

استفاده از رویکرد ترکیبی گاما- ماشین بردار پشتیبان به منظور پیش بینی بلندمدت آورد رودخانه
در حوضه زرینه رود

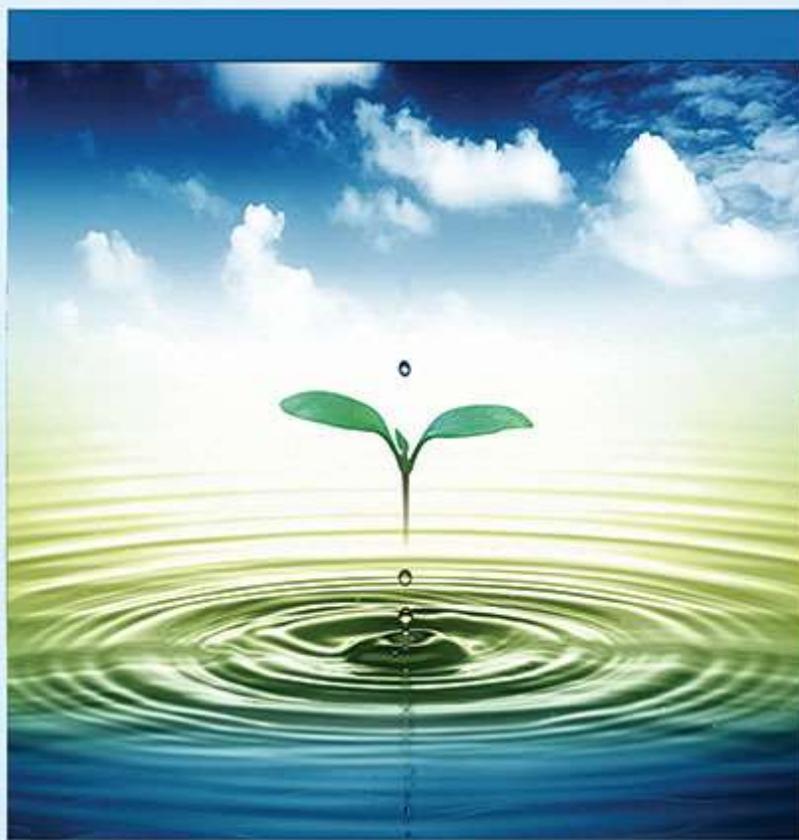
عباس عباسی، مجید دلاور و سعید مرید

دوره ۳، شماره ۴، زمستان ۱۳۹۶، صفحات ۳۶۶-۳۵۳

Vol. 3(4), Winter 2018, 353 – 366

**Application of Hybrid Gamma-SVM
Approach for River Flow Prediction in
Zarinehrud Basin**

Abbasi A., Delavar M. and Morid S.



www.jewe.ir

OPEN ACCESS

ارجاع به این مقاله: عباسی ع.، دلاور م.، مرید س. (۱۳۹۶). استفاده از رویکرد ترکیبی گاما- ماشین بردار پشتیبان به منظور پیش بینی بلندمدت آورد رودخانه در حوضه زرینه رود. محیط زیست و مهندسی آب، دوره ۳، شماره ۴، صفحات: ۳۶۶ – ۳۵۳

Citing this paper: Abbasi A., Delavar M. and Morid S. (2018). Application of a hybrid Gamma-SVM approach for river flow prediction in Zarinehrud basin. J. Environ. Water Eng., 3(4), 353 – 366.

استفاده از رویکرد ترکیبی گاما- ماشین بردار پشتیبان به منظور پیش‌بینی بلندمدت آورد رودخانه در حوضه زربینه رود

عباس عباسی^۱، مجید دلاور^{۳*} و سعید مرید^۴

^۱ کارشناسی ارشد، گروه مهندسی منابع آب، دانشکده کشاورزی، دانشگاه تربیت مدرس، تهران، ایران

^۲ دانشجوی دکتری، گروه مهندسی منابع آب، دانشکده کشاورزی، دانشگاه ارومیه، ارومیه، ایران

^۳ استادیار گروه مهندسی منابع آب، دانشکده کشاورزی، دانشگاه تربیت مدرس، تهران، ایران

^۴ استاد گروه مهندسی منابع آب، دانشکده کشاورزی، دانشگاه تربیت مدرس، تهران، ایران

* نویسنده مسئول: m.delavar@modares.ac.ir

مقاله اصلی

تاریخ دریافت: [۱۳۹۶/۰۸/۲۱]

تاریخ پذیرش: [۱۳۹۶/۰۹/۲۸]

چکیده

پیش‌بینی آورد رودخانه از موضوعات مهم در برنامه‌ریزی منابع آب و مدیریت عرضه و تقاضای آن محسوب می‌شود. از این رو در طراحی، برنامه‌ریزی، مدیریت و بهره‌برداری از تأسیسات آبی و همچنین مدیریت شرایط بحرانی از قبیل سیلاب و خشکسالی مورد توجه پژوهشگران است. در تحقیق حاضر با استفاده از یک رویکرد ترکیبی بر مبنای آزمون گاما و مدل ماشین بردار پشتیبان (GSVM)، پیش‌بینی میزان آورد، بررسی شد. بدین منظور با استفاده از آزمون گاما از ترکیبات مختلف ۱۰ متغیر هواشناسی و هیدرولوژیکی در سطح حوضه بهترین ترکیب ممکن برای پیش‌بینی آورد انتخاب شد. سپس با توجه به بهترین ترکیب پیش‌بینی‌کننده‌ها، آورد پتانسیل رودخانه با استفاده از ماشین بردار پشتیبان پیش‌بینی شد. مقایسه مقادیر پیش‌بینی‌شده و مشاهداتی بیانگر کارایی مطلوب رویکرد ترکیبی در پیش‌بینی آورد طبیعی حوضه و تنظیم برنامه مدیریتی حوضه بر مبنای آن است. براین اساس دقت کلی مدل در پیش‌بینی سطوح مدیریتی خشکسالی بر مبنای آورد رودخانه زربینه رود برابر با ۷۱/۴ درصد، خطای برآورد دست بالا برابر ۸/۲ درصد و خطای برآورد دست پایین نیز ۲۰/۴ است. این نتایج دقت قابل قبول مدل GSVM در پیش‌بینی آورد در شرایط مختلف هیدرولوژیکی حوضه را نشان داد.

واژه‌های کلیدی: آزمون گاما؛ پیش‌بینی جریان؛ زربینه رود؛ ماشین بردار پشتیبان

۱- مقدمه

پیش‌بینی جریان به‌عنوان یکی از چالش‌های مدیریت منابع آب به‌ویژه در زمینه برنامه‌ریزی و تخصیص منابع در دهه‌های اخیر است. به‌طوری که محققان روش‌های مختلفی را برای این امر در مقالات مختلف به کار گرفته‌اند. این روش‌ها طیف متنوعی از مدل‌های سری زمانی و داده‌محور را همچون ARMA و ARIMA (Kelmes 1973)، شبکه‌های عصبی مصنوعی (Silverman and Darcup 2000)، برنامه‌ریزی ژنتیک (Guven 2009)، نروفازی (Nayak et al. 2004)، مدل‌های ماشین بردار پشتیبان (SVM¹) و لگاریتم خطی شامل می‌شوند. نخستین کاربردهای SVM در زمینه منابع آب را می‌توان به تحقیقات (Dibike et al. 2001) نسبت داد که با استفاده از SVM پیش‌بینی هیدرولوژیکی و همچنین مدل‌سازی فرآیند بارش-رواناب را انجام دادند. Liong and Sivapragasam (2002) موفق به کاربرد SVM در پیش‌بینی پهنه‌بندی سیلاب شهر داکا در بنگلادش شدند و توانایی پیش‌بینی SVM در تخمین تراز سیل، در قالب دوره‌های ۱ تا ۷ روزه را بررسی کردند. نتایج کار آنها نشان داد که روش SVM برای پیش‌بینی درازمدت، نتایج بهتری نسبت به ANN در هر دو مرحله آموزش و آزمون ارائه می‌دهد. (et al. 2006) Asefa توانایی SVM را در طراحی بهینه شبکه نظارت آب زیرزمینی بر اساس ماشین‌های بردار پشتیبان آزمایش کردند. در تحقیق آنها مقادیر بهینه پارامترها به‌وسیله روش جستجوی شبکه تعیین و تابع هسته‌ای RBF مورد استفاده قرار گرفت. Behzad et al. (2009) نیز برای پیش‌بینی رواناب روزانه در رودخانه بختیاری، مدل SVM را با بهره گرفتن از مسئله بهینه‌سازی دوگانه و استفاده از داده‌های هواشناسی و بارندگی بکار گرفتند. آن‌ها

نتایج مدل را با مدل ANN و ANN-GA مقایسه و به این نتیجه رسیدند که روش ماشین بردار پشتیبان، خطای کمتری نسبت به دو روش دیگر دارد و از سرعت و دقت بالاتری برخوردار است. (Noori et al. 2011) با استفاده از ۱۸ متغیر ورودی دبی را پیش‌بینی کردند و سپس با استفاده از PCA تعداد متغیرهای ورودی را از ۱۸ به ۵ کاهش دادند. یافته‌های این تحقیق نشان داد که پیش‌پردازش متغیرهای ورودی با استفاده از PCA، بهبود عملکرد مدل SVM را به همراه داشته است. در خصوص پیش‌بینی بلندمدت جریان رودخانه‌ها می‌توان به تحقیق Kiani Flaverjani et al. (2011) اشاره کرد که پیش‌بینی رواناب ورودی به مخزن سد زاینده‌رود را با استفاده از ترکیب سیگنال‌های بزرگ مقیاس اقلیمی و روش SVM انجام دادند که نتایج این تحقیق بیانگر دقت قابل قبول روش SVM در پیش‌بینی بلندمدت رواناب می‌باشد. (Chu et al. 2016) نیز جهت پیش‌بینی بلندمدت و میان‌مدت رواناب در بالادست رودخانه زرد چین از روش رگرسیون بردار پشتیبان (SVR) استفاده کردند و نتایج آن را با روش AR و روش شبکه عصبی مصنوعی با تابع پایه هسته شعاعی (RBFNN) مقایسه نمودند. آن‌ها به این نتیجه رسیدند که روش SVR نسبت به دو روش دیگر از دقت بیشتری در پیش‌بینی بلندمدت و میان‌مدت رواناب برخوردار است. (Yilin et al. 2016) جهت پیش‌بینی بلندمدت جریان رودخانه از روش SVM استفاده و نتایج آن را با دو روش ARMA و ANN مقایسه و به این نتیجه رسیدند که روش SVM نسبت به دو روش دیگر در پیش‌بینی بلندمدت جریان رودخانه از عملکرد بهتری برخوردار است. (Remesan et al. 2008) با بهره‌گیری از آزمون گاما مهمترین عوامل تأثیرگذار بر میزان تابش روزانه

¹Support Vector Machine

انتخاب و کارایی مدل به‌منظور به‌کارگیری نتایج آن در برنامه مدیریت منابع آب حوضه مورد بررسی قرار گرفت.

۲- مواد و روش‌ها

۱-۲ - منطقه مورد مطالعه و داده‌های تحقیق

رودخانه زرینه‌رود با وسعت حوزه آبخیز ۱۱۰۰۰ کیلومتر مربع، یکی از مهمترین و طویل‌ترین رودخانه‌های حوزه آبخیز دریاچه ارومیه می‌باشد. زرینه‌رود در میان جریان‌های ورودی به دریاچه با بیشترین تخلیه آب، از اهمیت بالایی برخوردار است. شهرهای میاندوآب، شاهین‌دژ، تکاب و سقز از کانون‌های مهم شهری این حوضه می‌باشند. طول تقریبی رودخانه، ۳۰۰ کیلومتر است و از کوه‌های چهل چشمه واقع در استان کردستان سرچشمه گرفته و به‌موازات رودخانه سیمینه‌رود در جهت شمال جریان دارد. این رودخانه در نهایت وارد دریاچه ارومیه می‌شود. شکل (۱) موقعیت حوضه رودخانه زرینه‌رود را نشان می‌دهد. مقادیر پتانسیل آورد سالانه، بارش سالانه، متوسط دمای نقطه شب‌م سالانه، متوسط درصد رطوبت نسبی سالانه و دمای متوسط سالانه از جمله اطلاعات مورد استفاده در این تحقیق می‌باشد. اطلاعات هواشناسی از ایستگاه سینوپتیک سقز از سازمان هواشناسی کشور دریافت گردید و مقادیر پتانسیل آورد سالانه بر اساس طرح مدیریت خشکسالی دریاچه ارومیه (WGSMWRA 2012) مدنظر قرار گرفت. دوره آماری اطلاعات مورد استفاده در این مطالعه از سال ۱۳۳۶ تا ۱۳۸۵ می‌باشد.

خورشیدی در حوزه آبخیز برو در انگلستان را تعیین نمودند. (Moghaddamnia et al. (2009a) به‌منظور تعیین بهترین ترکیب از متغیرهای ورودی به مدل‌های شبکه عصبی مصنوعی و فازی جهت برآورد تبخیر روزانه دریاچه چاه نیمه زابل از آزمون گاما استفاده نمودند. نتیجه این تحقیق نشان داد که آزمون گاما از توانایی لازم در تعیین متغیرهای ورودی به مدل‌های پیش‌بینی برخوردار است. Wan Jaafar et al. (2011) از دو روش آزمون گاما و LOOCV^۱ برای انتخاب بهترین ترکیب متغیرهای ورودی جهت تخمین حداکثر جریان سالانه در جنوب غربی انگلستان استفاده و نشان داد آزمون گاما از قابلیت بالاتری در محاسبه واریانس داده‌های خروجی بدون نیاز به ایجاد یک مدل شبیه‌ساز برخوردار است. (Sharifi et al. (2013) نیز با استفاده از آزمون گاما از هشت متغیر ورودی که شامل بارش و رواناب با تأخیرهای زمانی ۱ تا ۳ ساله می‌باشد، ترکیب بهترین متغیرهای ورودی برای شبیه‌سازی رواناب در حوزه آبخیز امامه را تعیین نمودند.

(Kakaei-Lefdani et al. (2014) جهت پیش‌بینی حجم رسوبات معلق رودخانه دویرج واقع در استان ایلام از ترکیب روش آزمون گاما و SVM استفاده کردند که نتایج این تحقیق نشان داد که پیش‌پردازش متغیرهای ورودی با استفاده از آزمون گاما سبب افزایش دقت مدل پیش‌بینی SVM در هر دو مرحله آموزش و آزمون می‌شود. در این مطالعه بر اساس تلفیق نتایج حاصل از آزمون گاما و ماشین بردار پشتیبان، رویکردی ترکیبی به‌منظور بهبود نتایج پیش‌بینی آورد رودخانه‌ها ارائه شد. بدین منظور رودخانه زرینه‌رود به‌عنوان بزرگترین و مهم‌ترین سیستم آبی حوزه آبخیز دریاچه ارومیه به‌عنوان منطقه مطالعاتی جهت ارزیابی روش‌شناسی

¹ Leave-On-Out Cross Validation



شکل ۱- موقعیت منطقه مورد مطالعه

همراه خواهد بود. جهت محاسبه متغیرهای w و b بایستی تابع خطا در رابطه (۲) با لحاظ کردن قیود زیر به سمت کمترین مقدار خود بهینه شود (Hamel 2009).

$$\begin{aligned} \text{Minimize } & \frac{1}{2} w^T \cdot w + c \sum_{i=1}^n \xi_i + c \sum_{i=1}^n \bar{\xi}_i \\ & w^T \cdot \phi(x_i) + b - Y_i \leq \varepsilon + \xi_i \quad (2) \\ & Y_i - w^T \cdot \phi(x_i) - b \leq \varepsilon + \bar{\xi}_i \\ & \xi_i, \bar{\xi}_i \geq 0, i = 1, \dots, N \end{aligned}$$

در رابطه (۲) ثابت گنجایش^۱، ϕ تابع هسته‌ای یا کرنل، ξ_i و $\bar{\xi}_i$ ضرایب کمبود^۲ که حد بالا و پایین خطای آموزش مرتبط با مقدار خطای مجاز ε را تعیین می‌کنند. با استفاده از ضرایب لاگرانژ a_i و a_i^* مسئله بهینه‌سازی با سادگی بیشتری در فرم دوگانه به وسیله کاربرد الگوریتم برنامه نویسی درجه دوم مطابق رابطه (۳) حل شود.

بعد از اینکه ضرایب لاگرانژ a_i و a_i^* تعیین شدند می‌توان بردارهای پارامتر w و b را تحت شرایط KKT^۳ تعیین کرد که در آن $w = \sum_{i=1}^N \phi(x_i) \cdot (\alpha_i - \alpha_i^*)$. با توجه به اینکه ترم‌های لاگرانژ $(\alpha_i - \alpha_i^*)$ می‌تواند مقادیر صفر و غیر صفر را لحاظ کنند بنابراین تنها مجموعه داده‌هایی که ضرایب $\bar{\alpha}_i$ آنها غیر صفر باشد در معادله نهایی وارد می‌شوند. تابع SVM رگرسیونی به صورت رابطه (۴) بیان می‌شود (Yu et al. 2006).

۲-۲- ماشین بردار پشتیبان SVM

ماشین بردار پشتیبان یک مدل داده محور بر پایه یادگیری با نظارت است که در مسائل مربوط به طبقه‌بندی و رگرسیون از آن استفاده می‌شود و مبنای آن تئوری جداسازی خطی می‌باشد. رویکرد SVM به این صورت است که در فاز آموزش، سعی می‌شود که مرز تصمیم‌گیری به گونه‌ای انتخاب شود که حداقل فاصله آن با هر یک از دسته‌های دیگر بیشترین مقدار را به خود بگیرد که این نحوه انتخاب مرز بر اساس نقاطی به نام بردارهای پشتیبان انجام می‌شود. در مدل رگرسیونی SVM، تابعی از متغیر وابسته Y که خود تابعی از چندین متغیر مستقل x می‌باشد برآورد می‌شود. فرض می‌شود که رابطه بین متغیرهای وابسته و مستقل مطابق رابطه (۱) با تابع جبری $f(x)$ به علاوه مقداری اغتشاش تحت عنوان خطای مجاز (ε) تعریف می‌شود.

$$Y = f(x) + \varepsilon \quad (1)$$

$$f(x) = w^T \cdot \phi(x) + b$$

که در این رابطه w بردار ضرایب و b ثابت مشخصه‌های تابع رگرسیونی و ϕ نیز تابع هسته‌ای می‌باشد. هدف از رابطه (۱) جستجو برای تعیین فرم تابع $f(x)$ می‌باشد که بدین منظور با استفاده از داده‌هایی که برای آموزش مدل لحاظ می‌گردند این امر محقق و این روند با بهینه‌سازی متوالی تابع خطا

³ Karush-Khuan-Tucker

¹ Capacity Cinstant

² Slack Variables

می‌شود. با استفاده از آماره گاما می‌توان ترتیب اهمیت پارامترهای ورودی، بهترین ترکیب از میان ترکیبات ممکن و تعداد داده‌های لازم برای آموزش مدل را به‌دست آورد. با در نظر گرفتن دو نمونه (X, Y) و (X', Y') که X' نزدیک‌ترین همسایه به X است در نظر گرفتن این مسئله که Y' و Y باید به هم نزدیک باشند، قابل قبول است. آزمون گاما بر اساس $N[i, p]$ می‌باشد، که p امین $(1 \leq k \leq P)$ نزدیک‌ترین همسایه برای هر بردار ورودی X یا شرایط $(1 \leq k \leq M)$ می‌باشد. p حداکثر تعداد همسایگی‌ها است که معمولاً ۱۰ تا ۵۰ در نظر گرفته می‌شود. در این صورت $\delta(p)$ که میانگین مربعات فاصله از همین همسایگی می‌باشد به صورت رابطه (δ) تعریف می‌شود؛ که در آن $LN[i, p]$ طول $N[i, p]$ می‌باشد. بر این اساس آزمون گاما برای M سری داده مطابق رابطه (ϵ) انجام می‌شود:

$$1 \leq k \leq P \delta_M(p) = \frac{1}{M} \sum_{i=1}^M \frac{1}{LN[i, p]} \sum_{j=N[i, p]} |x_i - x_j|^2 = \frac{1}{M} \sum_{i=1}^M |x_i - x_{N[i, p]}|^2 \quad (5)$$

$$\gamma_M(p) = \frac{1}{2M} \sum_{i=1}^M \frac{1}{LN[i, p]} \sum_{j=N[i, p]} |y_i - y_j|^2 = \frac{1}{2M} \sum_{i=1}^M |y_i - y_{N[i, p]}|^2 \quad 1 \leq k \leq P \quad (6)$$

آزمون M تست با فرض آنکه M تعداد سری مجموعه مشاهده شده از پدیده مورد بررسی باشد، با استفاده از آزمون گاما می‌توان چگونگی تغییرات منحنی‌های مجانب گاما و خطای استاندارد را برای ترکیب منتخب با افزایش تعداد الگوهای ورودی M بررسی و تعداد الگوهای ورودی که مقدار این آماره‌ها را به حالت پایدار برای ایجاد مدل هموار می‌رساند، تعیین نمود (Sharifi et al. 2013).

۲-۴- آماره‌های ارزیابی

در این تحقیق به منظور ارزیابی مدل‌های مورد نظر از شاخص‌های کمی ضریب تعیین (R^2)، مجذور میانگین مربعات خطا (RMSE) و میانگین مطلق خطا (MAE) استفاده شد که نحوه محاسبه آنها در روابط (۸) تا (۹) به ترتیب آورده شده است.

$$\sum_{i=1}^N y_i (\alpha_i - \alpha_i^*) - \epsilon \sum_{i=1}^N (\alpha_i + \alpha_i^*) - \frac{1}{2} \sum_{i=1}^N (\alpha_i - \alpha_i^*) (\alpha_j - \alpha_j^*) \phi(x_i)^T \cdot \phi(x_j) \quad (3)$$

$$\sum_{i=1}^N y_i (\alpha_i - \alpha_i^*) = 0$$

$$0 \leq \alpha_i \leq C$$

$$0 \leq \alpha_i^* \leq C, i = 1, 2, \dots, N$$

$$f(x) = \sum_{i=1}^N \bar{\alpha}_i \cdot \phi(x_i)^T \cdot \phi(x) + b \quad (4)$$

۲-۳- آزمون گاما

آزمون گاما از جمله روش‌های آماری ناپارامتری می‌باشد که در مجموعه‌ای از داده‌های ورودی-خروجی برآوردی از بهترین میانگین مربعات خطا روی خروجی به وسیله یک مدل هموار انجام می‌دهد. این برآورد آماره گاما (Γ) نامیده

که $yN[i, p]$ مقدار خروجی نظیر y برای p امین نزدیک‌ترین همسایه برای بردار x_i می‌باشد. در این روش با تعیین مقدار نزدیک‌ترین همسایگی p ، مقادیر $\delta_M(p)$ و $\gamma_M(p)$ از روی داده‌های ورودی و خروجی محاسبه می‌شود. سپس با ایجاد رابطه رگرسیون خطی بین P مجموعه $\delta_M(p)$ و $\gamma_M(p)$ معادله آن از رابطه (γ) به‌دست می‌آید.

$$\gamma = A\delta + \Gamma \quad (7)$$

مقدار آماره Γ و A به ترتیب عرض از مبدأ و شیب خط رگرسیون فوق می‌باشد. همچنین مقدار خطای استاندارد (SE) آن هرچه به عدد صفر نزدیکتر باشد با اطمینان بیشتری می‌توان مقدار گاما را نشان‌دهنده اغتشاش موجود در داده‌ها دانست (Moghaddamnia et al. 2009b).

$V_{ratio} = \frac{\Gamma}{Var(\gamma)}$ معیار ثابتی است که به صورت نرمال بین صفر تا یک متغیر می‌باشد. در واقع $|V_{ratio}| - 1$ مشابه با ضریب تعیین R^2 می‌باشد که تخمین می‌زند تا چه حد اطلاعات متناسب با مدل می‌باشد (Jones 2004).

$$OA = \frac{\sum_{i=0}^4 X_{ii}}{N} \quad (13)$$

$$UA_i = \frac{X_{ii}}{\sum_{i=0}^4 X_{ij}} \quad j = 0, \dots, 4 \quad (14)$$

$$PA_i = \frac{X_{ij}}{\sum_{i=0}^4 X_{ij}} \quad i = 0, \dots, 4 \quad (15)$$

طوری که N تعداد کل مشاهدات در طول دوره می‌باشد. در اینجا دقت استفاده کننده (UA) از نظر مفهوم برابر با دقت پیش‌بینی مدل می‌باشد. X_{ij} یعنی تعداد دفعاتی که کلاس آورد در وضعیت واقعی i قرار داشته و مدل پیش‌بینی نیز به‌درستی آن را i اعلام کرده است. X_{ij} یعنی تعداد دفعاتی که کلاس آورد در سطح i قرار داشته و وضعیت j که مربوط به کلاس دیگری بوده است را اعلام کرده است. در واقع اگر $i=j$ پیش‌بینی کلاس درست بوده است. همچنین خطای برآورد دست پایین (UE) و خطای برآورد دست بالا (OE) با توجه به روابط (۱۶) و (۱۷) تعیین می‌شوند (Gholamzade et al. 2011):

$$UE = \frac{\sum \forall i \sum \forall i X_{ij}}{N} \quad i < j \quad (16)$$

$$OE = \frac{\sum \forall i \sum \forall i X_{ij}}{N} \quad i > j \quad (17)$$

۳- یافته‌ها و بحث

در مطالعه حاضر از روش ماشین بردار پشتیبان به‌منظور پیش‌بینی جریان برای گام زمانی سالانه استفاده شد. بدین منظور ۵ متغیر هواشناسی و هیدرولوژیکی در دوره آماری (۱۳۸۵-۱۳۳۶) شامل بارش سالیانه، دمای متوسط سالیانه، متوسط رطوبت نسبی سالانه، متوسط دمای نقطه شبنم سالانه و میزان آورد طبیعی سالانه استفاده شدند.

این متغیرها به ترتیب با ۳، ۳، ۱، ۱ و ۲ سال تأخیر زمانی که در مجموع ۱۰ متغیر ورودی را تشکیل می‌دادند، برای پیش‌بینی آورد منطقه مورد مطالعه مورد بررسی قرار گرفتند. این متغیرها پس از استانداردسازی به‌منظور کاربرد در مدل‌های پیش‌بینی ارزیابی شدند.

$$R^2 = \left(1 - \frac{\sum_{i=1}^n (P_i - \hat{Q}_i)^2}{\sum_{i=1}^n (Q_i - \hat{Q}_i)^2} \right) \quad (8)$$

$$RMSE = \left(\frac{\sum_{i=1}^n (Q_i - \hat{Q}_i)^2}{n} \right)^{0.5} \quad (9)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |Q_i - \hat{Q}_i| \quad (10)$$

در این روابط Q_i مقدار داده‌های مشاهده‌ای در گام زمانی i ام، \hat{Q}_i مقدار داده پیش‌بینی شده در همان زمان، n تعداد داده‌ها و \bar{Q}_i میانگین مقادیر مشاهده‌ای می‌باشد. هر چه مقادیر R و E به یک نزدیکتر باشد و $RMSE$ به صفر نزدیکتر باشد، نتایج دقیق‌تر می‌باشد. در این تحقیق جهت نشان دادن توزیع خطا از دو شاخص میانگین قدر مطلق خطای نسبی (MARE) و تحلیل آستانه خطا^۲ (TS) استفاده گردید. شاخص TS برای مقادیر x درصد از برآوردها، نشان دهنده توزیع خطا در مقادیر برآورد شده برای هر مدل می‌باشد. این شاخص که بر حسب درصد تعریف می‌شود برای مقادیر مختلف قدر مطلق خطای نسبی ارائه می‌شود. مقدار TS برای x درصد از برآوردها طبق رابطه (۱۱) و شاخص MARE از رابطه (۱۲) به‌دست می‌آیند:

$$TS_x = \frac{Y_x}{n} \times 100 \quad (11)$$

$$MARE = \left| \frac{O_i - E_i}{O_i} \right| \times 100 \quad (12)$$

که در آن Y_x تعداد پیش‌بینی شده (از کل تعداد n) برای هر مقدار مطلق خطای نسبی کمتر از x درصد، O_i معادل مقادیر واقعی و E_i مقادیر برآورد شده می‌باشد (Riahi et al. 2009).

۲-۵- تحلیل خطای پیش‌بینی

به‌منظور ارزیابی مدل و تحلیل خطای پیش‌بینی آن در این مطالعه از سه معیار ارزیابی شامل: دقت کلی^۳ (OA)، آمین دقت استفاده کننده^۴ (UA_i) و آمین دقت تولید شده^۵ (PA_j) استفاده شد که این معیارها به ترتیب به‌صورت روابط (۱۳)، (۱۴) و (۱۵) محاسبه می‌شوند:

⁴ User's Accuracy

⁵ Producer's Accuracy

¹ Mean Absolute Relative Error

² Threshold Statistics

³ Overall accuracy

۳-۱- انتخاب ترکیب‌ها برای پیش‌بینی آورد

برای رودخانه زرینه‌رود با در نظر گرفتن ده متغیر ورودی مؤثر برای پیش‌بینی آورد، تعداد ۱۰۰ ترکیب مختلف تولید شد. با استفاده از آزمون گاما با بررسی ترکیبات مختلف ایجاد شده، ترکیبی که دارای کم‌ترین مقدار آماره گاما، گرادیان، خطای استاندارد و VRatio باشد، تعیین شد. جدول (۱) نتایج آزمون گاما برای ۱۰ ترکیب که بر اساس اولویت مرتب شده‌اند را نشان می‌دهد. بر اساس آماره‌های موجود در جدول (۱)، بهترین ترکیب، ترکیبی است که شامل ۷ پارامتر ورودی شامل بارش با یک و سه تأخیر زمانی، رطوبت نسبی و دمای نقطه شبنم و همچنین آورد با یک تأخیر زمانی، دما با دو و سه سال تأخیر زمانی باشد زیرا این ترکیب کمترین مقادیر معیارهای ارزیابی را کسب کرده است. با محاسبه مقادیر دلتا و گاما از روی پارامترهای ورودی و خروجی، پراکنش آنها حول خط رگرسیون برای ترکیب منتخب (ترکیب شماره ۱) در شکل (۲) رسم شده است.

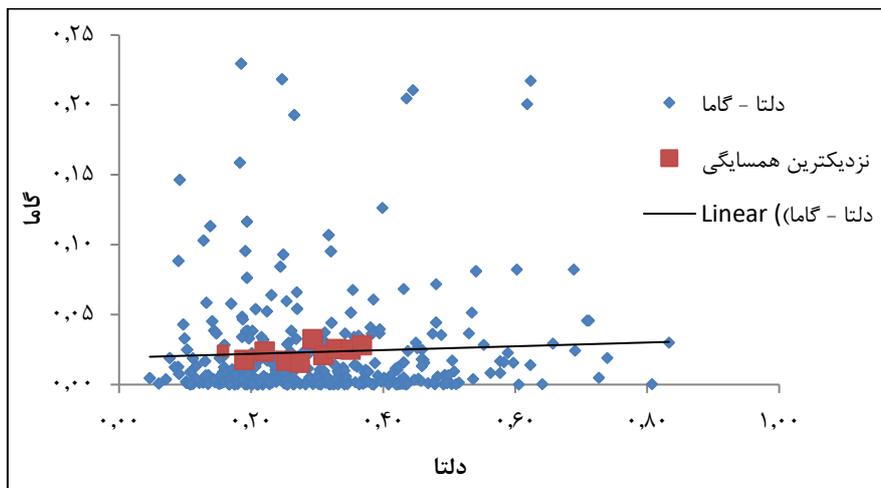
برای مشخص نمودن درجه اهمیت پارامترهای ورودی، ابتدا آزمون گاما برای بهترین ترکیب (ترکیب شامل هفت متغیر ورودی) انجام شد. سپس در هر مرحله یکی از ۷ پارامتر ورودی به‌دلخواه از مجموعه اولیه حذف و آزمون گاما با ۶ پارامتر باقی‌مانده محاسبه شد. این فرآیند برای تمام متغیرها به ترتیب انجام گرفت و در هر مرتبه مقدار آماره آزمون گاما محاسبه گردید. نتایج این آزمون برای پیش‌بینی آورد در جدول (۲) آورده شده است.

با توجه به جدول (۲) با حذف پارامتر بارش با یک تأخیر زمانی، آماره گاما دارای بیشترین مقدار می‌گردد. بنابراین

بارش با یک تأخیر مؤثرترین پارامتر برآورد می‌باشد و پارامترهای دمای نقطه شبنم با یک تأخیر، رطوبت نسبی با یک تأخیر، دما با دو تأخیر، دما با سه تأخیر، بارش با سه تأخیر و دبی با یک تأخیر، به ترتیب در رده‌های بعدی اهمیت قرار دارند. در شکل (۳) نتایج M-test برای بهترین ترکیب ورودی نشان داده شده است. در این شکل می‌توان مشاهده نمود که منحنی‌های مجانب آماره‌های گاما تقریباً در اطراف نقطه ۳۰ به حالت پایدار میل می‌کند و با افزایش تعداد داده‌ها این منحنی‌ها تغییرات چندانی به خود نمی‌بیند. بنابراین تعداد الگوهای لازم برای آموزش مدل SVM برابر با ۳۰ می‌باشد و از الگوی باقیمانده برای ارزیابی مدل‌های ایجاد شده با استفاده از آماره‌های خطا استفاده می‌شود.

جدول ۱- نتایج آزمون گاما

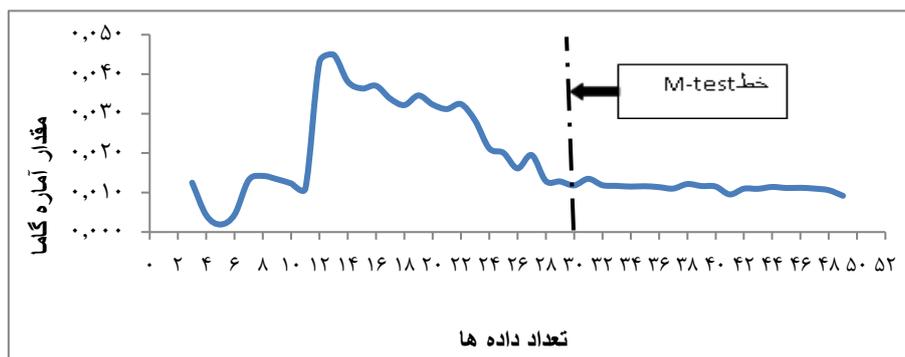
ردیف ترکیب	گاما (I)	VRatio	تأخیر (A)	خطای استاندارد	نقشه
۱	۰/۰۰۹	۰/۴۷۵	۰/۰۴۸	۰/۰۰۵	۰۱۱۱۱۰۱۰۱
۲	۰/۰۰۹	۰/۴۸۹	۰/۰۵۱	۰/۰۰۴	۱۱۱۰۱۱۱۱۰
۳	۰/۰۱۱	۰/۳۱۸	۰/۱۴۸	۰/۰۰۵	۰۱۱۰۰۱۱۰۰
۴	۰/۰۱۱	۰/۶۰۵	۰/۰۳۲	۰/۰۰۴	۱۱۱۱۰۱۱۰۱
۵	۰/۰۱۱	۰/۶۱۳	۰/۰۳۷	۰/۰۰۳	۱۱۱۱۱۱۰۰۱
۶	۰/۰۱۳	۰/۵۱۶	۰/۰۳۸	۰/۰۰۳	۱۱۱۱۰۱۱۱۰
۷	۰/۰۱۴	۰/۵۱۶	۰/۰۳۸	۰/۰۰۵	۱۱۱۱۱۰۱۱۰
۸	۰/۰۱۲	۰/۶۲	۰/۰۲۶	۰/۰۰۵	۱۱۱۱۱۱۱۱۰
۹	۰/۰۱۳	۰/۵۴	۰/۰۴۴	۰/۰۰۴	۱۱۱۱۱۱۰۱۰
۱۰	۰/۰۱۷	۰/۳۹	۰/۰۹۱	۰/۰۰۴	۱۱۱۱۰۱۰۰۰



شکل ۲- پراکنش مقادیر آماره گاما و دلتا حول خط رگرسیون

جدول ۲- تأثیر حذف متغیرهای ترکیب ورودی برتر بر مقدار آماره گاما

ترکیبات مختلف پارامترهای ورودی								پارامتر
حذف دما	حذف دما	حذف بارش	حذف دمای نقطه شب	حذف رطوبت نسبی	حذف بارش	ترکیب ۱	گاما	
با سه تأخیر	با دو تأخیر	بارش با سه تأخیر	آورد با یک تأخیر	با یک تأخیر	نسبی با یک تأخیر	با یک تأخیر		
۰/۰۱۵۳	۰/۰۱۶	۰/۰۱۴	۰/۰۱۰۴	۰/۰۲۱۷	۰/۰۱۷	۰/۰۲۳۲	۰/۰۰۹۱	گاما
۰/۰۱۸۷	۰/۰۰۸۶	۰/۰۳۰۹	۰/۰۴۷	-۰/۰۴۸	۰/۰۰۱۹	-۰/۰۱۷۸	۰/۰۴۸۸	گرادیان
۰/۰۰۳۳	۰/۰۰۳۸	۰/۰۰۲۹	۰/۰۰۳۸	۰/۰۰۳۴	۰/۰۰۱۹	۰/۰۰۲۹	۰/۰۰۵۲	خطای استاندارد
۰/۷۹۲۷	۰/۸۴۸۴	۰/۷۳۱۹	۰/۵۳۸	۰/۷۸۲	۰/۸۸	۰/۱۹۹	۰/۴۷۵۱	V _{Ratio}



شکل ۳- نتایج آزمون M-test

۲-۳- ارزیابی عملکرد مدل‌ها

بررسی چگونگی عملکرد و مقایسه قدرت مدل‌ها در پیش‌بینی جریان رودخانه زرينه‌رود در تحقیق حاضر لازم و ضروری است. برای این منظور جهت پیش‌بینی جریان، با استفاده از سه تابع هسته‌ای خطی، SVM توسعه مدل

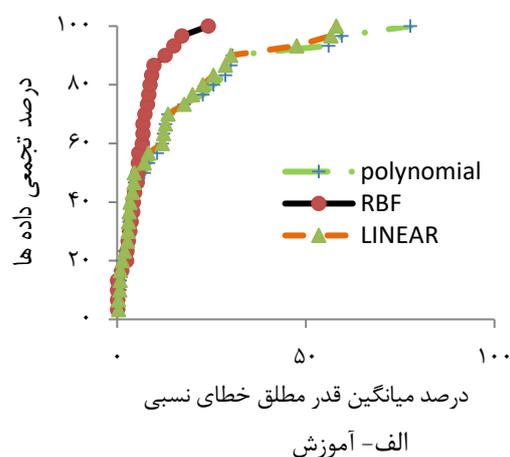
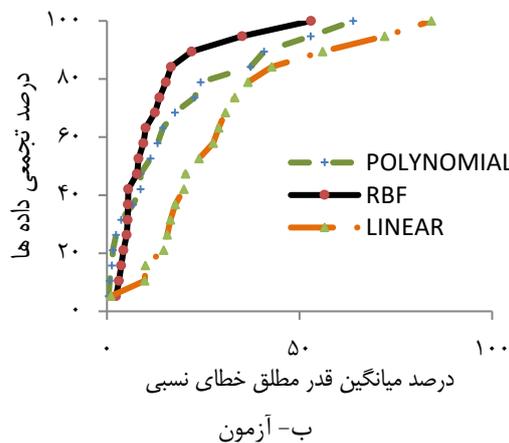
پارامترها به همراه نتایج ارزیابی مدل‌ها در جدول (۳) ارائه شده است. برای دو ترکیب برتر انتخاب شده، مدل RBF پل‌نومیال و نظر قرار گرفت. مقادیر پارامترهای مورد نیاز مدل و توابع هسته‌ای از طریق جستجوی شبکه‌ای تعیین شد که این

جدول ۳- نتایج ارزیابی مدل SVM

پارامتر	آزمون			آموزش			تابع هسته‌ای	ترکیب		
	C	ϵ	γ	MAE	RMSE	R ²			MAE	RMSE
۴	۰/۰۰۱	۲	۰/۰۳۷	۰/۰۴۸	۰/۸۸	۰/۸۸	۰/۰۴	۰/۹۵	RBF	۱
۱۶	۰/۰۰۱	۱	۰/۰۴۸	۰/۰۷	۰/۷۷	۰/۰۴	۰/۰۶	۰/۷۸	Polynomial	
۸	۰/۰۲	-	۰/۰۸۳	۰/۰۹	۰/۶	۰/۰۵	۰/۰۶۸	۰/۶۲	Linear	
۵۱۲	۰/۰۰۲	۵/۸	۰/۰۴۷	۰/۰۶	۰/۸	۰/۰۲	۰/۰۴	۰/۹۳	RBF	۲
۲۵۶	۰/۰۰۲	۹	۰/۰۴۷	۰/۰۶	۰/۸۴	۰/۰۳۷	۰/۰۵	۰/۹	Polynomial	
۲۵۶	۰/۱	-	۰/۰۶	۰/۰۷	۰/۷۸	۰/۰۴۸	۰/۰۷	۰/۷۶	Linear	

جدول (۳) ارائه شده‌اند، پراکنش زیادی در مورد نحوه توزیع خطا می‌دهند. به این منظور برای ارزیابی مدل‌ها از شاخص میانگین قدر مطلق خطای نسبی (MARE) و تحلیل آستانه خطا (TS) و نمودار پراکنندگی خطای مطلق نسبی نیز استفاده شد. در شکل (۴) نمودار توزیع خطای مربوط به مراحل آموزشی و آزمایشی سه روش مورد استفاده برای پیش‌بینی جریان آمده است.

با توجه به جدول (۳) تابع هسته‌ای RBF برای هر دو ترکیب مقادیر R² بالاتر و RMSE و MAE کمتری را نشان می‌دهد و این نشان از برتری این تابع دارد. رتبه‌بندی بهترین مدل‌ها بر اساس معیارهای ارزیابی فوق جهت پیش‌بینی جریان در رودخانه زربنده رود به ترتیب به صورت SVM-RBF، SVM، Polynomial و SVM-Linear می‌باشد. آماره‌های R² و MAE و RMSE که در



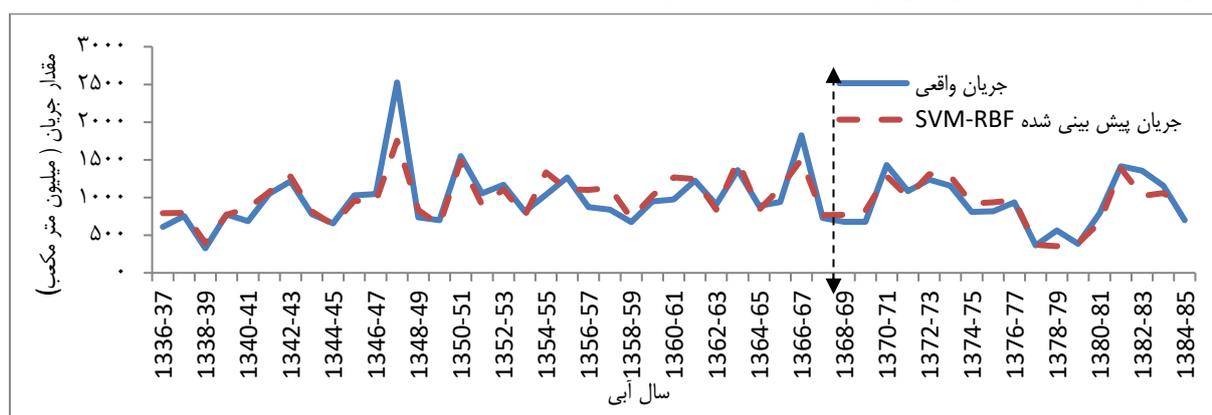
شکل ۴- نحوه توزیع و پراکنندگی میانگین قدر مطلق خطای نسبی در مرحله الف) آموزش و ب) آزمون رودخانه زربنده رود

۴۷ است و تنها برای ۵ درصد موارد است که قدر مطلق خطای نسبی به ترتیب تا ۲۴، ۷۷، ۵۸ می‌رسد. در مرحله آزمون نیز در ۹۵ درصد موارد خطای برآورد مدل‌های

شکل (۴) نشان می‌دهد که در مرحله آموزشی در ۹۵ درصد تجمعی داده‌ها خطای برآورد SVM-RBF، SVM، Polynomial، SVM-Linear به ترتیب کمتر از ۱۵، ۵۶،

برابر مقادیر واقعی جریان برای دو بخش آزمون و آموزش آورده شده است. در شکل (۵) مشاهده می‌شود که روند آنها از همبستگی بسیار خوبی برخوردار است. همچنین شکل (۵) نشان می‌دهد که در اکثر سال‌ها مقادیر پیش‌بینی شده و واقعی به هم نزدیک هستند و تنها در موارد معدودی مقادیر آنها با وجود اینکه روند یکسانی دارد، باهم تفاوت پیدا کرده‌اند. جهت بررسی دقت پیش‌بینی‌ها در این بخش و بررسی میزان کارایی آن‌ها در برنامه‌ریزی منابع آب حوضه، با بهره‌گیری از راهکارهای ارائه شده در طرح مدیریت ریسک خشکسالی دریاچه ارومیه (WGSMWRA 2012)، چهار سطح برای آورد رودخانه زرينه‌رود در نظر گرفته شد.

SVM-RBF، SVM-Polynomial، SVM-Linear به ترتیب کمتر از ۳۵، ۵۲، ۷۲ و در ۵ درصد موارد قدر مطلق خطای نسبی به ترتیب تا ۵۲، ۶۳، ۸۴ می‌رسد. بنابراین مدل SVM-RBF از هر دو مدل SVM-Polynomial، SVM-Linear دقت بیشتری داشته و میزان خطای کمتری تولید می‌کند که نتایج جدول (۳) نیز مؤید همین مطلب است. مقدار ضریب تعیین و RMSE برای مدل برتر انتخاب شده برای دوره آموزش به ترتیب برابر ۰/۹۵ و ۰/۰۳۹۲ و برای دوره آزمون به ترتیب برابر ۰/۸۸ و ۰/۰۴۷۷۴ می‌باشد که نشان از دقت مناسب مدل پیش‌بینی ماشین بردار پشتیبان در تخمین مقادیر جریان می‌باشد. این نتایج با مطالعه Kakaei-Lefdani et al. (2014) مطابقت دارد. در شکل (۵) روند تغییرات مقادیر جریان پیش‌بینی شده با استفاده از مدل برتر (SVM-RBF) در



شکل ۵- سری زمانی آورد واقعی و پیش‌بینی شده توسط مدل SVM-RBF

برای اعتبار سنجی و ارزیابی مدل GSVM در پیش‌بینی سطوح خشکسالی هیدرولوژیک از داده‌های تاریخی آورد سالانه از سال ۱۳۳۶ تا ۱۳۸۵ استفاده شد. بدین منظور از یک ماتریس خطا مطابق جدول (۵) استفاده شد که در آن دقت استفاده کننده (UA) و دقت تولید کننده (PA) برای کلاس‌های مختلف آورد درج شده است. به طوریکه λ نشان‌دهنده مقدار آورد به دست آمده از مدل پیش‌بینی SVM و λ مقادیر واقعی آن می‌باشد. مقدار پیش‌بینی در

این سطوح بر مبنای پتانسیل آورد رودخانه در سطوح مختلف خشکسالی و بر اساس سناریوهای مدیریتی پیشنهادی در طرح یاد شده تعیین شده است. نحوه تعیین طبقه قرارگیری آورد بر این اساس است که شرایط طبیعی در پتانسیل آورد به عنوان سطح صفر لحاظ می‌شود و شرایط مدیریتی که در آن زراعت به صورت کامل حذف گردد، به عنوان سطح ۴ معرفی می‌شود. بازه بین آنها به سه قسمت مساوی تقسیم و بر مبنای آن سه سطح دیگر تعیین می‌شود. جدول (۴) نتایج سطح بندی خشکسالی را برای پایین دست زرينه‌رود ارائه می‌دهد.

نشان‌دهنده قرارگیری سطح پیش‌بینی شده در سطح پایین‌تر نسبت به سطح واقعی می‌باشد، ۲۰/۴ درصد بدست آمد.

۴- نتیجه‌گیری

در این تحقیق به بررسی کارایی و دقت مدل‌های پیش‌بینی آورد پتانسیل رودخانه زرینه‌رود با استفاده از رویکرد ترکیبی بر مبنای آزمون گاما پرداخته و نتایج زیر حاصل شد.

۱- با بررسی طیف متنوعی از متغیرهای هواشناسی ورودی به مدل پیش‌بینی آورد و با تأخیرهای زمانی مختلف به‌وسیله آزمون گاما، متغیر بارش با یک تأخیر زمانی، مؤثرترین متغیر بر پیش‌بینی میزان آورد رودخانه زرینه‌رود شناسایی شد.

۲- مقایسه مقادیر مشاهداتی و پیش‌بینی شده آورد بر اساس رویکرد ترکیبی GSVM حاکی از عملکرد قابل قبول مدل پیش‌بینی می‌باشد. بر این اساس در دوره آموزش ضریب تعیین و RMSE پیش‌بینی به ترتیب برابر ۰/۹۵ و ۰/۰۴ و در دوره آزمون به ترتیب برابر ۰/۸۸ و ۰/۰۴۸ می‌باشد.

۳- نمودار توزیع خطای مربوط به مراحل آموزشی و آزمایشی مدل SVM با تابع هسته‌ای مختلف برای بهترین ترکیب انتخاب شده از آزمون گاما نشان داد که در هر دو بخش مدل SVM-RBF خطای کمتری را در پیش‌بینی ایجاد کرده است و توزیع خطای هموارتری دارد.

۴- دقت کلی مدل ترکیبی در پیش‌بینی سطوح مدیریتی خشکسالی بر مبنای آورد رودخانه زرینه‌رود نشان از دقت قابل قبول مدل GSVM در پیش‌بینی آورد در شرایط مختلف هیدرولوژیکی و همچنین قابلیت به‌کارگیری آن در مدیریت منابع آب حوضه دارد.

صورت $i=j$ "درست"، $i < j$ "برآورد دست پایین" و اگر $i > j$ "برآورد دست بالا" است.

جدول ۴- سطح بندی مقادیر جریان (WGSMWRA) (2012)

سطح خشکسالی	آورد متوسط رودخانه (میلیون متر مکعب)
۰	۱۰۷۰
۱	۹۴۰
۲	۸۱۰
۳	۶۸۰
۴	۵۶۰

جدول ۵- ماتریس خطای مدل GSVM در پیش‌بینی سطوح خشکسالی

سطح واقعی آورد (i)	سطح پیش‌بینی شده آورد (j)				
	۰	۱	۲	۳	۴
۰	۱۵	۱	۰	۰	۰
۱	۱	۵	۲	۰	۰
۲	۲	۱	۵	۰	۰
۳	۱	۰	۲	۴	۱
۴	۰	۰	۱	۲	۶
UA(%)	۷/۹۵	۷/۴۲	۵	۶/۷	۸/۷

دقت کلی (OA) پیش‌بینی آورد با روش GSVM برای رودخانه زرینه‌رود برابر با ۷۱/۴ درصد می‌باشد. این میزان دقت به این بدین معنی است که در ۷۱/۴ درصد موارد به‌درستی سطح آورد پیش‌بینی شده است. همچنین خطای برآورد دست بالا که نشان می‌دهد آورد پیش‌بینی شده در سطح بالاتری نسبت به سطح واقعی قرار دارد، برابر ۸/۲ درصد می‌باشد. خطای برآورد دست پایین نیز که

References

Asefa T., Kemblowski M. W., Mckee M. and Khalil A. (2006). Multi-time scale stream flow prediction: the support vector machine approach. J. Hydrol., 318, 7-16.

Behzad M., Asghari K., Eazi M. and Pallhang M. (2009). Generalization performance of support vector machines and neural networks in runoff modeling. Expert Sys. Appl., 36(4), 7624-7629.

² Upper estimate Error

¹ Under Estimate Error

- Chu H., Wei J., Li T. and Jia K. (2016). Application of support vector regression for mid- and long-term runoff forecasting in "Yellow River Headwater" region. HIC.2016. 12th Int. Conf. Hydroinformatics, Procedia Engineering 154,1251 – 1257.
- Dibike Y. B., Velickov S., Solomatine D. and Abbott M. B. (2001). Model induction with support vector machines: introduction and applications. *J. Comput. Civil Eng.*, 15(3), 208–216.
- Gholamzade M., Morid M. and Delava M. (2011). Use of early drought warning system for Zayandeh Rood Dam Operation. *J. Agri. Sci. Technol. Water Soil Sci.*, 56, 35-47 [In Persian].
- Guven A. (2009). Linear genetic programming for time-series modeling of daily flow rate. *J. Earth Syst. Sci.*, 118(2),157-173.
- Hamel L. (2009). Knowledge Discovery with Support Vector Machines. Hoboken NJ. John Wiley.
- Jones A. J. (2004). New tools in non-linear modeling and prediction. *Comput. Manag. Sci.*, 1(2), 109-149.
- Kakaei-Lefdani E., Moghadamnia A., Ahmadi A. and Ebrahimi H. (2014). Evaluation of the effect of pre-processing of input variables in the model of supporting vector machine by gamma proton method in order to precipitate suspended sediment volume. *J. Grassland Watershed*, 67(2), 289-303 [In Persian].
- Kelmes V. (1973). Watershed as semi infinitestorage reservoirs. *J. Irrig. Drain.*, 99, 477-491.
- Kiani Falavarjani M., Ahmadi A. and Gorji M. (2011). Long-term Prediction of Runoff into the Zayandehrood Dam Reservoir Using Large-scale Climatic Signals and intelligent Computing Methods. 6th Scientific Congress of Civil Engineering, Semnan University, Semnan [In Persian].
- Liong S. Y. and Sivapragasam C. (2002). Flood stage forecasting with support vector machines. *Am. Water Resour. Assoc.*, 38, 173-186.
- Moghaddamnia A., Ghafari Gousheh M., Piri J., Amin S. and Han D. (2009a). Evaporation estimation using artificial neural networks and adaptive neuro-fuzzy inference system techniques. *Adv. Water Resour.*, 32(1), 88-97.
- Moghaddamnia A., Remesan R., Hassanpour Kashani M., Mohammadi M., Han D. and Piri J. (2009b). Comparison of LLR, MLP, Elman, NNARX and ANFIS Models - with a case study in solar radiation estimation. *J. Atm. Solar-Terrest. Phys.*, 71, 975-982.
- Nayak P. C., Sudheer K. P., Rangan D. M. and Ramasastri K. S. (2004). A neuro fuzzy computing technique for modeling hydrological time series. *J. Hydrol.*, 291(2), 52-66.
- Noori R., Khakpour A., Dehghani M. and Farokhnia A. (2011). Monthly stream flow prediction using support vector machine based on principal component analysis. *J. Water Wastewater*, 22(1), 118-123 [In Persian].
- Remesan R., Shamim M.A. and Han D. (2008). Model data selection using gamma test for daily solar radiation estimation. *Hydrol. Proc.*, 22, 4301-4309.
- Riahi S., Pourbasheer E., Ganjali M.R. and Norouzi P. (2009). Investigation of different linear and nonlinear chemometric methods for modeling of retention index of essential oil components: concerns to support vector machine. *J. Hazard. Mater.*, 166(2), 853-859.
- Sharifi A.R., Dinpashoh Y., Fakheri-Fard. and Moghaddamnia A.R. (2013). Optimum combination of variables for runoff simulation in Amameh Watershed using gamma test. *Water Soil J.*, 23(4), 59-72 [In Persian].
- Silverman D. and Dracup J. A. (2000). Artificial neural network and long range precipitation rediction California. *J. Appl. Meteorol.*, 39(1), 57-66.

- Wan Jaafar W. Z., Liu J. and Han D. (2011). Input variable selection for median flood regionalization. *Water Resour. Res.*, 47, 1-18.
- Working Group on Sustainable Management of Water Resources and Agriculture (WGSWMRA), Regional Council of Lake Urmia Basin Management. (2012). Drought risk management plan for lake urmia basin, Tarbiat Modares university and Iranian conservation of wetlands plan, Volume 8 [In Persian].
- Lin J.Y., Cheng C.T. and Chau K.W. (2006). Using support vector machines for long-term discharge prediction. *Hydrol. Sci. J.*, 51(4), 599-612.
- Yu P.S., Chen S.T. and Chang I.F. (2006). Support vector regression for real-time flood stage forecasting. *J. Hydrol.*, 328, 704-716.

Application of Hybrid Gamma-SVM Approach for River Flow Prediction in Zarinehrud Basin

Abbas Abbasi^{1,2}, Majid Delavar^{3*} and Saeid Morid⁴

¹M.Sc., Department of Water Resources Engineering, Faculty of Agriculture, Tarbiat Modares University, Tehran, Iran

²Ph.D Scholar, Department of Water Resources Engineering, Faculty of Agriculture, Urmia University, Urmia, Iran

³Assist. Professor, Department of Water Resources Engineering, Faculty of Agriculture, Tarbiat Modares University of Tehran, Iran

⁴Professor, Department of Water Resources Engineering, Faculty of Agriculture, Tarbiat Modares University, Tehran, Iran

*Corresponding author: m.delavar@modares.ac.ir

Original Paper

Received: November 12, 2017

Accepted: December 19, 2017

Abstract

Prediction of river flow is an important issue in planning water resources and management of supply and demand in future conditions. Hence, it has attracted researchers' notice due to its importance in the designing, planning, management and operation of water facilities and also management the critical conditions such as flood and drought. In the present study, it was attempted to improve Zarinehrud river inflow prediction for use in water resource planning using a hybrid approach based on gamma test and supporting vector machine model (GSVM). For this purpose, the best possible combination of predictors was selected from the different combinations of 10 meteorological and hydrological variables in the basin. Then, based on the best combination of predictors, the potential of river inflow was predicted using a support vector machine. Comparison of predicted and observed flow indicated the good performance of hybrid approach in prediction of potential river inflow for application in basin management plans. In this case, the overall accuracy of the model to predict drought management levels based on Zarinehrud flow is 71.4%, and the upper and under estimation error are 8.2 and 20.4% respectively. These results show the acceptable precision of GSVM model for flow prediction in different hydrological situations of basin.

Keywords: Gamma Test; Flow Prediction; Zarinehrud; Support Vector Machine