظور تعیین ارتباط با بارش منطقه و تدوین مدل پیش‌بینی بارش استفاده شده است. داده‌های دمای سطح آب از داده‌های مؤسسه NCAR^۱ که از زیربخش‌های NOAA^۲ است، استخراج شده است. دقت مکانی این

¹ National Center for Atmospheric Research (NCAR)

² National Oceanic and Atmospheric Administration (NOAA)



شکل ۱- موقعیت حوضه آبریز رودخانه اهرچای و سد ستارخان در استان آذربایجان شرقی



شکل ۲- فلوچارت مراحل انجام پژوهش

۲-۱-۲-۲- آزمون گاما

آزمون گاما حداقل متوسط مجذور خطایی که در مدل‌های غیرخطی پیوسته با داده‌های مشاهده نشده قابل حصول است را تخمین می‌زند [۱۲]. آزمون گاما اولین بار توسط استفانسون و همکاران در سال ۱۹۹۷ توسعه داده شد و بعدها توسط بسیاری از دانشمندان استفاده و بحث شد [۱۹-۲۲]. اگر یک سری داده مشاهداتی به صورت زیر وجود داشته باشد

$$((x_1, \dots, x_m), y) = (X, y) \quad (1)$$

که در آن $X = (x_1, \dots, x_m)$ بردار ورودی در محدوده $C \in R^m$ و y بردار خروجی است. چنانچه رابطه زیر بین اعضای جامعه برقرار باشد

$$y = f(x_1, \dots, x_m) + r \quad (2)$$

که در آن r متغیر تصادفی است. آزمون گاما تخمینی برای واریانس خروجی مدلی غیرهموار است. این آزمون گاما بر اساس $N[i, k]$ که شامل لیستی از k امین همسایگی برای هر بردار X_i ($1 \leq i \leq M$) است، می‌باشد. تابع دلتا میانگین مجذور فاصله k امین همسایگی را محاسبه می‌کند [۶]

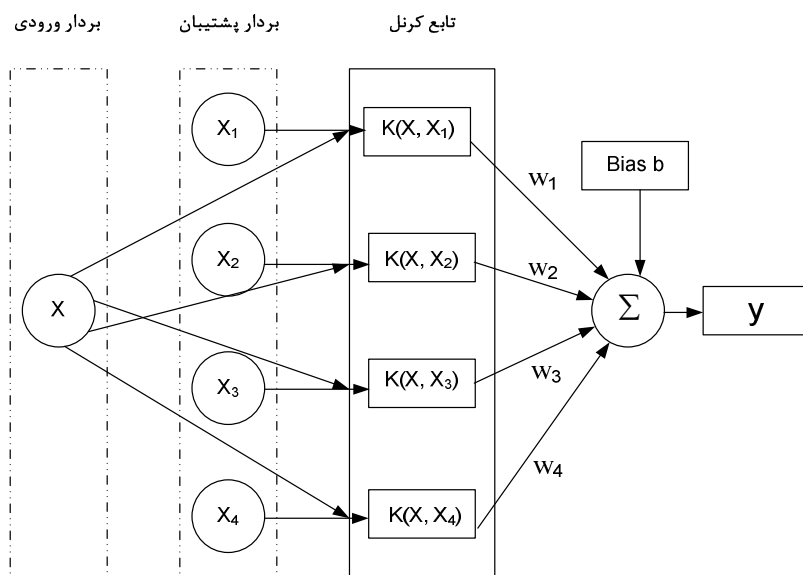
$$\delta_M(k) = \frac{1}{M} \sum_{i=1}^M |X_{N[i,k]} - X_i|^2 \quad (3)$$

که در آن $| \cdot |$ نشان‌دهنده فاصله اقلیدسی^۱ است و تابع گاما متناظر آن

$$\gamma_M(k) = \frac{1}{2M} \sum_{i=1}^M (y_{N[i,k]} - y_i)^2 \quad (4)$$

که در آن $y_{N[i,k]}$ مقدار y متناظر با k امین همسایگی X_i در C است. به منظور محاسبه Γ ، یک خط رگرسیون خطی از p نقطه بر مقادیر $\delta_M(k)$ و $\gamma_M(k)$ برازش داده می‌شود

¹ Euclidean distance



شکل ۳- ساختار مدل SVM

داده‌های آموزشی انتخاب می‌شود). بردارهای پشتیبان کاهش یافته قابلیت تعمیم مدل را افزایش می‌دهند و بار محاسباتی آن را کاهش می‌دهند. ساختار عمومی این مدل‌ها در شکل ۳ نشان داده شده است. از آنجا که ماشین‌های بردار پشتیبان در ابتدا توسط انجمن یادگیری ماشینی تولید شد، این نوع از مدل‌ها، ماشین‌های بردار پشتیبان نامیده می‌شوند [۲۴].

ساختار ویژه تابع کرنل برای تبدیل غیرخطی، قادر به رها شدن از بیشترین بردارهای آموزشی است. تابع کرنل امکان استفاده از یک تابع غیرخطی در فضای ورودی‌ها را برای تبدیل به تابع خطی در فضای ویژگی‌ها فراهم می‌کند. توابع کرنل خطی، چندضلعی و بر پایه شعاعی^۲ از انواع توابع کرنل است. تابع بر پایه شعاعی بر خلاف تابع کرنل خطی، در مواقعی که رابطه غیرخطی پیچیده در فضای ورودی وجود دارد، مناسب است. همچنین تابع کرنل بر پایه شعاعی، دارای پارامترهای تنظیمی کمتری نسبت به توابع کرنل چندضلعی است و پیچیدگی کمتری در مدل‌سازی دارد. انتخاب پارامترهای مدل و تابع کرنل، در فرایند سعی و خطا و بر اساس میزان خطای حداقل در مرحله آزمون انجام شد.

۳- نتایج و بحث

۳-۱- شناسایی الگوهای پیوند دور اقلیمی مؤثر بر منطقه

مورد پژوهش

کارآموز در پژوهشی، با استفاده از اطلاعات متوسط دمای سطح دریا، فشار هوا در سطح دریا و اختلاف فشار بین نقاط مختلف

$$\gamma = A\delta + \Gamma \quad (5)$$

عرض از مبدا این خط $\delta = 0$ ، مقدار Γ را نشان می‌دهد و $\gamma_M(k)$ برابر با واریانس خطاها است.

۲-۲-۲- ماشین بردار پشتیبان

ماشین‌های بردار پشتیبان، دسته‌ای از روش‌های یادگیری با ناظر^۱ هستند که برای مسائل طبقه‌بندی و رگرسیونی به‌کار می‌روند. این روش در سال ۱۹۹۵ توسط واپنیک و همکاران معرفی شده و بر پایه نظریه یادگیری محاسباتی توسعه یافته است [۲۳]. در سال‌های اخیر تعداد زیادی از انواع مدل‌های ماشین بردار پشتیبان در دسته‌بندی و رگرسیون توسعه داده شده است که به‌عنوان مدل مقایسه‌ای مبنا برای شبکه‌های عصبی مصنوعی به‌کار رفته است. نتایج نشان داده است که کارکرد تجربی این مدل‌ها عموماً بهتر از مدل شبکه عصبی است. این بهبود در نتایج می‌تواند به دلیل تعداد پارامترهای مورد استفاده کمتر برای بهینه‌سازی رویکرد ماشین بردار پشتیبان، کاهش امکان پذیری آموزش بیش از اندازه شبکه در مرحله یادگیری شبکه و افزایش کارکرد واقعی آن باشد.

در مدل ماشین بردار پشتیبان، تمامی داده‌های ورودی به صورت بردار ساماندهی می‌شوند و برخی از آنها در فرایند مدل‌سازی استفاده می‌شوند. این مدل، داده‌های آموزشی را برای کالیبره نمودن مدل و تخمین پارامترهای آن استفاده می‌کند ولی مهم‌ترین قسمت داده‌های بردار ورودی را در مدل نگه می‌دارد. این بردارها، بردارهای پشتیبان نامیده می‌شود (تنها بخش کوچکی از

² Radial Basis Function(RBF)

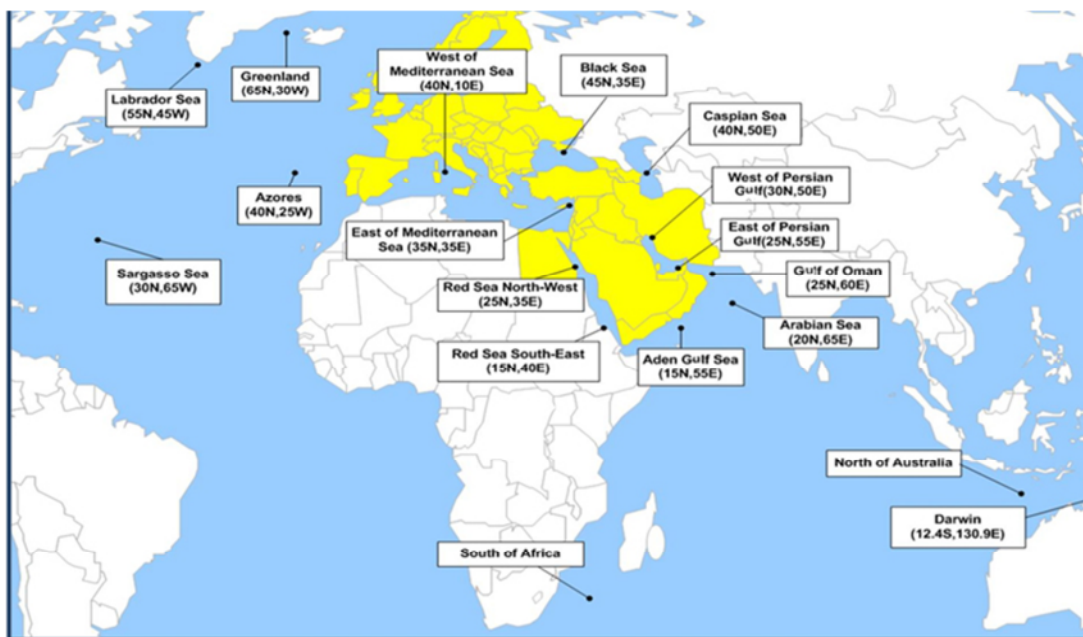
¹ Supervised Learning

گرفته شده‌اند، در جدول ۱ نشان داده شده است [۶]. برای انجام تحلیل آزمون گاما از نرم افزار وین گاما^۱ استفاده شد. از کدهای در دسترس در محیط متلب^۲ موجود در کتابخانه SVM، برای مدل سازی ماشین بردار پشتیبان استفاده شده است [۲۱ و ۲۵].

¹ WinGamma

² Matlab

انتخابی نشان داده شده در شکل ۴، پانزده الگوی پیوند دور اقلیمی مؤثر بر اقلیم منطقه شمال غرب را شناسایی نمود [۶]. کارآموز در سال ۲۰۰۴ و ۲۰۰۹، الگوهای پیوند دور اقلیمی مؤثر بر ایران را بر اساس مشاهده تغییرات بازه‌ای و استفاده از روش‌های آماری از جمله تحلیل عاملی و همبستگی شناسایی کرد. الگوهای پیوند دور اقلیمی که به‌عنوان بهترین پیش‌بینی‌کنندگان برای بارش شمال غرب ایران پیشنهاد شده‌اند و برای هر دو فصل سال یکسان در نظر



شکل ۴- تأیید نقاط مشخص شده برای بررسی اثر اختلاف فشار در سطح دریا و دمای سطح دریا بر متوسط بارش حوضه‌های آبریز شمال غرب، غرب و جنوب غرب [۷]

جدول ۱- الگوهای پیوند دور اقلیمی‌های مؤثر بر اقلیم شمال غرب ایران [۶]

شماره پیش‌بینی‌کننده	گرادیان فشار بین جنوب گرینلند و ...	گرادیان فشار بین جنوب سبیری و ...	دمای سطح آب در نقاط
۱	آزور	---	---
۲	غرب مدیترانه	---	---
۳	شرق مدیترانه	---	---
۴	دریای سیاه	---	---
۵	---	سودان	---
۶	---	شرق خلیج فارس	---
۷	---	دریای مازندران	---
۸	---	دریای عدن	---
۹	---	دریای سیاه	---
۱۰	---	شرق مدیترانه	---
۱۱	---	اقیانوس هند	---
۱۲	---	شمال غربی دریای سرخ	---
۱۳	---	دریای عرب	---
۱۴	---	غرب خلیج فارس	---
۱۵	---	شرق خلیج فارس	---

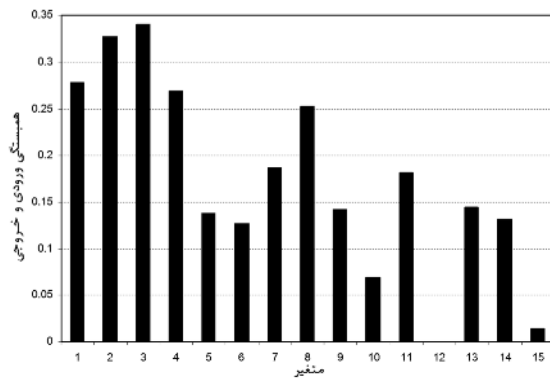
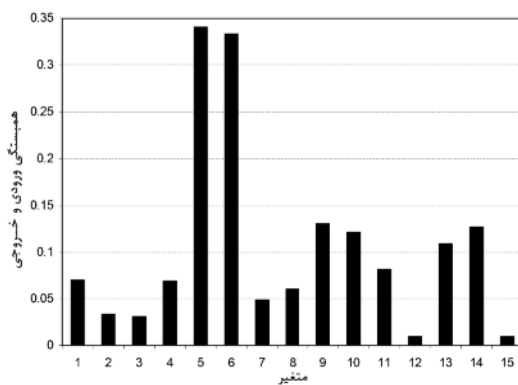
۲-۳- انتخاب پیش‌بینی‌کننده‌های مناسب

به‌منظور کاهش پیچیدگی مدل، تعداد متغیرهای لازم برای مدل‌سازی شش متغیر از بین ۱۵ متغیر پیش‌بینی‌کننده، در نظر گرفته شد. برای انتخاب متغیرهای مؤثر از روش سنتی رگرسیون بین متغیرهای ورودی و خروجی و روش جدید آزمون گاما استفاده شد. نتایج همبستگی‌ها برای بارش فصل تر و فصل خشک در شکل ۵ نشان داده شده است.

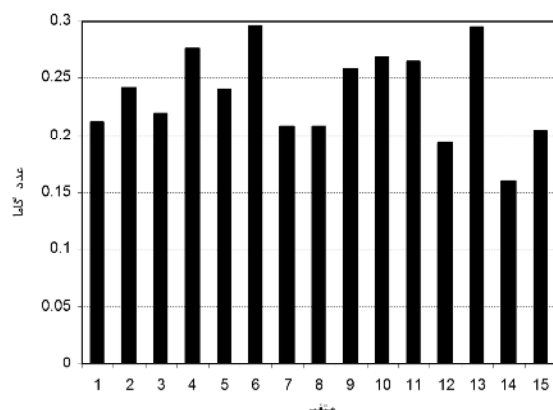
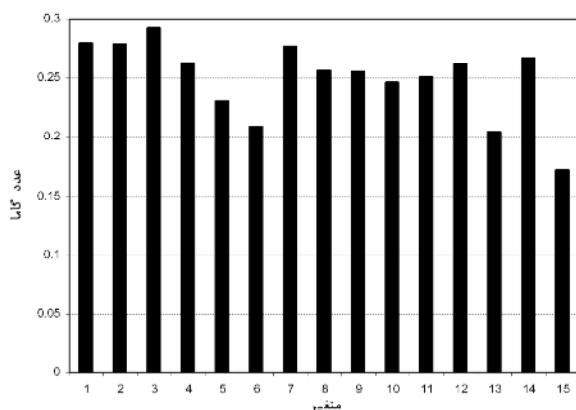
در فصل تر (ماه‌های دسامبر تا می) متغیرهای ۵، ۶، ۹، ۱۴، ۱۰ و ۱۳، بیشترین همبستگی را با بارش حوضه اهرچای دارند. اما ممکن است بین بعضی از این متغیرها همبستگی وجود داشته باشد و یکی از آنها بتواند معرف یک یا چند متغیر دیگر باشد. بنابراین ورودی‌های مدل پیش‌بینی، با توجه به ماتریس کواریانس آنها و بر اساس خود همبستگی بین متغیرها انتخاب می‌شوند. در ماه‌های دسامبر تا می، از آنجا که همبستگی زیادی (برابر با ۰/۹۳) بین متغیر ۵ و ۶ وجود داشت، متغیر ۶ حذف و به‌جای آن متغیر ۱۱ به ورودی‌های مدل اضافه شد. بنابراین متغیرهای ۵، ۹، ۱۴، ۱۰، ۱۳ و

۱۱ به‌عنوان ورودی مدل پیش‌بینی (در فصل تر) با استفاده از روش رگرسیون انتخاب شدند. معیار قابل قبول بودن همبستگی بین متغیرها، به میزان ۰/۸۵ در نظر گرفته شد. عدد گاما برای بارش منطقه در فصل تر و خشک به ازای الگوهای پیوند دور اقلیمی‌های مختلف در شکل ۶ آمده است. با توجه به شکل ۶ در فصل تر، الگوهای پیوند دور اقلیمی‌هایی که کمترین عدد گاما با بارش منطقه را دارند، شامل متغیرهای ۱۵، ۱۳، ۶، ۵، ۱۰ و ۱۱ است. بنابراین مدل اول آزمون گاما شامل متغیرهای ۱۵، ۱۳، ۶، ۵، ۱۱ و ۱۰ است.

در مدل دوم آزمون گاما، با توجه به همبستگی بین متغیرهای ۵ و ۶ (برابر با ۰/۹۳)، متغیر ۵ حذف و به‌جای آن متغیر ۹ جایگزین می‌شود. بنابراین برای مدل دوم آزمون گاما با در نظر گرفتن همبستگی بین متغیرها، متغیرهای ۱۵، ۱۳، ۶، ۱۰، ۱۱ و ۹ به‌عنوان متغیر ورودی در نظر گرفته شدند. با توجه به متغیرهای شناسایی شده توسط آزمون گاما و روش همبستگی برای فصل تر، سه مدل رگرسیون خطی محلی تدوین شد که در آن‌ها از ۳۲ داده در



شکل ۵- نتایج همبستگی بین ۱۵ الگوهای پیوند دور اقلیمی در نظر گرفته شده و میزان بارش منطقه ماه‌های دسامبر تا می (راست) و ماه‌های ژوئن تا نوامبر (چپ)



شکل ۶- عدد گاما بین سری زمانی بارش می تا دسامبر و الگوهای پیوند دور اقلیمی‌های ۱۵‌گانه (راست) و سری زمانی بارش ژوئن تا نوامبر و الگوهای پیوند دور اقلیمی‌های ۱۵‌گانه (چپ)

مرحله آموزش و ۸ داده برای بررسی عملکرد هر مدل در مرحله آزمون استفاده شده است. با توجه نتایج ارائه شده در جدول ۲، مشاهده می‌شود که در فصل تر، تعیین پیش‌بینی‌کننده‌ها به روش آزمون گاما با حذف متغیرهای وابسته، بهترین مدل در مرحله آموزش شبکه بوده است؛ زیرا دارای حداقل خطای انحراف متوسط، خطای ریشه متوسط مربعات و حداکثر همبستگی بین مقادیر پیش‌بینی و واقعی است. ولی در مرحله آزمون مدل، با متغیرهای ورودی تعیین شده توسط آزمون گاما و بدون حذف متغیرهای وابسته، دارای بهترین عملکرد بوده است و این مطلب نشان‌دهنده کارایی مدل آزمون گاما در انتخاب متغیرهای ورودی برای مدل پیش‌بینی در فصل تر (دسامبر تا می) است.

با توجه به شکل ۵، در فصل خشک بیشترین همبستگی بین بارش منطقه با متغیرهای ۳، ۲، ۱، ۴، ۸ و ۷ وجود دارد. در ماه‌های ژوئن تا نوامبر با توجه به ماتریس کواریانس ورودی‌های مدل پیش‌بینی، همبستگی زیادی بین متغیرهای ۳ و ۱ (برابر با ۰/۹۴)، بین متغیرهای ۳ و ۴ (برابر با ۰/۹۷) و بین متغیرهای ۴ و ۲ (برابر با ۰/۹۵) وجود دارد. لذا متغیرهای ۱، ۲ و ۴ حذف شدند و به جای آن متغیر ۱۳، متغیر ۹ و متغیر ۵، به لیست ورودی‌های مدل اضافه شدند. بنابراین متغیرهای ۳، ۸، ۷، ۱۳، ۹ و ۵ به‌عنوان ورودی‌های مدل پیش‌بینی در فصل خشک، با استفاده از روش رگرسیون انتخاب شدند.

با توجه به شکل ۶ در فصل خشک (ماه‌های ژوئن تا نوامبر)، الگوهای پیوند دور اقلیمی که کمترین عدد گاما با بارش منطقه را دارند، شامل متغیرهای ۱۴، ۱۲، ۱۵، ۷، ۸ و ۱ هستند. به دلیل اینکه همبستگی معنی‌داری بین متغیرهای پیشنهادی وجود ندارد، بنابراین متغیرهای ۱۴، ۱۲، ۱۵، ۷، ۸ و ۱ به‌عنوان ورودی‌های مدل پیش‌بینی در نظر گرفته شدند. با توجه به متغیرهای شناسایی شده توسط هر مدل، دو مدل رگرسیون خطی محلی تدوین شد (جدول ۳).

با توجه به نتایج ارائه شده در جدول ۳، مشاهده می‌شود که در فصل خشک (ژوئن تا نوامبر)، تعیین پیش‌بینی‌کننده‌ها به روش همبستگی بهترین مدل در مرحله آموزش و آزمون شبکه بوده است زیرا دارای حداقل خطای انحراف متوسط، خطای ریشه متوسط مربعات و حداکثر همبستگی بین مقادیر پیش‌بینی و واقعی است.

۳-۳- انتخاب بهترین ترکیب پیش‌بینی‌کننده‌ها

هدف از انتخاب بهترین ترکیب پیش‌بینی‌کننده‌ها، شناسایی و حذف الگوهای پیوند دور اقلیمی‌هایی است که وجود آنها تنها بر پیچیدگی مدل می‌افزاید و تأثیر قابل توجهی در نتایج مدل ندارد. برای این منظور از روش آزمون گاما برای دستیابی به مدل با حداقل خطای گاما استفاده شد. ترکیبات مختلف از وجود یا عدم وجود هر

جدول ۲- مقایسه نتایج آزمون گاما و همبستگی در انتخاب بهترین پیش‌بینی‌کننده‌ها برای ماه‌های دسامبر تا می (فصل تر)

مرحله	آموزش		آزمون	
الگوهای پیوند دور اقلیمی‌های مؤثر	۱۰، ۱۴، ۹، ۵	۶، ۵، ۱۳، ۱۵	۱۰، ۱۴، ۹، ۵	۶، ۵، ۱۳، ۱۵
مدل	همبستگی	آزمون گاما	همبستگی	آزمون گاما
Bias (MBE)	-۸	-۶	-۳۲	-۱۲
RMSE	۲۵	۲۸	۵۵	۵۲
MAE	۱۹	۲۲	۳۵	۲۹
R ²	۰/۸۶	۰/۸۴	۰/۴۹	۰/۵۳

سه معیار اول بر حسب میلی‌متر و R² بر حسب درصد است.

جدول ۳- مقایسه نتایج تست گاما و همبستگی در انتخاب بهترین پیش‌بینی‌کننده‌ها برای ماه‌های ژوئن تا نوامبر (فصل خشک)

	آموزش		آزمون	
	همبستگی	آزمون گاما	همبستگی	آزمون گاما
Bias (MBE)	۱	-۸	۸	-۱۳
RMSE	۲۷	۲۹	۳۹	۵۳
MAE	۲۰	۲۳	۳۰	۳۸
R ²	۰/۸۶	۰/۸۵	۰/۳۰	۰/۰۶

سه معیار اول بر حسب میلی‌متر و R² بر حسب درصد است.

در فصل خشک (ماه‌های ژوئن تا نوامبر) ترکیب چهار متغیر شامل متغیرهای ۳، ۷، ۹ و ۵، بهترین مدل است.

۳-۴- مقایسه نتایج مدل پیش‌بینی با مدل رگرسیون نتایج حاصل از پیش‌بینی بارش در فصل تر و فصل خشک با استفاده از مدل‌های ماشین بردار پشتیبان و مدل رگرسیون چند متغیره، به‌عنوان مدل مبنا در جدول ۵ نشان داده شده است. برای پیش‌بینی با استفاده از مدل SVM- γ و کرنل RBF، لازم است مقادیر بهینه سه پارامتر C ، ϵ و γ تعیین شود. پارامترهای C و ϵ بر کیفیت و زمان آموزش تأثیر می‌گذارند. همچنین مقدار پارامتر γ می‌تواند در وقوع بیش‌برازش و دون‌برازش شبکه تأثیر گذار باشد. در این پژوهش مقادیر این پارامترها با استفاده از روش سعی و خطا تعیین شد. در انتخاب مقادیر بهینه پارامترها، مدل با استفاده از این مقادیر، کمترین خطا را در مرحله آزمون داشت. پارامترهای C ، ϵ و γ در مدل به ترتیب برابر 0.007 ، $2/5$ و $3/4$ انتخاب شدند. نتایج حاصل از مدل ماشین بردار پشتیبان در مقایسه با داده‌های تاریخی فصل تر و فصل خشک متناظر آن در شکل ۷ نشان داده شده است. همانگونه که از نتایج ارائه شده در جدول ۵ دیده می‌شود، در هر دو فصل تر و فصل خشک، مدل

مدل پیش‌بینی بارش در منطقه مورد پژوهش، $2^6 - 1 = 63$ است. در جدول ۴ نتایج ترکیبات منتخب مورد استفاده توسط مدل آزمون گاما آورده شده است. همانگونه که مشاهده می‌شود، در فصل تر، عدد گاما در ترکیب چهار ورودی اول (۱۱۱۱۰۰)، کمترین مقدار را دارد، بنابراین سناریوی پیشنهادی توسط آزمون گاما، ترکیب ۱۱۱۱۰۰ است. در فصل خشک، عدد گاما در ترکیب چهار ورودی (۱۰۱۰۱۱)، کمترین مقدار را دارد. بنابراین الگوهای پیوند دور اقلیمی‌های مؤثر بر بارش فصل تر (ماه‌های دسامبر تا می) محدوده مورد پژوهش، شامل متغیرهای ۱۵، ۶، ۱۳ و ۵ بود.

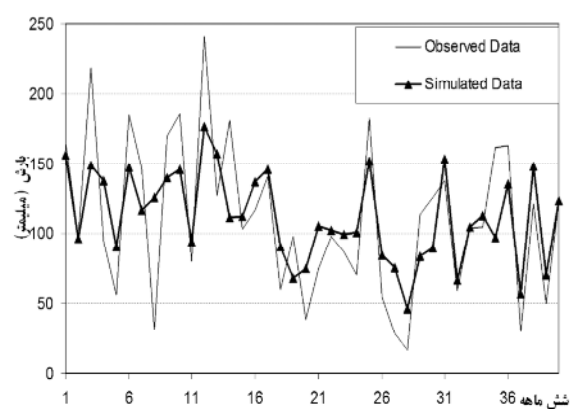
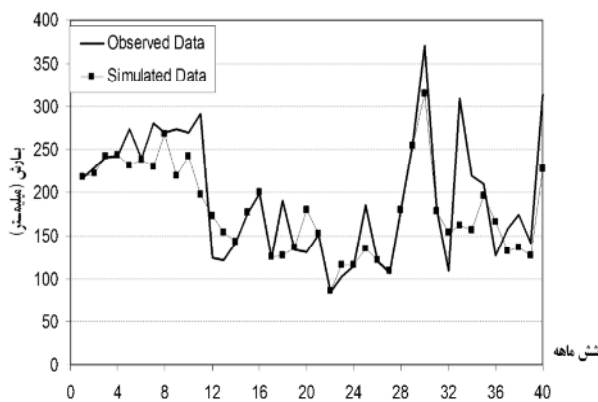
جدول ۴- نتایج آزمون گاما برای ترکیبات منتخب ورودی

مدل پیش‌بینی بارش		
گاما	ترکیب ورودی	
۰/۱۷۲	۱۰۰۰۰۰	۱
۰/۱۷۵	۰۱۰۱۰۰	۲
۰/۲۲۵	۱۱۱۱۱۱	۳
۰/۱۷۰	۱۱۱۱۰۰	۴
۰/۱۴۶	۱۰۱۰۱۰	۱
۰/۲۰۰	۱۱۱۱۱۱	۲
۰/۱۵۷	۱۰۱۱۱۱	۳
۰/۱۳۹	۱۰۱۰۱۱	۴

جدول ۵- نتایج مدل‌سازی به روش‌های مختلف در فصل تر و فصل خشک

آزمون		فصل تر		فصل خشک		فصل تر		فصل خشک	
SVM	Regr	SVM	Regr	SVM	Regr	SVM	Regr	SVM	Regr
-۱	۴	-۲	-۳۴	۱	۰	-۳	۰		
۳۴	۴۶	۴۰	۵۰	۱۳	۲۰	۳۱	۴۳		
۲۸	۳۸	۳۱	۳۴	۹	۱۳	۲۲	۲۷		
۰/۷۶	۰/۴۶	۰/۶۰	۰/۴۹	۰/۷۸	۰/۵۹	۰/۶۹	۰/۶۶		

سه معیار اول بر حسب میلی‌متر و R^2 بر حسب درصد است.



شکل ۷- مقایسه نتایج حاصل از مدل ماشین بردار پشتیبان و داده‌های تاریخی بارش در ماه‌های ژوئن تا نوامبر (راست) و ماه‌های دسامبر تا می (چپ)

ماشین بردار پشتیبان نتایجی بهتر از مدل رگرسیون چند متغیره در مرحله آموزش و آزمون دارد. همچنین این مدل دارای خطای مدل سازی کمتری نسبت به مدل رگرسیون چند متغیره است که نشان دهنده کارایی آن در مدل سازی غیر خطی است.

۴- نتیجه گیری

در این پژوهش، بارش با استفاده از الگوهای پیوند دور اقلیمی مؤثر بر اقلیم شمال غرب ایران در طول یک دوره ۴۰ ساله برای دو فصل تر و خشک پیش بینی شد. تعداد بهینه ورودی مدل پیش بینی با استفاده از روش های همبستگی و روش آزمون گاما و بهترین ترکیب آن با استفاده از روش آزمون گاما تعیین شد. سپس بارش با استفاده از مدل ماشین بردار پشتیبان پیش بینی شد و نتایج حاصل از آن با نتایج مدل رگرسیون چند متغیره مقایسه شد. نتایج نشان می دهد، مدل ماشین بردار پشتیبان، هم در مرحله آموزش و هم در مرحله آزمون دارای خطای کمتری در پیش بینی مقادیر بارندگی نسبت به سایر مدل ها است. در فصل تر در صورتی که ورودی های مدل با استفاده از روش آزمون گاما انتخاب شوند، مدل، خطاهای

شبه سازی کمتری نسبت به انتخاب با روش رگرسیون در برآورد میزان بارندگی منطقه دارد. بنابراین مؤثرترین الگوهای پیوند دور اقلیمی ها بر بارندگی فصل تر (ماه های دسامبر تا می) شامل الگوهای پیوند دور اقلیمی های گرادیان فشار بین جنوب سبیری و سودان (پیش بینی کننده شماره ۵) و شرق خلیج فارس (پیش بینی کننده شماره ۶)، و دمای سطح آب در دریای عرب (پیش بینی کننده شماره ۱۳) و شرق خلیج فارس (پیش بینی کننده شماره ۱۵) است. همچنین در فصل خشک، تعیین پیش بینی کننده ها با استفاده از روش رگرسیون بهترین مدل در مرحله آموزش و آزمون بوده است. مؤثرترین الگوهای پیوند دور اقلیمی ها برای بارندگی فصل خشک (ماه های ژوئن تا نوامبر) عبارت اند از گرادیان فشار بین جنوب گرینلند و شرق مدیترانه (پیش بینی کننده شماره ۳)، گرادیان فشار بین جنوب سبیری و سودان (پیش بینی کننده شماره ۵) و دمای سطح آب در دریای سیاه (پیش بینی کننده شماره ۹) و دریای مازندران (پیش بینی کننده شماره ۷) است. بنابراین با استفاده از مدل های تدوین شده می توان بارش شش ماهه فصل تر و خشک سال را بر اساس متغیرهای اقلیمی شش ماه گذشته پیش بینی نمود.

۵- مراجع

1. Araghinejad, S., and Burn, D.H. (2005). "Probabilistic forecasting of hydrological events using gestatistical analysis." *Hydrol. Sci. J.*, 50(5), 838-856.
2. Ghosh, S., and Mujumdar, P.P. (2008). "Statistical downscaling of GCM simulations to streamflow using relevance vector machine." *Adv. Water Resour.*, 31, 132-146.
3. Najafi, M., Moradkhani, H., and Wherry, S. (2011). "Statistical downscaling of precipitation using machine learning with optimal predictor selection." *J. Hydrol. Eng.*, 16(8), 650-664.
4. Nagesh Kumar, D., Janga Reddy, M., and Maity, R. (2007). "Regional rainfall forecasting using large scale climate teleconnections and artificial intelligence techniques." *J. Intell. Syst.*, 16(4), 307-322.
5. Omogbai, B.E. (2010). "Prediction of northern Nigeria rainfall using sea surface temperature." *J. Hum Ecol.*, 32(2), 127-133.
6. Karamouz, M. (2009). *Model and software preparation for long term rainfall forecasting of area impacted by large scale climate signals*, Iranian Meteorological Organization. (In Persian)
7. Karamouz, M. (2004). *Development of long lead rainfall prediction models based on climate patterns of large scale climate signals*, Iranian Meteorological Organization. (In Persian)
8. Nikbakht Shahbazi, A., Zahraie, B., and Nasserli, M. (2011). "Seasonal meteorological drought prediction using support vector machine." *J. of Water and Wastewater*, 75, 73-85. (In Persian)
9. Noori, R., Farokhnia, A., Morid, S., and Riahi Madvar, H. (2008). "Effect of variables preprocessing in artificial neural network on monthly flow prediction by PCA and wavelet transformation." *J. of Water and Wastewater*, 69, 13-22. (In Persian)
10. Nazemosadat, M.J., and Shirvani, A. (2005). "Prediction of persian gulf sea surface temperature using multiple regressions and principal components analysis." *JWSS - Isfahan University of Technology*, 9 (3), 1-11. (In Persian).

11. Schoof, J.T., and Pryor, S.C. (2001) "Downscaling temperature and precipitation: A comparison of regression-based methods and artificial neural networks". *Int. J. Climatol.*, 21, 773-790.
12. Ahmadi, A., Han, D., Karamouz, M., and Remesan, R. (2009). "Input data election for solar radiation estimation." *J. Hydrol Process*, 23, 2754-2764.
13. Ahmadi, A., and Han, D. (2013). "Identification of dominant sources of sea level pressure for precipitation forecasting over Wales." *J. Hydroinform*, doi:10.2166/hydro.2012.110
14. Moghaddamia, A., Ghafari, M., Piri, J., and Han, D. (2009). "Evaporation estimation using support vector machines technique." *Int. J. Eng., Sci.*, 5 (7), 415-423.
15. Noori, R., Khakpour, A., Dehghani, M., and Farokhnia, A. (2010). "Monthly stream flow prediction using support vector machine based on principal component analysis." *J. of Water and Wastewater*, 80, 118-123. (In Persian)
16. Noori, R., Karbassi, AR., Moghaddamia, A., Han, D., Zokaei-Ashtiani, MH., Farokhnia, A., and Ghafari Gousheh, M. (2011). "Assessment of input variables determination on the SVM model performance using PCA, Gamma test and forward selection techniques for monthly stream flow prediction." *J. Hydrol.*, 401, 177-189.
17. Dastorani, M.T., Sharifi Darani, H., Talebi, A., and Moghaddam Nia, A.R. (2011). "Evaluation of the application of artificial neural networks and adaptive neuro-fuzzy inference for rainfall-runoff modeling in Zayandeh_rod dam basin." *J. of Water and Wastewater*, 80, 114-125. (In Persian)
18. Karamouz, M. (2009). *Re-evaluation of water supply and demand and drought management in Ahar-Chaie river basin*, Azarbayjan Regional Water Authority. (In Persian)
19. Kakaie Lafdani, E., Moghadam Nia, A.R., and Ahmadi, A. (2012). "Daily suspended sediment load prediction using artificial neural networks and support vector machines." *J. Hydrol.*, 478 (25), 50-62.
20. Stefanesson, A., Koncar, N., and Jones, A.J. (1997). "A note on the gamma test." *Neural Comput Appl.*, 5, 131-133.
21. Durrant, P.J. (2001). "WinGamma: A non-linear data analysis and modelling tool with applications to flood prediction." PhD. Thesis, Department of Computer Science, Cardiff University, Wales, UK.
22. Jones, A.J., Tsui, A., and de Oliveira, A.G. (2002). "Neural models of arbitrary chaotic systems: construction and the role of time delayed feedback in control and synchronization." *Complex Int.*, 9, 1-9.
23. Vapnik V. (1995). *The nature of statistical learning theory*, Springer, New York.
24. Bray, M., and Han, D. (2004). "Identification of support vector machines for runoff modeling." *J. Hydroinform*, 6(4), 265-280.
25. <http://www.csie.ntu.edu.tw/~cjlin/libsvm/>