

استفاده از تئوری آنتروپی و آزمون گاما در تعیین متغیرهای ورودی برای تخمین تبخیر روزانه (مطالعه موردی: ایستگاه‌های سینوپتیک رشت، بندرانزلی و آستارا)

سید مصطفی بی‌آزار^۱، محمدعلی قربانی^{۲*}، صابره دربندی^۳

۱. دانشجوی دکتری مهندسی منابع آب، دانشکده کشاورزی، دانشگاه تبریز

۲. دانشیار گروه مهندسی منابع آب، دانشکده کشاورزی، دانشگاه تبریز

۳. استادیار گروه مهندسی منابع آب، دانشکده کشاورزی، دانشگاه تبریز

(تاریخ دریافت ۱۳۹۶/۰۶/۰۱؛ تاریخ تصویب ۱۳۹۶/۱۱/۱۵)

چکیده

در این تحقیق قابلیت تئوری آنتروپی و آزمون گاما برای تعیین ورودی مدل‌های شبکه عصبی مصنوعی و ماشین بردار پشتیبان به منظور تخمین تبخیر ایستگاه‌های سینوپتیک رشت، آستارا و بندر انزلی در استان گیلان بررسی شده است. با توجه به نتایج پژوهش، برای ایستگاه‌های سینوپتیک رشت، آستارا و انزلی، تئوری آنتروپی وجود همه متغیرها را در مدل‌سازی مؤثر تشخیص داده است. آزمون گاما برای ایستگاه رشت دو متغیر رطوبت حداکثر و رطوبت متوسط، برای ایستگاه انزلی سه متغیر دمای حداقل، دمای متوسط و رطوبت متوسط و برای ایستگاه آستارا یک متغیر سرعت باد را از ترکیب بهینه خارج کرد. بنا بر نتایج، عملکرد هر دو مدل شبکه عصبی مصنوعی و ماشین بردار پشتیبان در هر دو بخش قابل قبول بوده است. در تعیین تک‌ورودی در ایستگاه رشت روش تئوری آنتروپی و در ایستگاه انزلی روش آزمون گاما بهتر عمل کرده‌اند. برای ایستگاه آستارا هر دو روش عملکرد مناسبی داشته‌اند. به طور کلی، با توجه به نتایج می‌توان گفت که تئوری آنتروپی نسبت به آزمون گاما عملکرد قوی‌تری داشته است.

کلیدواژگان: آزمون گاما، آنتروپی، استان گیلان، تبخیر روزانه، مدل تخمین.

مقدمه

تبخیر از مهم‌ترین فرایندها در هواشناسی و هیدرولوژی به شمار می‌رود [۱]. در سال‌های اخیر هرچند از شبکه عصبی مصنوعی و ماشین بردار پشتیبان به طور گسترده در تخمین تبخیر استفاده شده است، ولی پیش‌پردازش داده‌های ورودی برای بهبود تخمین با استفاده از آزمون گاما و تئوری آنتروپی مورد توجه قرار نگرفته است. روش‌های پیش‌پردازش داده‌ها با شناسایی ورودی‌های مؤثر کاهش ابعاد ورودی‌های مدل و از بین بردن روند نایستایی موجود در ورودی‌ها موجب افزایش کارایی مدل برای پیش‌بینی پدیده‌ها برای آینده می‌شود. آزمون گاما روشی برای یافتن ترکیبی مناسب از پارامترهای ورودی برای مدل‌سازی فرایندهای غیرخطی است. اساساً این آزمون به‌عنوان یک ابزار می‌تواند برای کمک به ساخت مدلی هموار از رفتار سیستمی آن پدیده براساس سری داده‌های اندازه‌گیری شده به کار رود. همچنین، از آنجا که ارتباط بین بسیاری از پدیده‌های هیدرولوژیکی با یکدیگر توأم با پیچیدگی بوده و غیرخطی است، بنابراین روش تئوری آنتروپی، که ارتباط غیرخطی بین پارامترها را می‌سنجد، برای مشخص کردن قدرت ارتباط غیرخطی بین آنها توصیه می‌شود [۲].

از مطالعات صورت گرفته در زمینه آزمون گاما و تئوری آنتروپی در تعیین پارامترهای ورودی، می‌توان به موارد زیر اشاره کرد: احمدی و همکارانش [۳] روش‌های متعدد گزینش پارامترهای ورودی مانند آزمون گاما، نظریه آنتروپی، معیارهای AIC و BIC را روی پارامترهای مؤثر بر تابش خورشیدی در حوضه برو در انگلستان بررسی کردند. ایشان برای مقایسه روش‌های یادشده از دو روش مدل‌سازی غیرخطی به نام‌های رگرسیون خطی محلی شبکه عصبی مصنوعی استفاده کردند. آنها ضمن مقایسه روش‌های مختلف نتیجه گرفتند که روش آزمون گاما در مقایسه با سایر روش‌های انتخاب بهترین ترکیب از ورودی‌ها، عملکرد بهتری دارد. افزون بر این، آزمون گاما تعداد داده‌های مورد نیاز برای مدل‌سازی را به‌طور درخور توجهی کاهش می‌دهد. مقدم‌نیا و همکارانش [۴] برای یافتن بهترین ترکیب پارامترهای ورودی برای مدل‌سازی تبخیر روزانه دریاچه چاه نیمه واقع در شمال شرقی استان سیستان و بلوچستان، از بین پارامترهای هواشناسی شامل میانگین روزانه دمای هوا،

کمبود فشار بخار اشباع، رطوبت نسبی و سرعت باد در ارتفاع دو متری در دوره آماری ۱۹۸۳-۲۰۰۵ از آزمون گاما استفاده کردند. ایشان برای مدل‌سازی تبخیر روزانه دریاچه از دو روش شبکه عصبی مصنوعی و سیستم استنتاجی عصبی- فازی تطبیقی استفاده کردند. نتایج پژوهش آنها نشان داد بهترین ترکیب شامل پارامترهای کمبود فشار بخار اشباع، رطوبت نسبی و سرعت باد در ارتفاع دو متری است. همچنین، روش شبکه عصبی مصنوعی برای پیش‌بینی تبخیر نسبت به مدل تجربی و سیستم استنتاج فازی-عصبی تطبیقی کارایی بیشتری دارد. نورانی و همکارانش [۲] در مدل‌سازی پدیده بارش-رواناب برای انتخاب زیرسری‌های مهم به‌منظور اعمال به مدل ترکیبی شبکه عصبی-موجکی از روش آنتروپی انتقال اطلاعات بهره گرفتند. نتایج پژوهش آنها نشان داد استفاده از روش آنتروپی انتقال اطلاعات نسبت به ابزارهای خطی دیگر سبب بهبود مدل‌سازی می‌شود. شریفی و همکارانش [۵] با استفاده از مدل‌های خطی و غیرخطی پیش‌بینی رواناب روزانه حوضه امامه را بررسی کردند. ایشان در مطالعه خود از آزمون گاما، انتخاب پیش‌رو و تحلیل عاملی برای انتخاب بهترین ترکیب ورودی استفاده کردند. نتایج پژوهش آنها نشان داد آزمون گاما عملکرد بهتری نسبت سه روش دیگر داشته است. از میان چهار روش شبکه عصبی مصنوعی، رگرسیون خطی محلی، سیستم استنتاج فازی-عصبی تطبیقی و ماشین بردار پشتیبان، مدل ماشین بردار پشتیبان از سه روش دیگر عملکرد بهتری در تخمین رواناب در این حوضه داشته است. از مطالعات در زمینه تخمین تبخیر نیز می‌توان به موارد زیر اشاره کرد:

سیفی و همکارانش [۶] تحقیقی در زمینه معرفی و کاربرد حداقل مربعات ماشین بردار پشتیبان در برآورد تبخیر- تعرق مرجع و تحلیل عدم قطعیت نتایج آن در محدوده شهر کرمان انجام دادند. در این تحقیق از روش حداقل مربعات ماشین بردار پشتیبان مبتنی بر آزمون گاما با شبکه عصبی تابع پایه شعاعی، خطی و چندجمله‌ای برای محاسبه تبخیر- تعرق لایسمتری استفاده شده است. نتایج با دو مدل شبکه عصبی مصنوعی (ANN)، سیستم استنتاجی عصبی- فازی تطبیقی و داده‌های لایسمتری مقایسه شد. نتایج حداقل مربعات ماشین بردار پشتیبان بیان‌کننده برتری شبکه عصبی تابع پایه شعاعی نسبت به سایر مدل‌ها بوده است. کیم و

است. روش بهینه‌سازی گرادیان مزدوج برای واسنجی معادلات ریاضی برای زیاد کردن دقت تخمین تبخیر استفاده شده است. نتایج به دست آمده از معادلات ریاضی با عملکرد مدل نروفازی و مدل درختی M5 مقایسه شد. نتایج پژوهش آنها نشان داد مدل‌های ریاضی که توسط روش گرادیان مزدوج بهینه شده بودند، عملکرد بهتری نسبت به مدل‌های درختی M5 و مدل نروفازی داشته‌اند.

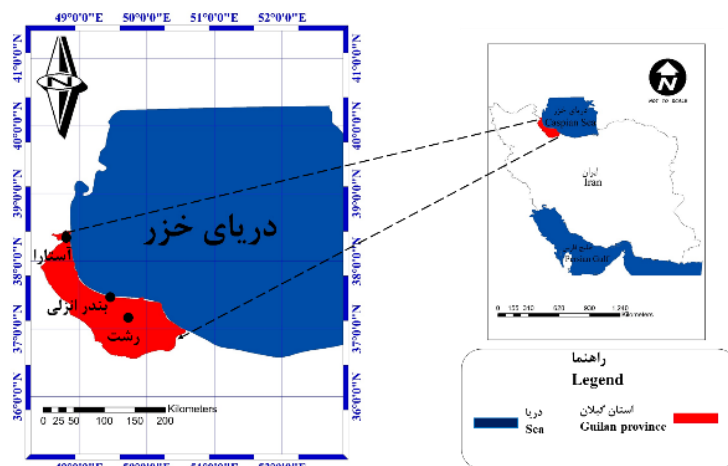
هدف از انجام این تحقیق، ارزیابی عملکرد روش‌های آزمون گاما و تئوری آنتروپی در انتخاب متغیرهای ورودی برای تخمین تبخیر با استفاده از مدل‌های ماشین بردار پشتیبان و شبکه عصبی مصنوعی در ایستگاه‌های سینوپتیک رشت، آستارا و بندر انزلی است که در خط ساحلی استان گیلان واقع شده‌اند.

مواد و روش‌ها

منطقه مطالعه شده و داده‌های استفاده شده

استان گیلان به مرکزیت رشت در شمال کشور قرار گرفته است و از شمال به دریای خزر، از شرق به استان مازندران، از جنوب به رشته کوه‌های البرز و استان‌های زنجان و قزوین و از غرب به استان اردبیل محدود است. توزیع ناهمگون نزولات جوی و حجم جریانات سطحی در طول سال و از طرفی تأثیر آب در کشاورزی منطقه به خصوص تولید محصول استراتژی برنج ضرورت بررسی و مطالعه سیستماتیک منابع آب اعم از سطحی و زیرزمینی را ایجاب کرده است. متوسط بارندگی استان گیلان ۱۱۰۰ میلی‌متر و میانگین حجم بارش ۱۵۰۲۷ میلیون مترمکعب است که از حجم یادشده میزان ۷۸۸۷ میلیون مترمکعب جز منابع غیرقابل دسترس (تبخیر و تعرق) و حجمی معادل ۷۱۴۰ میلیون مترمکعب قابل دسترس و تجدیدپذیر است. در این مطالعه از داده‌های سه ایستگاه سینوپتیک رشت، آستارا و انزلی، که واقع در مناطق مرکزی استان گیلان و نوار ساحلی دریای خزر هستند، استفاده شده است (شکل ۱). طول دوره آماری داده‌ها برای ایستگاه سینوپتیک رشت، آستارا و انزلی از سال ۱۳۸۳ تا سال ۱۳۹۳ است. داده‌های گرفته شده شامل تبخیر، سرعت باد در ارتفاع دو متری، دما (حداقل، حداکثر و متوسط)، رطوبت (حداقل، حداکثر و متوسط)، ساعت آفتابی و بارندگی بوده است.

همکارانش [۷] تبخیر روزانه را با استفاده از شبکه عصبی پرسپترون چندلایه و شبکه‌های عصبی همبستگی آبخاری و بر اساس دمای هوا، تابش خورشیدی، سرعت باد، رطوبت نسبی و فشار سطح دریا در چهار ایستگاه هواشناسی جمهوری کره مدل‌سازی کردند. نتایج پژوهش آنها نشان داد شبکه عصبی مصنوعی پرسپترون چندلایه عملکرد بهتری داشته است. گویال و همکارانش [۸] تبخیر روزانه حوضه کارسو در هندوستان را با استفاده از مدل‌های شبکه عصبی مصنوعی، حداقل مربعات رگرسیون بردار پشتیبان، منطق فازی و سیستم استنتاجی عصبی-فازی تطبیقی مدل‌سازی کردند. از آزمون گاما برای انتخاب پارامترهای مؤثر بارندگی، دمای حداقل و حداکثر، رطوبتی نسبی حداکثر و حداقل و ساعت‌های آفتابی استفاده شد. نتایج پژوهش آنها نشان داد مدل حداقل مربعات ماشین بردار پشتیبان و منطق فازی عملکرد بهتری نسبت به روش‌های دیگر داشته‌اند. تزال و بویکلدیز [۹] با استفاده از ماشین بردار پشتیبان و شبکه عصبی مصنوعی تبخیر ماهانه را در بهسیر تخمین زدند. در این تحقیق از شبکه‌های عصبی مصنوعی پرسپترون چندلایه، تابع پایه شعاعی^۱ و ماشین بردار پشتیبان استفاده شده است. داده‌های استفاده شده شامل رطوبت نسبی، سرعت باد و بارندگی بوده است. نتایج پژوهش آنها نشان داد شبکه عصبی مصنوعی پرسپترون چندلایه عملکرد بهتری نسبت به ماشین بردار پشتیبان داشته است. کیشی [۱۰] تبخیر از تشت را با استفاده از مدل‌های حداقل مربعات ماشین بردار پشتیبان، رگرسیون چندگانه انطباقی و مدل درختی M5 تخمین زد. نتایج پژوهش ایشان بیان‌کننده دقت زیاد روش حداقل مربعات ماشین بردار پشتیبان است. کیشی و همکارانش [۱۱] تبخیر روزانه از تشت در آنکارا و پولاتی در ترکیه را با چند مدل هوشمند مدل‌سازی کردند. نتایج پژوهش آنها نشان داد عملکرد مدل شبکه عصبی مصنوعی از سایر مدل‌ها بهتر بوده است. کشتکار و همکارانش [۱۲] تبخیر روزانه را با استفاده از روش‌های غیرخطی ریاضی براساس روش گرادیان مزدوج مدل‌سازی کردند. در این تحقیق با استفاده از تحلیل رگرسیون ورودی ترتیب متغیرهای ورودی به ترتیب دمای هوا، رطوبت نسبی، سرعت باد، و ساعت آفتابی برای دو ایستگاه مطالعاتی چاه نیمه زابل و ایرانشهر پیشنهاد شده



شکل ۱. منطقه مطالعه شده

آزمون گاما

آزمون گاما یک روش غیرخطی است که نتایج آن صرف نظر از روش های خاص برای مدل سازی به کار برده می شود [۱۳]. فرض می شود مجموعه داده های مشاهداتی به صورت رابطه ۱ است.

$$\{(x_i - y_i)\}, 1 \leq i \leq M \tag{1}$$

که در آن x_i بردار مشاهداتی ورودی (مانند رواناب، بارش، دمای هوا با تأخیرهای مختلف، تبخیر و هر متغیر مؤثر دیگر)، y_i خروجی مشاهداتی (مانند میانگین دبی رواناب روزانه) و M تعداد کل مشاهدات است.

از بردار x_i (در فضای m بعدی) برای مدل سازی y_i (در فضای R) استفاده می شود. بنابراین، با توجه به فرض یادشده رابطه بین ورودی و خروجی به صورت رابطه ۲ در نظر گرفته می شود.

$$y = f(x) + r \tag{2}$$

که در آن $f(x)$ یک تابع هموار و r متغیر تصادفی نشان دهنده خطا است. همچنین، فرض شده است میانگین توزیعی که r از آن پیروی کرده، صفر و واریانس خطا $Var(r)$ کران دار است. به این ترتیب، مدل نشان داده شده با رابطه ۲ به مدلی هموار با مشتقات جزئی مرتبه اول محدود می شود و آماره گاما (Γ)، واریانس بخشی از داده ها که تابع $f(x)$ قادر به توجیه آن نیست را نشان می دهد.

برای بردار ورودی x_i مجموعه $N[i, k]$ ($1 \leq k \leq p$) به نام مجموعه نزدیک ترین همسایگی است. پایه و اساس

آزمون گاما بر مبنای مجموعه $N[i, k]$ است. $x_{N[i, k]}$ به عنوان k امین نزدیک ترین همسایگی ها برای بردار مشاهداتی ورودی x_i تعریف شده به طوری که $(1 \leq k \leq p)$ و $(1 \leq i \leq M)$ است. حداکثر تعداد همسایگی ها است که معمولاً بین ۱۰ تا ۵۰ در نظر گرفته می شود [۱۴]. به منظور به دست آوردن مقدار آماره گاما (Γ) ابتدا مقدار $(\delta_M(k))$ ، با توجه به داده های ورودی به صورت رابطه ۳ محاسبه می شود.

$$\delta_M(k) = \frac{1}{M} \sum_{i=1}^M |x_{N[i, k]} - x_i|^2, 1 \leq k \leq p \tag{3}$$

که در آن $|....|$ به عنوان فاصله اقلیدسی بردار $x_{N[i, k]}$ از همسایگی های آن است. همچنین، مقدار آماره گامای کوچک $\gamma_M(k)$ با استفاده از داده های خروجی به صورت رابطه ۴ به دست می آید:

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

که در آن $y_{N[i, k]}$ مقدار خروجی نظیر k امین نزدیک ترین همسایگی برای بردار x_i است. به این ترتیب، p مقدار برای $(\delta_M(k))$ و p مقدار برای $(\gamma_M(k))$ به دست آورده می شود. سپس، بین p زوج مرتب $\{\delta_M(k), \gamma_M(k)\}$ رابطه رگرسیونی خطی ساده به صورت رابطه ۵ برقرار خواهد شد، طوری که در آن δ و γ به ترتیب به عنوان متغیرهای مستقل و وابسته در نظر گرفته شده است.

$$\gamma = A \delta + \Gamma \tag{5}$$

شانون و ویور [۱۶] در سال ۱۹۴۹ از یک متغیر تصادفی گسسته x اولین آنتروپی مرزی را به نام $H(x)$ مشخص کردند.

$$H(x) = -K \sum_{i=1}^n P(x_i) \log P(x_i) \quad (۷)$$

مقدار پارامتر $K=1$ خواهد بود، به شرطی که مقدار لگاریتم $H(x)$ بر پایه عدد نپر (e) بیان شود. n بیان‌کننده تعداد وقایع با احتمال $P(x_i) (i=1, \dots, N)$ است.

آنتروپی مجموع دو متغیر تصادفی مستقل x و y از مجموع میزان آنتروپی مرزی دو متغیر به دست می‌آید.

$$H(x, y) = H(x) + H(y) \quad (۸)$$

وقتی که دو متغیر تصادفی x و y وابسته باشند، مقدار آنتروپی مشترک کمتر از مقدار آنتروپی مجموع رابطه ۹ می‌شود.

آنتروپی انتقال اطلاعات نوعی دیگر از آنتروپی است که ارتباط متقابل بین x و y را مشخص می‌کند. که با تفاضل بین آنتروپی مجموع و آنتروپی مشترک از دو متغیر وابسته x و y به دست می‌آید.

$$T(x, y) = H(x) + H(y) - H(x, y) \quad (۹)$$

$$T(x, y) = H(x) - H(x|y) = H(y) - H(y|x) \quad (۱۰)$$

که در آن $H(x|y)$ بیان‌کننده آنتروپی شرطی است. آنتروپی شرطی x به شرط y بیان‌کننده عدم قطعیت باقی‌مانده در x به شرط معلوم بودن y می‌باشد و به عکس.

با توجه به توضیحات یادشده می‌توان محاسبات آنتروپی را با M متغیر تعمیم داد [۱۷] آنتروپی مجموع برای متغیر مستقل $X_m (m=1, \dots, M)$ برابر است با:

$$H(x_1, x_2, \dots, x_M) = \sum_{m=1}^M H(x_m) \quad (۱۱)$$

اگر متغیرها وابسته باشند، آنتروپی مشترک آنها از رابطه ۱۲ محاسبه می‌شود.

$$H(x_1, x_2, \dots, x_M) = H(x_1) + \sum_{m=2}^M H(x_m | x_1, \dots, x_{m-1}) \quad (۱۲)$$

مقدار آماره Γ در واقع عرض از مبدأ خط رگرسیون فوق است. A شیب خط رگرسیون است که نشان‌دهنده پیچیدگی مدلی است و از روی داده‌های مشاهداتی ایجاد می‌شود. متغیر دیگری که معمولاً در آزمون گاما به کار برده می‌شود، متغیر بدون بعد V_{ratio} نام دارد که مقدار آن بین صفر و یک است و مقدار آن از رابطه ۶ محاسبه می‌شود [۱۵].

$$V_{ratio} = \frac{\Gamma}{\sigma^2(y)} \quad (۶)$$

که در آن $\sigma^2(y)$ واریانس داده‌های مشاهداتی است. هرچه مقدار V_{ratio} به صفر نزدیک‌تر باشد، ترکیب انتخابی برای مدل‌سازی مناسب‌تر است. ثابت شده است که اگر این مقدار از یک کم شود، مقدار ضریب تبیین مدل به دست می‌آید.

تئوری آنتروپی

تعریف آنتروپی مشتق شده از تئوری اطلاعات است، که یک سری از ساختارهای آماری از سری‌های عددی، علامت‌ها یا نمادها که بین سیگنال‌ها ارتباط برقرار می‌کند را تحلیل می‌کند. پایه اصلی تئوری اطلاعات توسط شانون و ویور در سال ۱۹۴۸ مطرح شد. طبق نظریه شانون، اطلاعات، به دست آمده از عدم قطعیت‌های موجود در هر واقعه است. شانون در سال ۱۹۴۸ نشان داد وقایع با احتمال وقوع زیاد اطلاعات کمتری در اختیار می‌گذارند و به عکس هر چقدر احتمال وقوع یک رخداد کمتر باشد، اطلاعات حاصل از آن بیشتر است. با به دست آوردن اطلاعات جدید، در واقع عدم قطعیت کاهش داده شده است و ارزش اطلاعات جدید برابر با مقداری است که از عدم قطعیت کاسته شده است. در نتیجه، عدم قطعیت و اطلاعات پارامترهای وابسته به هم هستند. به بیان دیگر، از تئوری آنتروپی می‌توان به عنوان شاخصی برای کمی کردن میزان ناآگاهی و دانش نسبت به مشخصات یک سامانه، استفاده کرد.

در کل، برای کمی کردن محتوای اطلاعات چهار نوع آنتروپی معرفی شده است. آنتروپی مرزی، آنتروپی مشترک، آنتروپی شرطی و آنتروپی انتقال اطلاعات. در این تحقیق از آنتروپی انتقال اطلاعات استفاده شده است که به شرح مختصری از آن می‌پردازیم.

در معادله ۱۵، میانگین ضرایب لاگرانژ است. محاسبه $\phi(x)$ در فضای مشخصه آن ممکن است بسیار پیچیده باشد. برای حل این مشکل روند معمول در مدل ماشین بردار پشتیبان انتخاب یک تابع کرنل است. با توجه به اینکه در ماشین بردار پشتیبان معمولاً از تابع کرنل پایه شعاعی استفاده می‌شود، در تحقیق حاضر نیز از این تابع استفاده شده است. فرمول کرنل پایه شعاعی به صورت رابطه ۱۶ آورده شده است [۲۱].

$$k(x, x_i) = \exp\left(-\|x - x_i\|^2 / \sigma^2\right) \quad (16)$$

درخور یادآوری است که فرایند محاسبات ماشین بردار پشتیبان، بر اساس کدنویسی در محیط متلب انجام و پارامترهای تابع کرنل از طریق سعی و خطا بهینه شدند.

معیارهای ارزیابی

با استفاده از روش‌های مختلفی می‌توان عملکرد مدل‌ها را ارزیابی و مقایسه کرد. یکی از این روش‌ها استفاده از معیارهای ارزیابی است. از جمله معیارهای ارزیابی پرکاربرد در علوم مهندسی آب، ضرایب همبستگی (CC)، ریشه میانگین مربعات خطا (RMSE) و معیار نش-ساتکلیف (NS) هستند که در روابط ۱۷-۱۹ آورده شده‌اند. دقیق‌ترین مدل با توجه به این معیارها، مدلی خواهد بود که مقدار این سه معیار برای آن به ترتیب، نزدیک به یک، صفر و یک باشد.

$$CC = \sqrt{\frac{(\sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y}))^2}{\sum_{i=1}^N (x_i - \bar{x})^2 \sum_{i=1}^N (y_i - \bar{y})^2}} \quad (17)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - y_i)^2} \quad (18)$$

$$NS = 1 - \frac{\sum_{i=1}^N (x_i - y_i)^2}{\sum_{i=1}^N (x_i - \bar{x})^2} \quad (19)$$

در روابط یادشده \bar{x} و \bar{y} به ترتیب، مقادیر مشاهداتی جریان و میانگین مقادیر مشاهداتی، y_i مقادیر محاسباتی جریان، \bar{y} میانگین مقادیر محاسباتی و N تعداد داده‌هاست. مقدار شاخص NS از یک تا منفی بی‌نهایت تغییر می‌کند. بازه‌های ۰/۷۵-۱، ۰/۳۶-۰/۷۵ و کمتر از ۰/۳۶ برای مقدار این شاخص در یک شبیه‌سازی به ترتیب نشان‌دهنده عملکرد بسیار خوب، رضایت‌بخش و ضعیف مدل استفاده‌شده هستند [۲۲].

شبکه عصبی مصنوعی

از اوایل دهه ۱۹۹۰ تا کنون، این شبکه‌های عصبی مصنوعی در زمینه هیدرولوژی نیز به طور موفقیت‌آمیزی به کار گرفته شده‌اند [۱۸]. حدود ۹۰ درصد شبکه‌هایی که در مسائل هیدرولوژی کاربرد دارد، شبکه‌های پیشرو هستند [۱۹]. معماری معمول این شبکه متشکل از سه لایه ورودی، پنهانی و خروجی است. تعداد نورون‌های موجود در لایه ورودی و خروجی بستگی به نوع مسئله دارد، ولی تعداد گره‌های لایه پنهان با سعی و خطا به دست می‌آید [۲۰].

درخور یادآوری است که در تحقیق حاضر از شبکه عصبی پیشرو با آموزش لوانبرگ-مارکوات و تعداد نورون‌های یک تا ۲۰ برای لایه پنهان شبکه استفاده شد تا از این طریق تأثیر ساختار مختلف شبکه بر عملکرد آن در تخمین تبخیر را ارزیابی شود. از آنجا که متداول‌ترین نوع توابع محرک تابع تانژانت سیگموئید است، در مطالعه حاضر از آن برای نگاشت اطلاعات از لایه ورودی به لایه پنهان و همچنین برای نگاشت اطلاعات از لایه پنهان به لایه خروجی استفاده شد.

ماشین بردار پشتیبان

ماشین بردار پشتیبان یک سیستم یادگیری کارآمد بر مبنای تئوری بهینه‌سازی مقید است که از اصل استقرای کمینه‌سازی خطای ساختاری استفاده می‌کند و به یک جواب بهینه کلی منجر می‌شود. در مدل ماشین بردار پشتیبان تابعی مرتبط با متغیر وابسته y که خود تابعی از چند متغیر مستقل x است، برآورد می‌شود. مشابه سایر مسائل رگرسیونی فرض می‌شود رابطه میان متغیرهای مستقل و وابسته با تابع جبری مانند $f(x)$ به علاوه مقداری اغتشاش (خطای مجاز ϵ) مشخص شود.

$$f(x) = W^T \cdot \phi(x) + b \quad (13)$$

$$y = f(x) + noise \quad (14)$$

چنانچه W بردار ضرایب، b ثابت مشخصه‌های تابع رگرسیونی و $\phi(x)$ نیز تابع کرنل باشد، آن‌گاه هدف پیدا کردن فرم تابعی برای $f(x)$ است. تابع رگرسیون بردار پشتیبان را می‌توان به صورت رابطه ۱۵ بازنویسی کرد:

$$f(x) = \sum_{i=1}^N \bar{\alpha}_i \cdot \phi(x_i)^T \cdot \phi(x) + b \quad (15)$$

نتایج

آزمون گاما

برای تعیین پارامترهای ورودی بهینه به منظور تخمین تبخیر روزانه در این بخش از آزمون گاما استفاده شد. ابتدا، آزمون گاما برای کل داده‌ها انجام گرفت (ردیف ۱) و سپس، این آزمون براساس حذف هر یک از متغیرها از کل نه متغیر دیگر صورت گرفته است (ردیف‌های ۲ تا ۱۰). نتایج در جدول ۲ ارائه شد. با توجه به جدول ۲ برای تخمین تبخیر در ایستگاه سینوپتیک رشت دو پارامتر رطوبت حداکثر و رطوبت متوسط دارای مقدار آماره گامای کمتر از ترکیب اول هستند. بنابراین، به غیر از این دو پارامتر از بقیه پارامترها می‌توان برای تخمین تبخیر در این ایستگاه استفاده کرد. برای ایستگاه آستارا به‌جز پارامتر سرعت باد از باقی پارامترها که با حذف از ترکیب، مقدار آماره گامای آنها بیشتر از ترکیب اول شده را می‌توان برای برآورد تبخیر در این ایستگاه استفاده کرد. در ایستگاه سینوپتیک انزلی با حذف پارامترهای دمای حداکثر، رطوبت حداکثر، رطوبت حداقل، بارندگی، سرعت باد و ساعت‌های آفتابی مقدار پارامتر گاما نسبت به حالت اول بیشتر شده است. در این ایستگاه با حذف پارامترهای دمای حداقل، دمای متوسط و رطوبت متوسط مقدار آماره گاما بیشتر از مقدار گامای ترکیب اول نشده است. پس برای تخمین تبخیر در این ایستگاه می‌توان از پارامترهای ساعت‌های آفتابی، سرعت باد، بارندگی، رطوبت حداقل، رطوبت حداکثر و دمای حداکثر استفاده کرد.

با توجه به جدول ۲ مشاهده می‌شود که در هر سه ایستگاه ترکیبی که در آن پارامتر ساعت‌های آفتابی حذف شده است، بدترین ترکیب است. بهترین ترکیب برای ایستگاه انزلی، آستارا و رشت به ترتیب ترکیباتی هستند که از آنها رطوبت متوسط، سرعت باد و رطوبت حداکثر حذف شده‌اند.

تئوری آنتروپی

با توجه به توضیحات بخش آنتروپی، در جدول‌های ۲-۴ نتایج مقادیر آنتروپی انتقال اطلاعات برای متغیرهای ورودی مختلف در تخمین تبخیر در سه ایستگاه رشت، آستارا و انزلی آورده شده است. در این ایستگاه‌ها داده‌های ورودی شامل دمای حداکثر، حداقل و متوسط، رطوبت حداکثر، حداقل و متوسط، بارندگی، سرعت باد و ساعت‌های آفتابی است. در جدول‌های یادشده ابتدا مقدار آنتروپی برای کل داده‌ها محاسبه شد و سپس متغیرها تک‌به‌تک حذف شده و مقادیر آنتروپی هر ترکیب دوباره محاسبه شد. بیشترین مقدار برای آنتروپی انتقال اطلاعات در هر سه ایستگاه هنگامی مشاهده می‌شود که همه ورودی‌های در دسترس در ترکیب وجود داشته باشند (ترکیب ۱).

برای ایستگاه رشت (جدول ۲) ترکیب ۸ با حذف متغیر بارندگی از مدل با مقدار آنتروپی ۵/۸۲۳ به‌عنوان بهترین ترکیب بعد از ترکیب نخست شناخته شده است. ترکیب ۴ نیز با حذف متغیر دمای متوسط به‌عنوان بدترین ترکیب با مقدار آنتروپی ۴/۲۲۸ بوده است.

جدول ۱. نتایج آزمون گاما برای ایستگاه‌های سینوپتیک مطالعه‌شده

متغیر حذف شده	شماره ترکیب	رشت		آستارا		بندر انزلی	
		گاما	نسبت ۷	گاما	نسبت ۷	گاما	نسبت ۷
-	۱	۰/۰۵۶۳۶	۰/۲۲۵۴	۰/۰۳۹۷۷۱	۰/۱۵۹۰۸	۰/۰۴۰۰۱	۰/۱۶۰۰۵
دمای حداکثر	۲	۰/۰۶۰۲۲	۰/۲۴۰۹	۰/۰۴۰۲۱۳	۰/۱۶۰۸۵	۰/۰۴۰۱۴	۰/۱۶۰۵۸
دمای حداقل	۳	۰/۰۵۸۵۸	۰/۲۳۴۳	۰/۰۴۱۵۵۸	۰/۱۶۶۲۳	۰/۰۳۸۸۳	۰/۱۵۵۳۰
دمای متوسط	۴	۰/۰۶۰۳۴	۰/۲۴۱۴	۰/۰۴۱۱۲۹	۰/۱۶۴۵۲	۰/۰۳۹۵۶	۰/۱۵۸۲۳
رطوبت حداکثر	۵	۰/۰۵۴۳۴	۰/۲۱۷۴	۰/۰۴۳۲۰۸	۰/۱۷۲۸۳	۰/۰۴۰۰۶	۰/۱۶۰۲۴
رطوبت حداقل	۶	۰/۰۵۸۰۱	۰/۲۳۲۰	۰/۰۴۱۴۷۲	۰/۱۶۵۸۹	۰/۰۴۰۵۲	۰/۱۶۲۰۹
رطوبت متوسط	۷	۰/۰۵۵۷۲	۰/۲۲۲۹	۰/۰۴۰۹۷۵	۰/۱۶۳۹۰۰	۰/۰۳۸۴۱	۰/۱۵۳۶۵
بارندگی	۸	۰/۰۵۸۰۹	۰/۲۳۲۴	۰/۰۴۱۷۱	۰/۱۶۶۸۴	۰/۰۴۱۶۶	۰/۱۶۶۶۴
سرعت باد	۹	۰/۰۵۹۱۹	۰/۲۳۶۸	۰/۰۳۹۵۶	۰/۱۵۸۲۴	۰/۰۴۳۳۴	۰/۱۷۳۳۶
ساعت‌های آفتابی	۱۰	۰/۰۶۵۸۵	۰/۲۳۶۴	۰/۰۴۵۱۵	۰/۱۸۰۶۰	۰/۰۴۶۹۶	۰/۱۸۷۸۳

جدول ۲. نتایج آزمون تئوری آنژیروپی برای ایستگاه‌های سینوپتیک رشت

ایستگاه	شماره ترکیب	متغیر حذف شده	آنژیروپی مرزی	آنژیروپی مشترک	آنژیروپی شرطی	آنژیروپی انتقال اطلاعات
	۱	-	۱۴/۴۷۲	۸/۵۲۱	۵/۹۹۸	۵/۹۵۱
	۲	دمای حداکثر	۱۲/۳۵۳	۸/۰۶۲	۵/۵۳۹	۴/۲۹۲
	۳	دمای حداقل	۱۲/۵۲۸	۸/۰۷۹	۵/۵۵۶	۴/۴۴۹
	۴	دمای متوسط	۱۲/۴۰۴	۸/۱۴۶	۵/۶۲۳	۴/۲۲۸
رشت	۵	رطوبت حداکثر	۱۴/۲۹۵	۸/۴۸۹	۵/۹۶۶	۵/۸۰۶
	۶	رطوبت حداقل	۱۲/۵۵۳	۷/۸۳۸	۵/۳۱۵	۴/۷۱۵
	۷	رطوبت متوسط	۱۳/۱۳۱	۸/۱۷۸	۵/۶۵۵	۴/۹۵۳
	۸	بارندگی	۱۴/۲۳۵	۸/۴۱۲	۵/۸۸۹	۵/۸۲۳
	۹	سرعت باد	۱۳/۳۹۶	۷/۸۲۳	۵/۳۰۰	۵/۵۷۳
	۱۰	ساعت‌های آفتابی	۱۲/۳۵۶	۷/۴۵۰	۴/۹۲۸	۴/۹۰۶

در ایستگاه آستارا (جدول ۳) ترکیب ۸ با داشتن مقدار آنژیروپی ۶/۶۲۴ به عنوان بهترین ترکیب بعد از ترکیب نخست با حذف متغیر بارندگی از ورودی مدل است. ترکیب ۶ به عنوان بدترین ترکیب با حذف متغیر رطوبت حداقل بوده است که مقدار آنژیروپی آن ۴/۴۸۰ است.

با توجه به جدول ۴ براساس حذف هر یک از متغیرها بهترین ترکیب برای ایستگاه انزلی بعد از ترکیب ۱، ترکیب ۸ است که با حذف متغیر بارندگی به دست آمده است. بدترین ترکیب در این ایستگاه، ترکیب ۴ که با حذف دمای متوسط به دست آمده است. همان طور که گفته شد، متغیر بارندگی در این ایستگاه کم‌اثرترین متغیر است که با

حذف آن از ترکیب مقدار آنژیروپی انتقال اطلاعات در ترکیب به صورت درخور توجهی تغییر نمی‌کند. همان طور که مشاهده می‌شود، با حذف متغیر بارندگی از ترکیب در هر سه ایستگاه مقدار آنژیروپی انتقال اطلاعات در ترکیب نخست تغییر زیادی ندارد. بنابراین، می‌توان چنین گفت که بعد از ترکیب ۱، از ترکیب ۸ نیز در تخمین تبخیر در هر سه ایستگاه می‌توان استفاده کرد. بنابراین، در جدول ۵ سه ترکیب قرار گرفته است که به ترتیب برای هر ایستگاه شامل آنژیروپی (EN1)، ترکیبی که همه متغیرها به عنوان ورودی انتخاب شده‌اند، آنژیروپی (EN2)، ترکیبی که متغیر ورودی بارندگی از ورودی‌ها حذف شده است و ترکیب انتخابی آزمون گاما (GT) است.

جدول ۳. نتایج آزمون تئوری آنژیروپی برای ایستگاه‌های سینوپتیک آستارا

ایستگاه	شماره ترکیب	متغیر حذف شده	آنژیروپی مرزی	آنژیروپی مشترک	آنژیروپی شرطی	آنژیروپی انتقال اطلاعات
	۱	-	۱۵/۸۲۶	۹/۰۶۶	۶/۴۵۳	۶/۷۶۰
	۲	دمای حداکثر	۱۳/۶۸۲	۸/۷۶۷	۶/۱۵۴	۴/۹۱۶
	۳	دمای حداقل	۱۳/۷۵۵	۸/۶۰۹	۵/۹۹۷	۵/۱۴۵
	۴	دمای متوسط	۱۳/۶۸۴	۸/۸۵۵	۶/۲۴۳	۴/۸۲۹
	۵	رطوبت حداکثر	۱۴/۸۰۸	۸/۶۸۸	۶/۰۷۵	۶/۱۲۱
بندر آستارا	۶	رطوبت حداقل	۱۳/۹۱۶	۸/۴۳۶	۵/۸۲۳	۴/۴۸۰
	۷	رطوبت متوسط	۱۴/۱۵۲	۸/۶۷۸	۶/۰۶۶	۵/۴۷۳
	۸	بارندگی	۱۵/۵۵۸	۸/۹۳۴	۶/۳۲۲	۶/۶۲۴
	۹	سرعت باد	۱۵/۱۱۵	۸/۷۰۸	۶/۰۹۵	۶/۴۰۷
	۱۰	ساعت‌های آفتابی	۱۳/۶۳۸	۷/۹۲۶	۵/۳۱۳	۵/۷۱۲

جدول ۴. نتایج آزمون تئوری آنتروپی برای ایستگاه‌های سینوپتیک انزلی

ایستگاه سینوپتیک	شماره ترکیب	متغیر حذف شده	آنتروپی مرزی	آنتروپی مشترک	آنتروپی شرطی	آنتروپی انتقال اطلاعات
	۱	-	۱۵/۷۴۸	۹/۲۲۶	۶/۴۴۶	۶/۵۲۲
	۲	دمای حداکثر	۱۳/۶۱۸	۸/۸۹۲	۶/۱۱۳	۴/۷۲۶
	۳	دمای حداقل	۱۳/۷۰۷	۸/۸۵۰	۶/۰۷۱	۴/۸۵۶
	۴	دمای متوسط	۱۳/۶۱۸	۹/۰۱۱	۶/۲۳۲	۴/۶۰۶
بندر انزلی	۵	رطوبت حداکثر	۱۵/۱۳۸	۹/۰۳۷	۶/۲۵۷	۶/۱۰۱
	۶	رطوبت حداقل	۱۴/۱۰۹	۸/۶۰۶	۵/۸۲۶	۵/۵۰۳
	۷	رطوبت متوسط	۱۴/۴۳۶	۸/۸۴۷	۶/۰۶۸	۵/۵۸۹
	۸	بارندگی	۱۵/۳۳۷	۹/۰۵۳	۶/۲۷۴	۶/۲۸۳
	۹	سرعت باد	۱۴/۲۷۶	۸/۲۶۸	۵/۴۸۹	۶/۰۰۷
	۱۰	ساعات‌های آفتابی	۱۳/۵۸۲	۸/۰۱۶	۵/۲۳۶	۵/۵۶۶

جدول ۵. ترکیب‌های مشخص شده توسط تئوری آنتروپی و آزمون گاما برای هر ایستگاه سینوپتیک مطالعه شده

ایستگاه	شماره ترکیب	روش	ترکیب
رشت	۱	آنتروپی (EN1)	همه ورودی‌ها
	۲	آنتروپی (EN2)	همه ورودی‌ها به جز متغیر بارندگی
	۳	گاما (GT)	دمای حداکثر، دمای متوسط، دمای حداقل، رطوبت حداقل، بارندگی، سرعت باد و ساعات‌های آفتابی
آستارا	۱	آنتروپی (EN1)	همه ورودی‌ها
	۲	آنتروپی (EN2)	همه ورودی‌ها به جز متغیر بارندگی
	۳	گاما (GT)	دمای حداکثر، دمای حداقل، دمای متوسط، رطوبت حداکثر، رطوبت حداقل، رطوبت متوسط، بارندگی و ساعات‌های آفتابی
انزلی	۱	آنتروپی (EN1)	همه ورودی‌ها
	۲	آنتروپی (EN2)	همه ورودی‌ها به جز متغیر بارندگی
	۳	گاما (GT)	دمای حداکثر، رطوبت حداکثر، رطوبت حداقل، بارندگی، سرعت باد و ساعات‌های آفتابی

تخمین میزان تبخیر با مدل‌های ANN و SVM

مدل شبکه عصبی مصنوعی پرسپترون چندلایه (MLP)

با توجه به جدول ۶ در تخمین تبخیر ایستگاه رشت، ترکیب‌های ۳ و ۱ عملکرد بهتری نسبت به ترکیب ۲ در مدل شبکه عصبی مصنوعی داشته‌اند. با توجه به معیارهای ارزیابی ترکیب ۳ به عنوان ترکیب برتر انتخاب شد که دارای ضریب همبستگی ۰/۸۸۸، مجذور میانگین مربعات خطای ۰/۸۸۵ و ضریب نش-ساتکلیف ۰/۷۸۵ در دوره صحت‌سنجی بوده است. همچنین، درخور یادآوری است که ترکیب ۱ نیز نتیجه نزدیکی به ترکیب ۳ دارد.

مدل شبکه عصبی مصنوعی به ازای ترکیب‌های ۱ و ۳ براساس مجذور میانگین مربعات خطا و ضریب نش-ساتکلیف، پارامتر تبخیر را در ایستگاه آستارا، با دقت قابل قبولی تخمین زده است. با توجه به آماره‌های ارزیابی، ترکیب

۱ با ضریب همبستگی، مجذور میانگین مربعات خطا و ضریب نش-ساتکلیف به ترتیب برابر ۰/۹۲۱، ۱/۱۶۴ و ۰/۷۹۰ در دوره صحت‌سنجی مقادیر پارامتر تبخیر را بهتر از سایر ترکیب‌های استفاده شده تخمین زده است.

تخمین تبخیر در ایستگاه انزلی به ازای ترکیب‌های ۱ و ۲ نتایج بهتری نسبت به ترکیب ۳ داشته است. البته، ترکیب ۱ با مجذور میانگین مربعات خطای کمتر نسبت به دو ترکیب دیگر به عنوان ترکیب برتر در تخمین تبخیر در ایستگاه انزلی انتخاب شد. این ترکیب در دوره صحت‌سنجی ضریب همبستگی برابر با ۰/۸۹۷، مجذور میانگین مربعات خطای برابر ۱/۰۶۲ و ضریب نش-ساتکلیف برابر ۰/۷۹۸ دارد.

مدل ماشین بردار پشتیبان (SVM)

با توجه به جدول ۶ مدل ماشین بردار پشتیبان تخمین

ترکیب برتر انتخاب شد. بنابراین، ترکیب ۳ مقادیر آماره‌های ضریب همبستگی، مجذور میانگین مربعات خطا و ضریب نش- ساتکلیف را در دوره صحت‌سنجی به ترتیب برابر با ۰/۹۱۹، ۱/۲۲۲ و ۰/۷۶۸ به دست آورده است. البته، عملکرد ترکیب ۱ بسیار نزدیک به ترکیب ۳ بوده است. در ایستگاه سینوپتیک انزلی مدل ماشین بردار پشتیبان با استفاده از ترکیب ۱ پارامتر تبخیر را با دقت مناسب تری نسبت به دو ترکیب دیگر تخمین زده است. مدل ماشین بردار پشتیبان با استفاده از ترکیب ۱ پارامتر تبخیر را با ضریب همبستگی، مجذور میانگین مربعات خطا و ضریب نش- ساتکلیف به ترتیب برابر با ۰/۹۱۲، ۰/۹۷۱ و ۰/۸۲۸ در دوره صحت‌سنجی تخمین زده است.

پارامتر تبخیر را به ازای ترکیب ۱ در ایستگاه رشت با دقت زیادی انجام داده است. این مدل با استفاده از ترکیب ۱ در دوره صحت‌سنجی مقادیر آماره‌های ضریب همبستگی، مجذور میانگین مربعات خطا و ضریب نش- ساتکلیف را به ترتیب برابر با ۰/۸۸۰، ۰/۹۱۲ و ۰/۷۷۱ به دست آورده است. در تخمین تبخیر در این ایستگاه، مدل ماشین بردار پشتیبان به ترتیب به ازای دو ترکیب ۲ و ۳ نتایج نسبتاً مشابهی به دست آورده است.

مدل ماشین بردار پشتیبان نیز با استفاده از ترکیب‌های ۱ و ۳، پارامتر تبخیر در ایستگاه آستارا را به خوبی تخمین زده است. ترکیب ۳ با توجه به ضریب نش- ساتکلیف و مجذور میانگین مربعات خطا به عنوان

جدول ۶. نتایج مدل‌سازی ماشین بردار پشتیبان و شبکه عصبی مصنوعی

ایستگاه	شماره ترکیب	مدل	آموزش			صحت‌سنجی			ساختار
			NS	RMSE	CC	NS	RMSE	CC	
رشت	۱	ANN-EN1	۰/۷۵۲	۰/۹۵۰	۰/۸۶۷	۰/۷۸۲	۰/۸۹۱	۰/۸۸۶	(۱-۲-۱)
	۲	ANN-EN2	۰/۷۵۵	۰/۹۴۴	۰/۸۶۹	۰/۷۸۰	۰/۸۹۴	۰/۸۸۵	(۱-۴-۱)
	۳	ANN-GT	۰/۷۴۹	۰/۹۵۶	۰/۸۶۶	۰/۷۸۵	۰/۸۸۵	۰/۸۸۸	(۱-۶-۱)
	۱	SVM-EN1	۰/۷۶۲	۰/۹۳۰	۰/۸۷۳	۰/۷۷۱	۰/۹۱۲	۰/۸۸۰	-
	۲	SVM-EN2	۰/۷۶۱	۰/۹۳۴	۰/۸۷۲	۰/۷۷۰	۰/۹۱۴	۰/۸۷۹	-
	۳	SVM-GT	۰/۷۶۴	۰/۹۲۸	۰/۸۷۴	۰/۷۶۴	۰/۹۲۷	۰/۸۷۶	-
آستارا	۱	ANN-EN1	۰/۸۳۰	۰/۸۴۳	۰/۹۱۳	۰/۷۹۰	۱/۱۶۴	۰/۹۲۱	(۱-۱۲-۱)
	۲	ANN-EN2	۰/۸۳۲	۰/۸۴۰	۰/۹۱۲	۰/۷۷۰	۱/۲۱۶	۰/۹۲۰	(۱-۱۵-۱)
	۳	ANN-GT	۰/۸۳۳	۰/۸۳۶	۰/۹۱۳	۰/۷۷۶	۱/۲۰۲	۰/۹۱۹	(۱-۱۶-۱)
	۱	SVM-EN1	۰/۸۴۷	۰/۸۰۰	۰/۹۲۱	۰/۷۶۷	۱/۲۲۵	۰/۹۲۰	-
	۲	SVM-EN2	۰/۸۳۴	۰/۸۳۵	۰/۹۱۳	۰/۷۵۹	۱/۲۴۵	۰/۹۱۷	-
	۳	SVM-GT	۰/۸۳۷	۰/۸۲۸	۰/۹۱۵	۰/۷۶۸	۱/۲۲۲	۰/۹۱۹	-
انزلی	۱	ANN-EN1	۰/۸۴۲	۰/۹۰۶	۰/۹۱۸	۰/۷۹۸	۱/۰۶۲	۰/۸۹۷	(۱-۶-۱)
	۲	ANN-EN2	۰/۸۳۷	۰/۹۲۱	۰/۹۱۵	۰/۷۹۶	۱/۰۶۸	۰/۸۹۵	(۱-۶-۱)
	۳	ANN-GT	۰/۸۳۱	۰/۹۳۶	۰/۹۱۲	۰/۷۸۸	۱/۰۸۹	۰/۸۸۹	(۱-۵-۱)
	۱	SVM-EN1	۰/۸۳۴	۰/۹۳۸	۰/۹۱۳	۰/۸۲۸	۰/۹۷۱	۰/۹۱۲	-
	۲	SVM-EN2	۰/۸۴۱	۰/۹۱۰	۰/۹۱۷	۰/۷۹۳	۱/۰۷۶	۰/۸۹۵	-
	۳	SVM-GT	۰/۸۵۰	۰/۸۸۳	۰/۹۲۲	۰/۷۹۵	۱/۰۷۲	۰/۸۹۶	-

ترکیب‌ها نشان داده است. در این ایستگاه مدل شبکه عصبی مصنوعی-تئوری آنتروپی (بدون حذف متغیر بارندگی) به عنوان مدل برتر انتخاب شده است. ترکیب انتخابی توسط آزمون گاما در ایستگاه آستارا همانند ترکیب دوم تئوری آنتروپی یک متغیر را حذف کرده، اما عملکرد بهتری نسبت به ترکیب دوم تئوری آنتروپی در هر دو مدل داشته است. در ایستگاه انزلی آزمون گاما با هر دو مدل نسبت به ترکیب انتخابی توسط تئوری آنتروپی ضعیف عمل کرده است. در این

همان طور که در جدول ۶ مشاهده می‌شود، استفاده از ترکیب انتخابی توسط آزمون گاما برای تخمین تبخیر در ایستگاه رشت در مدل شبکه عصبی مصنوعی دقت بیشتری در تخمین تبخیر داشته است. در واقع، در ایستگاه رشت مدل شبکه عصبی مصنوعی-آزمون گاما نسبت به سایر مدل‌ها عملکرد بهتری داشته است. در ایستگاه آستارا ترکیب انتخابی توسط تئوری آنتروپی بدون حذف متغیر بارندگی در مدل شبکه عصبی مصنوعی عملکرد بهتری نسبت به سایر

آنتروپی نشان‌دهنده مؤثرتر بودن پارامتر مد نظر است. با توجه به شکل ۲ برای ایستگاه رشت پارامترهای مؤثر در تخمین تبخیر دمای حداکثر، دمای متوسط، ساعت‌های آفتابی، دمای حداقل، رطوبت حداقل، رطوبت متوسط و رطوبت حداکثر، بارندگی و سرعت باد به ترتیب به‌عنوان پارامترهای مؤثر در تخمین تبخیر شناخته شده‌اند. پارامترهای دمای حداکثر، دمای متوسط، دمای حداقل، ساعت‌های آفتابی، رطوبت متوسط، رطوبت حداقل، رطوبت حداکثر، بارندگی و سرعت باد برای ایستگاه آستارا برای تخمین تبخیر معرفی شدند. همچنین، برای ایستگاه انزلی می‌توان به‌ترتیب از پارامترهای دمای حداکثر، دمای متوسط، دمای حداقل، ساعت‌های آفتابی، رطوبت متوسط، رطوبت حداقل، رطوبت حداکثر، بارندگی و سرعت باد به‌عنوان مؤثرترین پارامتر در تخمین تبخیر در این ایستگاه نام برد. تخمین تبخیر در هر یک از سه ایستگاه رشت، آستارا و انزلی با استفاده از مدل‌های ANN و SVM انجام شده است که نتایج دوره تست آنها به طور خلاصه در شکل ۲ نشان داده شده است.

تقریباً، در همه شکل‌ها روند افزایش آنتروپی همراه با کاهش خطای مدل‌ها همراه بوده است. البته، در برخی موارد این روند صادق نیست که مقدار آن بسیار اندک و قابل چشم‌پوشی است. این تناقض‌ها در ایستگاه رشت بین رطوبت متوسط و رطوبت حداقل، ایستگاه آستارا دمای حداقل و ساعت‌های آفتابی و در ایستگاه انزلی بین رطوبت حداقل و رطوبت حداکثر بوده است که مقدار آنها بسیار ناچیز است. دلیل اصلی این تطابق‌نداشتن را می‌توان تأثیر یکسان دو پارامتری که مقدار آنتروپی آنها نزدیک بهم است، دانست. بنابراین، نمی‌توان این تطابق‌نداشتن بسیار کم را به عنوان نقضی برای عملکرد تئوری آنتروپی دانست. درخور یادآوری است که با توجه شکل ۲ در همه ایستگاه‌ها بنا بر تشخیص آنتروپی پارامتر سرعت باد کم‌اثرترین پارامتر شناخته شده است.

آزمون گاما

در گام بعدی عملکرد آزمون گاما در تعیین تک‌ورودی برای تخمین تبخیر در سه ایستگاه رشت، آستارا و بندرانزلی ارزیابی شد. در این بخش هرچقدر مقدار پارامتر گامی یک ورودی کمتر باشد، آزمون گاما آن پارامتر را مؤثرتر در تخمین تبخیر در ایستگاه مد نظر معرفی

ایستگاه مدل ماشین بردار پشتیبان-تئوری آنتروپی (بدون حذف متغیر بارندگی) به‌عنوان مدل برتر انتخاب شده است. درخور یادآوری است که منظور از ساختار تعیین‌شده در جدول ۶ نشان‌دهنده تعداد نورون‌ها در هر لایه است. برای مثال، برای ایستگاه رشت ترکیب ۱ در مدل شبکه عصبی مصنوعی ساختار (۱-۲-۹) دارد که بیان‌کننده ۹ ورودی، یک خروجی و تعداد ۲ نورون است که به صورت سعی و خطا برای لایه پنهان به‌دست آمده است.

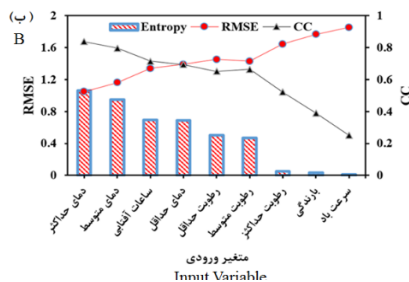
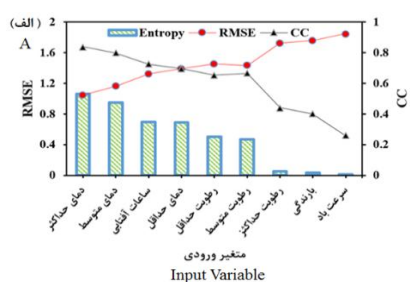
با توجه به جدول ۶ حذف پارامتر بارندگی در ترکیب انتخابی توسط تئوری آنتروپی سبب کاهش دقت هر دو مدل شبکه عصبی مصنوعی و ماشین بردار پشتیبان شده است، اما این کاهش دقت در مقابل کاهش پیچیدگی مدل و صرفه‌جویی که در وقت و هزینه توسط حذف این پارامتر از ترکیب می‌شود، قابل چشم‌پوشی است. در واقع، می‌توان گفت که به جای استفاده از همه متغیرها در تخمین تبخیر که توسط تئوری آنتروپی برای هر سه ایستگاه سینوپتیک انتخاب شده بود، می‌توان پارامتر بارندگی را حذف کرد و تبخیر را با دقت قابل قبولی در هر سه ایستگاه سینوپتیک تخمین زد. با توجه به جدول و موارد یادشده هر دو مدل شبکه عصبی مصنوعی و آزمون گاما در تخمین تبخیر در ایستگاه‌های انزلی، آستارا و رشت عملکرد قابل قبولی داشته‌اند. از بین دو روش انتخاب ورودی، از آزمون گاما به دلایل مدیریتی (صرفه‌جویی در وقت و هزینه، کاهش پیچیدگی مدل‌ها و...) به‌عنوان روش برتر نسبت به تئوری آنتروپی می‌تواند استفاده شود. درخور یادآوری است که مدل ماشین بردار پشتیبان در بخش آموزش بهتر از مدل شبکه عصبی مصنوعی عمل کرده است.

تعیین بهترین تک‌ورودی با تئوری آنتروپی و آزمون گاما تئوری آنتروپی

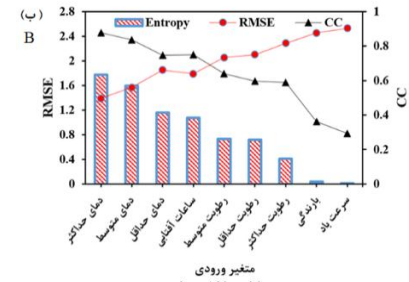
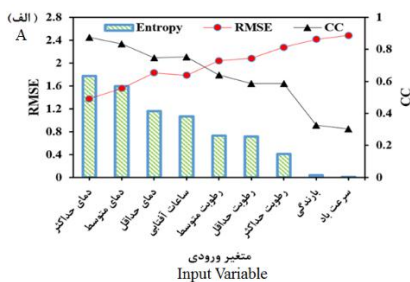
با توجه به شکل ۲ مقدار آنتروپی انتقال اطلاعات برای تک‌ورودی‌های مختلف در تخمین تبخیر برای هر یک از سه ایستگاه رشت، آستارا و انزلی نشان داده شده است. تک‌ورودی‌ها شامل دمای حداکثر، دمای حداقل، دمای متوسط، رطوبت حداکثر، رطوبت حداقل و رطوبت متوسط، بارندگی، ساعت‌های آفتابی و سرعت باد بوده است. با توجه به اینکه برای ارزیابی عملکرد تئوری آنتروپی از آنتروپی انتقال اطلاعات استفاده شده است، بنابراین افزایش مقدار این

همان طور که در شکل ۳ نشان داده شده است. تقریباً با افزایش مقدار گاما، مقدار خطای تخمین هر یک از دو مدل ANN و SVM در تخمین مقدار تبخیر در ایستگاه رشت افزایش پیدا می‌کند. این روند افزایش خطا در کنار افزایش مقدار گاما در تخمین تبخیر در دو ایستگاه آستارا و انزلی مشهود است. با توجه به شکل ۳ نتایج کلیه تخمین‌ها نشان می‌دهد آزمون گاما حداقل در تعیین سه تک‌ورودی اول کاملاً دقیق است. البته، نزدیکی مقدار خطای تک‌ورودی‌ها نیز به علت مقدار گامای بسیار نزدیک به یکدیگر آن پارامترهاست. البته، در خور یادآوری است که در برخی مواقع روند افزایش خطا با روند افزایش گاما هم‌خوانی ندارد (ایستگاه رشت در پارامترهای وسط و انتهایی، ایستگاه آستارا در پارامترهای وسط)، علت اصلی این هم‌خوانی نداشتن می‌تواند ناشی از تفاوت ساختاری مدل‌های به‌کاررفته (ANN و SVM) با آزمون گاما دانست. بنابراین، نمی‌توان برخی تطابق نداشتن‌ها در افزایش روند خطا با روند افزایش مقدار گاما را ناشی از عملکرد ضعیف آزمون گاما دانست.

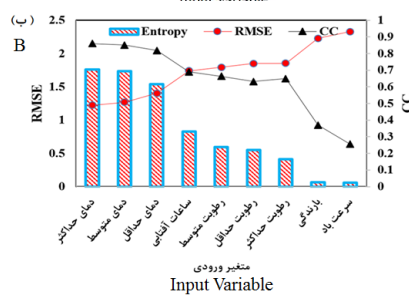
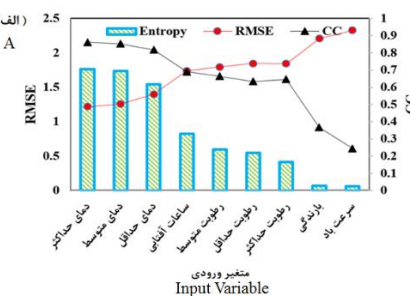
می‌کند. بنابراین، با توجه به شکل ۳، آزمون گاما به‌ترتیب استفاده از تک‌ورودی پارامترهای دمای حداکثر، دمای متوسط، ساعت‌های آفتابی، دمای حداقل، رطوبت حداقل، رطوبت متوسط، رطوبت حداکثر، بارندگی و سرعت باد را برای تخمین تبخیر در ایستگاه رشت پیشنهاد می‌دهد. آزمون گاما برای ایستگاه آستارا (به‌وسیله یک پارامتر ورودی) به‌ترتیب پارامترهای دمای حداکثر، دمای متوسط، دمای حداقل، ساعت‌های آفتابی رطوبت متوسط، رطوبت حداقل، رطوبت حداکثر، بارندگی و سرعت باد را پیشنهاد می‌دهد. معرفی می‌کند. همچنین، برای ایستگاه انزلی، به‌ترتیب پارامترهای دمای حداکثر، دمای متوسط، دمای حداقل، ساعت‌های آفتابی، رطوبت متوسط، رطوبت حداکثر، بارندگی و سرعت باد را پیشنهاد می‌دهد. تخمین تبخیر در هر یک از سه ایستگاه رشت، آستارا و انزلی با استفاده از مدل‌های ANN و SVM انجام شده است که نتایج دوره تست آنها به‌طور خلاصه در شکل ۳ نشان داده شده است.



رشت
Rasht



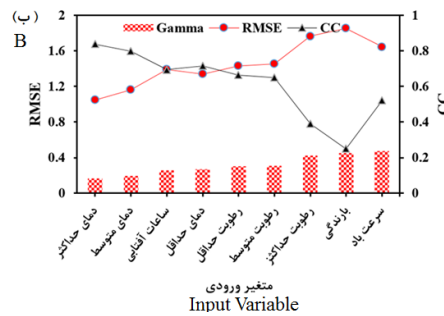
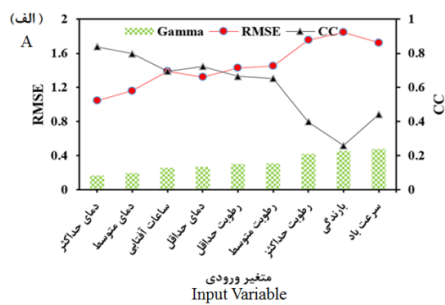
آستارا
Astara



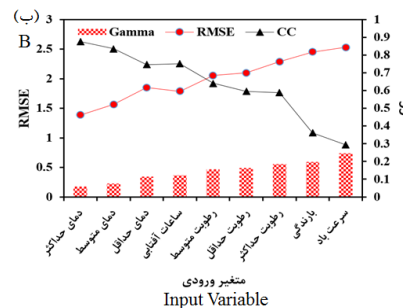
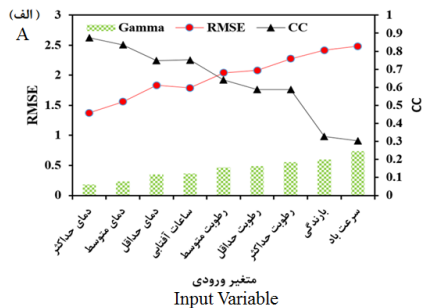
انزلی
Anzali

شکل ۲. نتایج تک‌متغیر ورودی برای تئوری آنتروپی

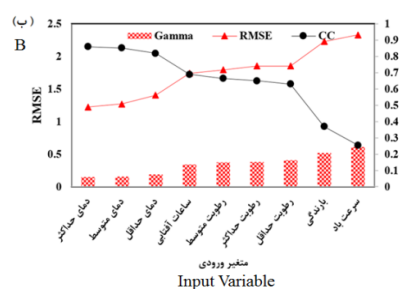
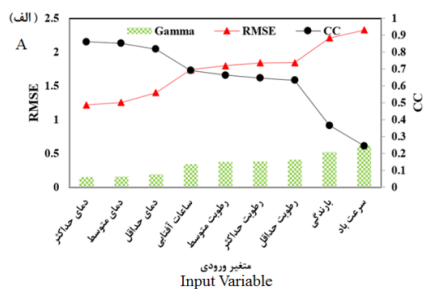
(شکل الف مربوط به شبکه عصبی مصنوعی و شکل ب مربوط به آزمون گاما است)، (واحد RMSE میلی‌متر بر روز است)



رشت
Rasht



آستارا
Astara



انزلی
Anzali

شکل ۳. نتایج تک‌متغیر ورودی برای آزمون گاما

(شکل الف مربوط به شبکه عصبی مصنوعی و شکل ب مربوط به آزمون گاما است)، (واحد RMSE میلی‌متر بر روز است)

صحیح منابع آب، در این تحقیق سعی شد با استفاده از مدل‌های شبکه عصبی مصنوعی و ماشین بردار پشتیبان مقادیر تبخیر در سه ایستگاه رشت، آستارا و بندر انزلی واقع در استان گیلان تخمین زده شود. به این منظور، از دو روش تئوری آنتروپی و آزمون گاما در بخش اول برای انتخاب ترکیب مؤثر استفاده شد. پارامترهای انتخابی توسط روش آزمون گاما در ایستگاه رشت دمای حداکثر، دمای متوسط، دمای حداقل، رطوبت حداقل، بارندگی، سرعت باد و ساعات‌های آفتابی، ایستگاه آستارا دمای حداکثر، دمای حداقل، دمای متوسط، رطوبت حداکثر، رطوبت حداقل، رطوبت متوسط، بارندگی و ساعات‌های آفتابی و برای ایستگاه انزلی دمای حداکثر، رطوبت حداکثر، رطوبت حداقل، بارندگی، سرعت باد و ساعات آفتابی انتخاب شدند. در روش تئوری آنتروپی همه متغیرها در هر سه ایستگاه به عنوان متغیرهای ورودی به

با توجه به شکل‌های ۲ و ۳ می‌توان گفت که پارامتر دمای حداکثر به عنوان مؤثرترین پارامتر توسط هر دو روش آزمون گاما و تئوری آنتروپی در تخمین تبخیر برای هر سه ایستگاه شناخته شده است. البته، درخور یادآوری است که تبخیر به عواملی غیر از دما نیز وابسته است. با توجه به شکل‌های ۲ و ۳ و عملکرد دو روش آزمون گاما و تئوری آنتروپی در تعیین تک‌ورودی مؤثر، می‌توان تئوری آنتروپی را به‌عنوان روش برتر در ایستگاه رشت و روش آزمون گاما را برای ایستگاه انزلی انتخاب کرد. برای ایستگاه آستارا هر دو روش یکسان عمل کرده‌اند. بنابراین، بسته به منطقه مطالعه شده این دو روش عملکرد متفاوتی دارند. در کل، هر دو روش سه تک‌ورودی اول را به‌خوبی تشخیص داده‌اند.

نتیجه‌گیری

با توجه به اهمیت اطلاع از میزان تبخیر در مدیریت

- models. *Water Science and Technology*. 2017:wst2017234.
- [6]. Seefi A, Mirlatifi M, Reahi H. Introduction and application of Least Square Support Vector Machine (LSSVM) for simulation of reference evaporation and uncertainty analysis of results, a case study fo the Kerman city. *Irrigation & Water Engineering*. 2013; 13 (5):67-78. (In Persian)
- [7]. Kim S, Shiri J, Kisi O, Singh VP. Estimating daily pan evaporation using different data-driven methods and lag-time patterns. *Water resources management*. 2013; 27(7):2267-86.
- [8]. Goyal MK, Bharti B, Quilty J, Adamowski J, Pandey A. Modeling of daily pan evaporation in sub-tropical climates using ANN, LS-SVR, Fuzzy Logic, and ANFIS. *Expert systems with applications*. 2014; 41(11):5267-76.
- [9]. Tezel G, Buyukyildiz M. Monthly evaporation forecasting using artificial neural networks and support vector machines. *Theoretical and applied climatology*. 2016; 124(1-2):69-80.
- [10]. Kisi O. Pan evaporation modeling using least square support vector machine, multivariate adaptive regression splines and M5 model tree. *Journal of Hydrology*. 2015; 528:312-20.
- [11]. Kisi O, Genc O, Dinc S, Zounemat-Kermani M. Daily pan evaporation modeling using chi-squared automatic interaction detector, neural networks, classification and regression tree. *Computers and Electronics in Agriculture*. 2016; 122:112-7.
- [12]. Keshtegar B, Piri J, Kisi O. A nonlinear mathematical modeling of daily pan evaporation based on conjugate gradient method. *Computers and Electronics in Agriculture*. 2016; 127:120-30.
- [13]. Sharifi A.R, Dinpashih Y, Fahkerifard A, Moghadamnia, AR. Optimum combination of variables for runoff simulation in Amameh Watershed using Gamma Test, *Journal of Soil and Water*. 2013; 23(3):72-59. (In Persian)
- [14]. Durrant PJ. winGamma: A non-linear data analysis and modelling tool with applications to flood prediction. Unpublished PhD thesis, Department of Computer Science, Cardiff University, Wales, UK. 2001 Jun 25.
- [15]. Evans D, Jones AJ. A proof of the Gamma test. In *Proceedings of the Royal Society of London a: Mathematical, Physical and Engineering Sciences* 2002 Nov 8 (Vol. 458, No. 2027, pp. 2759-2799). The Royal Society.
- [16]. Shannon C. E, & Weaver W, Urban:University of Illinois Press; 1949.
- مدل انتخاب شدند. با توجه به اینکه در روش تئوری آنتروپی با حذف متغیر بارندگی میزان آنتروپی انتقال اطلاعات تغییر چندانی نسبت به ترکیب اول رخ نداده است، بنابراین یک ترکیب دیگر توسط تئوری آنتروپی انتخاب شد. در ترکیب دوم انتخابی توسط تئوری آنتروپی متغیر بارندگی حذف شده است.
- نتایج مدل‌سازی نشان داد در ایستگاه رشت، مدل شبکه عصبی مصنوعی-آزمون گاما، در ایستگاه آستارا شبکه عصبی مصنوعی- تئوری آنتروپی (بدون حذف متغیر بارندگی) و برای ایستگاه انزلی ماشین بردار پشتیبان- تئوری آنتروپی (بدون حذف متغیر بارندگی) را می‌توان به‌عنوان مدل‌های برتر نام برد. در بخش دوم قدرت تئوری آنتروپی و آزمون گاما در تعیین تک‌ورودی ارزیابی شد. براساس نتایج، می‌توان تئوری آنتروپی را به‌عنوان روش برتر در ایستگاه رشت و روش آزمون گاما را برای ایستگاه انزلی انتخاب کرد. برای ایستگاه آستارا هر دو روش یکسان عمل کرده‌اند. بنابراین، بسته به منطقه مطالعه‌شده این دو روش عملکرد متفاوتی دارند. در کل، هر دو روش سه تک‌ورودی اول را به‌خوبی تشخیص داده‌اند. در نهایت، با توجه به اینکه روش آزمون گاما از لحاظ مدیریتی نسبت به تئوری آنتروپی عملکرد بهتری داشته (با تعداد کم ورودی عملکرد قابل قبول داشته است) و همچنین در تخمین تبخیر نیز عملکرد قابل قبولی داشته است.

منابع

- [1]. Ejlali F, *Weather and climatology*. Iran: Payamnoor University Press; 2004. (In Persian)
- [2]. Nourani V, Baghanam AH, Adamowski J, Kisi O. Applications of hybrid wavelet-Artificial Intelligence models in hydrology: A review. *Journal of Hydrology*. 2014; 514:358-77.
- [3]. Ahmadi A, Han D, Karamouz M, Remesan R. Input data selection for solar radiation estimation. *Hydrological processes*. 2009; 23(19):2754-64.
- [4]. Moghaddamnia A, Gousheh MG, Piri J, Amin S, Han D. Evaporation estimation using artificial neural networks and adaptive neuro-fuzzy inference system techniques. *Advances in Water Resources*. 2009; 32(1):88-97.
- [5]. Sharifi A, Dinpashoh Y, Mirabbasi R. Daily runoff prediction using the linear and non-linear

- [17]. Dawson CW, Abrahart RJ, Shamseldin AY, Wilby RL. Flood estimation at ungauged sites using artificial neural networks. *Journal of hydrology*. 2006; 319(1):391-409.
- [18]. Harmancioglu NB, Alpaslan N. WATER QUALITY MONITORING NETWORK DESIGN: A PROBLEM OF MULTI-OBJECTIVE DECISION MAKING. *JAWRA Journal of the American Water Resources Association*. 1992; 28(1):179-92.
- [19]. Coulibaly P, Anctil F, Bobee B. Daily reservoir inflow forecasting using artificial neural networks with stopped training approach. *Journal of Hydrology*. 2000; 230(3):244-57.
- [20]. ASCE Task Committee. Artificial neural networks in hydrology. I: Preliminary concepts. *Journal of Hydrologic Engineering*. 2000; 5(2):115-23.
- [21]. Kavzoglu T, Colkesen I. A kernel functions analysis for support vector machines for land cover classification. *International Journal of Applied Earth Observation and Geoinformation*. 2009; 11(5):352-9.
- [22]. Nash J.E, Sutcliffe I.V. River flow forecasting through conceptual models, Part I, A discussion of principles, *Journal of Hydrology*. 1970; 10(2):282-290.