

برآورد تبخیر روزانه از تشت تبخیر با استفاده از روش‌های یادگیری ماشینی

سعید صمدیان فرد^{۱*}، سجاد هاشمی^۲، مجتبی ایزدیاری^۳

تاریخ دریافت: ۱۳۹۶/۱۲/۲۳ تاریخ پذیرش: ۱۳۹۷/۳/۲۲

چکیده

تبخیر از اصلی‌ترین فرآیندها در چرخه آبی طبیعت و از مهم‌ترین عوامل در مطالعات کشاورزی، هیدرولوژیکی، هواشناسی، بهره‌برداری مخازن، طراحی سیستم‌های آبیاری و زهکشی، زمان‌بندی آبیاری و مدیریت منابع آب می‌باشد. بنابراین برآورد دقیق مقدار تبخیر از اهمیت بالایی در مطالعات هیدرولوژیکی برخوردار است. در این راستا و در تحقیق حاضر، از روش‌های هوشمند برنامه‌ریزی ژنتیک، رگرسیون بردار پشتیبان و شبکه عصبی مصنوعی برای برآورد و شبیه‌سازی مقادیر تبخیر از تشت در ایستگاه‌های هواشناسی تبریز و جلفا استفاده شد بدین منظور، از داده‌های هواشناسی تبخیر، دما، رطوبت نسبی، سرعت باد و تابش خورشیدی ایستگاه‌های مذکور در بازه زمانی بیست ساله (۱۳۷۱-۱۳۹۰) استفاده شد و دقت روش‌های مورد مطالعه با استفاده از پارامترهای آماری جذر میانگین مربعات خطا، میانگین خطای مطلق، ضریب همبستگی و دیاگرام تیلور مورد بررسی قرار گرفت. نتایج حاصل از این پژوهش نشان دادند که در بهینه‌ترین حالت و به ترتیب در ایستگاه‌های تبریز و جلفا، برنامه‌ریزی ژنتیک با دارا بودن خطای ۲/۶۸ و ۲/۱۸، رگرسیون بردار پشتیبان با خطای ۲/۱۹ و ۲/۲۲ و شبکه عصبی مصنوعی با خطای ۲/۱۴ و ۲/۲۱ عملکرد مناسبی در شبیه‌سازی مقدار تبخیر داشته‌اند. در نهایت برای ایستگاه تبریز سناریو دوم روش شبکه عصبی مصنوعی با پارامترهای ورودی دما و سرعت باد و برای ایستگاه جلفا سناریو هفتم روش شبکه عصبی مصنوعی با پارامترهای ورودی دما، رطوبت نسبی، سرعت باد و تابش خورشیدی و دارا بودن بهترین عملکرد، به‌عنوان مدل‌هایی با دقت مناسب برای شبیه‌سازی مقدار تبخیر از تشت پیشنهاد گردید.

واژه‌های کلیدی: تحلیل آماری، توابع کرنل، شبکه عصبی مصنوعی، مدیریت منابع آب

مقدمه

تبخیر از اصلی‌ترین فرآیندها در چرخه آبی طبیعت و از مهم‌ترین عوامل در مطالعات کشاورزی، هیدرولوژیکی، هواشناسی، بهره‌برداری مخازن، طراحی سیستم‌های آبیاری و زهکشی، زمان‌بندی آبیاری و مدیریت منابع آب می‌باشد. هر ساله میلیون‌ها مترمکعب آب شیرین از مخازن سدها که با هزینه زیادی جمع‌آوری گردیده، تبخیر شده و از دست می‌رود و املاح و نمک برجای مانده از حجم آب تبخیر شده، کیفیت آب را کاهش می‌دهد (کارآموز و عراقی‌نژاد، ۱۳۹۳). از جمله روش‌های اندازه‌گیری تبخیر، استفاده از تشت تبخیر استاندارد است که در اکثر ایستگاه‌های هواشناسی موجود می‌باشد، به طوری که این روش به‌عنوان یکی از روش‌ها با دقت بالا برای برآورد تبخیر و تعرق

گیاه مرجع توسط رحیمی خوب (۱۳۸۶) معرفی شده است. قره‌خانی و همکاران (۱۳۸۹) روند تغییرات تبخیر از تشت را در سراسر ایران بررسی کرده و بیان داشتند که علی‌رغم افزایش دما، در بلندمدت کاهش میزان تبخیر از تشت مشاهده می‌شود. همچنین در پژوهشی که توسط زارع ایبانه و همکاران (۱۳۸۹) انجام شد، داده‌های تشت تبخیر با نتایج حاصل از ۹ روش تجربی برآورد تبخیر مقایسه شد و در نهایت چنین نتیجه گرفته شد که روش ایوانف بیش‌ترین همخوانی را با داده‌های اندازه‌گیری شده داراست. نبود آمار کافی در سنوات گذشته و بالا بودن هزینه‌ی راه‌اندازی تشت تبخیر در پروژه‌های تحقیقاتی، محققان را بر آن داشته از روش‌های تجربی و تحلیلی که داده‌های هواشناسی را بکارمی‌برند، به جای روش موجود استفاده کنند. از جمله روش‌های تجربی، استفاده از الگوریتم‌های هوش مصنوعی همانند ماشین بردار پشتیبان، شبکه‌ی عصبی مصنوعی و برنامه‌ریزی ژنتیک است که برای استخراج رابطه بین داده‌های هواشناسی و مقدار تبخیر از تشت کاربرد گسترده‌ای بین محققان پیدا کرده است. روش‌های الهام گرفته شده از طبیعت از جمله برنامه‌ریزی ژنتیک، جزو مدل‌هایی هستند که در تحقیقات پیچیده و دقیق از آن‌ها استفاده می‌شود. برنامه‌ریزی ژنتیک، تکنیک برنامه‌ریزی خودکار می‌باشد که راه حل

۱- استادیار، گروه مهندسی آب، دانشکده کشاورزی، دانشگاه تبریز
۲- دانشجوی کارشناسی مهندسی آب، گروه مهندسی آب، دانشکده کشاورزی، دانشگاه تبریز
۳- دانشجوی کارشناسی مهندسی آب، گروه مهندسی آب، دانشکده کشاورزی، دانشگاه تبریز
(* - نویسنده مسئول: (Email: s.samadian@tabrizu.ac.ir

اقلیم و آب و هوای اکثر شهرهای این استان کوهستانی است ولی با این حال می‌توان شهرهایی با آب و هوای متفاوت نیز پیدا کرد. از این رو، در این تحقیق از آمارهای هواشناسی دو ایستگاه سینوپتیک تبریز و جلفا استفاده شد. این دو شهر از نظر تقسیم‌بندی دوماتن به ترتیب دارای اقلیم کشاورزی نیمه‌خشک و خشک می‌باشند و با توجه به تفاوت اقلیم این دو شهر می‌توان تجزیه و تحلیل بهتری نسبت به نتایج به‌دست آمده انجام داد. ایستگاه تبریز دارای طول جغرافیایی ۳۶ درجه و ۱۷ دقیقه و عرض جغرافیایی ۳۸ درجه و ۵ دقیقه و ارتفاع ۱۳۶۴ متر از سطح دریا و ایستگاه جلفا دارای طول جغرافیایی ۴۵ درجه و ۳۶ دقیقه و عرض جغرافیایی ۳۸ درجه و ۵۶ دقیقه و ارتفاع ۷۳۶ متر از سطح دریا می‌باشد. در این تحقیق، از پارامترهای هواشناسی رطوبت نسبی (درصد)، درجه حرارت (درجه سانتی‌گراد)، سرعت باد (متر بر ساعت)، تابش خورشیدی (مگاژول بر مترمربع بر روز) و تبخیر (میلی‌متر) ایستگاه‌های استفاده شد که از سازمان هواشناسی استان آذربایجان شرقی اخذ گردید و از بازه‌ی زمانی بیست ساله (۱۳۷۱-۱۳۹۰) استفاده شد. مقادیر برخی از کمیت‌های آماری این داده‌های هواشناسی در جدول ۱ آمده است.

برنامه‌ریزی ژنتیک^۱ (GP)

برنامه‌ریزی ژنتیک تعمیم یافته الگوریتم ژنتیک است که در سال ۱۹۹۲ توسط کوزا برای اولین بار ارائه شد (Kozá, 1992). مبنای این روش بر اساس نظریه تکاملی داروین استوار است. به این ترتیب که الگوریتم یاد شده اقدام به تعریف یک تابع هدف در قالب معیارهای کیفی نموده و سپس تابع هدف را برای اندازه‌گیری و مقایسه روش‌های مختلف حل کرده و در یک فرآیند گام به گام تصحیح ساختار داده‌ها به‌کار گرفته و در نهایت، روش حل مناسب را ارائه می‌نماید. برنامه‌ریزی ژنتیک با استفاده از برنامه‌ی کامپیوتری و یک تکنیک برنامه‌ریزی خودکار راه حل مسئله را ارائه می‌کند. در این روش در ابتدای فرآیند هیچ‌گونه رابطه‌ی تابعی در نظر گرفته نشده و این روش می‌تواند ساختار مدل و مولفه‌های آن را بهینه کند. تفاوت اساسی موجود بین برنامه‌ریزی ژنتیک و الگوریتم ژنتیک در این است که برنامه‌ریزی ژنتیک روی ساختار درختی فرمول‌ها به‌جای سلسله ارقام دو دویی عمل می‌کند. ساختارهای درختی از مجموعه توابع (عمل‌گرهای ریاضی مورد استفاده در فرمول‌ها) و ترمینال‌ها (متغیر-های مسئله و اعداد ثابت) ایجاد می‌شوند. براساس تحقیقات گذشته می‌توان با استفاده از محاسبات نرم (مثل برنامه‌ریزی ژنتیک، شبکه‌های عصبی، رگرسیون بردار پشتیبان و ...) بر بعضی از مشکلات مدل‌های فیزیکی غلبه کرد و این روش‌ها را به عنوان یک راه حل موثر و کارآمد برای مدل کردن فرآیندهای فیزیکی در حالت فقدان

مسئله را با استفاده از برنامه‌نویسی کامپیوتری ارائه می‌کند. الگوریتم-های تکاملی که براساس تئوری داروین پایه‌ریزی شده‌اند توانایی مدل‌سازی فرآیندهای کاملاً غیرخطی را دارند که برنامه‌ریزی ژنتیک نیز عضوی از خانواده الگوریتم‌های تکاملی می‌باشد. سلطانی و همکاران (۱۳۸۹) در پژوهشی نشان دادند برنامه‌ریزی ژنتیک علاوه بر داشتن توانایی استخراج رابطه بین متغیرهای ورودی و خروجی به‌طور خودکار و هوشمند، متغیرهایی که در مدل بیش‌ترین تاثیر را دارند انتخاب می‌کند. ترزی با استفاده از روش برنامه‌ریزی ژنتیک، تبخیر روزانه از سطح دریاچه ایدر ترکیه را مدل کرد، او بیان داشت که مدل برنامه‌ریزی ژنتیک می‌تواند جایگزین روش تشت تبخیر گردد (Terzi, 2010). ستاری و اسمعیل‌زاده (۱۳۹۵) میزان تبخیر تعرق را با سه روش فائو پنمن - مانیت، برنامه‌ریزی ژنتیک و مدل درختی M5 را برای منطقه‌ی تبریز برآورد کرده و بیان نمودند که برنامه‌ریزی ژنتیک نتایج تقریباً مشابهی با روش استاندارد (روش فائو پنمن - مانیت) بدست می‌دهد. قالحری و شاکری (۱۳۹۴) در پژوهشی در مورد پیش‌بینی بارش زمستانه بیان داشتند که شبکه عصبی مصنوعی قادر به پیش‌بینی بارش با دقت قابل‌قبولی است. مهدوی و همکاران (۱۳۹۰) در پژوهشی در حاشیه‌ی دریای خزر، میزان تبخیر را به وسیله‌ی پارامترهای هواشناسی با ۸ ترکیب مختلف توسط شبکه عصبی مصنوعی مدل کرده و با مقادیر ماهانه تبخیر از تشت مقایسه کردند. آن‌ها بیان داشتند پارامترهای دمای بیشینه و کمینه هوا، رطوبت نسبی، سرعت باد و ساعات آفتابی، حداقل داده‌های هواشناسی برای برآورد تبخیر از تشت هستند. شایان‌نژاد و ساداتی‌نژاد (۱۳۸۷) کارایی شبکه‌ی عصبی مصنوعی در برآورد تبخیر - تعرق گیاه چمن را با روش پنمن - مانیت مقایسه کردند. ایشان اذعان داشتند که روش شبکه‌ی عصبی مصنوعی نتایج بهتری را نسبت به روش مذکور بدست می‌دهد. صمدیان‌فرد و اسدی (۱۳۹۶) نمایه‌ی خشک‌سالی SPI را به دو روش رگرسیون بردار پشتیبان و خطی چندگانه پیش‌بینی کردند. آن‌ها بیان داشتند که روش رگرسیون بردار پشتیبان در دوره‌های ۶، ۹ و ۲۴ ماهه نتایج بهتری را نسبت به رگرسیون خطی چندگانه ارائه می‌دهد. با عنایت به موارد فوق هدف از پژوهش حاضر یافتن دقیق‌ترین و مناسب‌ترین ترکیب پارامترهای هواشناسی به عنوان ورودی روش‌های برنامه‌ریزی ژنتیک، شبکه‌ی عصبی مصنوعی و رگرسیون بردار پشتیبان و تخمین هر چه بهتر تبخیر روزانه از تشت می‌باشد.

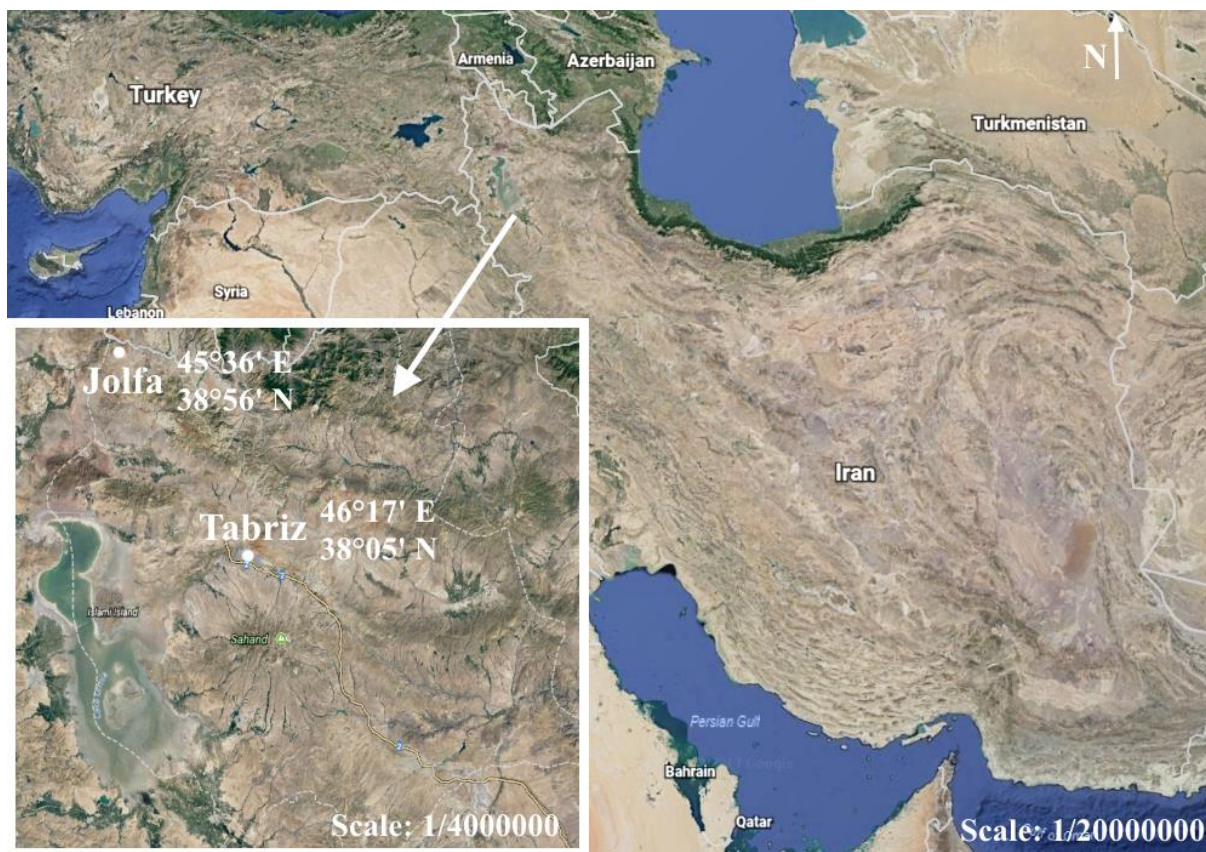
مواد و روش‌ها

منطقه مورد مطالعه

استان آذربایجان شرقی از استان‌های مهم ایران است که با استان‌های آذربایجان غربی، اردبیل و زنجان هم‌جوار است (شکل ۱).

مدل که بر مبنای آن می‌توان مشخص نمود که توانایی مدل در حل یک مسئله خاص تا چه اندازه می‌باشد (۴) مولفه‌های کنترل: مقادیر مولفه‌های عددی و متغیرهای کیفی که برای کنترل اجرای برنامه‌ها استفاده می‌شوند (۵) شرط توقف اجرای برنامه: که معیاری برای حصول نتیجه و توقف اجرای برنامه می‌باشد.

داده کافی معرفی کرد. قبل از مراحل اجرایی برنامه‌ریزی ژنتیک گام-های مقدماتی زیر باید توسط کاربر تعیین شوند (Koza., 1992). (۱) انتخاب مجموعه ترمینال: که همان متغیرهای مستقل مسئله و متغیرهای حالت سامانه می‌باشند (۲) انتخاب مجموعه توابع: که شامل عملگرهای حسابی و توابع آزمون می‌باشد (۳) شاخص اندازه‌گیری دقت



شکل ۱- تصویر منطقه مورد مطالعه

جدول ۱- کمیت‌های آماری پارامترهای هواشناسی ایستگاه‌های مورد پژوهش از سال ۱۳۷۱ تا ۱۳۹۰

پارامتر	کمینه		بیشینه		میانگین		انحراف معیار	
	تبریز	جلفا	تبریز	جلفا	تبریز	جلفا	تبریز	جلفا
دما	-۱۴/۹۸	۲/۱۳	۳۴/۰۳	۳۷/۱۲	۱۳/۰۳	۲۱/۹۱	۱۰/۲۲	۷/۰۰
رطوبت نسبی	۱۳/۸۸	۱۵/۲۵	۹۶/۳۸	۹۳/۱۳	۵۰/۸۹	۴۷/۶۸	۱۷/۵۷	۱۳/۸۲
سرعت باد	۰	۰	۱۳/۲۵	۱۴/۳۸	۳/۳۴	۳/۱۰	۱/۵۹	۲/۵۵
تابش خورشیدی	۰	۰	۹۹/۱۵	۱۳/۵۰	۱۵/۲۵	۹/۱۱	۱۰/۹۶	۳/۳۱
تبخیر	۰	۰/۲۰	۳۰/۴۰	۳۲/۰۰	۵/۴۰	۹/۹۷	۵/۳۷	۵/۶۲

لازم به ذکر است که پارامترهای مورد مطالعه دارای واحدهای دما (سانتی‌گراد)، رطوبت نسبی (درصد)، سرعت باد (متر بر ثانیه)، تابش خورشیدی (مگاژول بر مترمربع بر روز) و تبخیر (میلی‌متر) می‌باشد.

یادگیری آماری است که بوسه و همکاران معرفی کردند و در طی سال‌های بعد تئوری ابر صفحه‌ی بهینه‌را به عنوان طبقه‌بندی کننده خطی ارایه دادند (Boser et al., 1992). آن‌ها همچنین با کمک

رگرسیون ماشین بردار پشتیبان^۱ (SVR)

ماشین بردار پشتیبان یک روش یادگیری بر پایه‌ی تئوری

1- Support Vector Machine

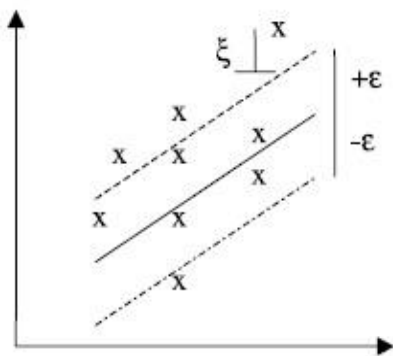
تابع زیان $(-\epsilon)$ غیرحساس که به صورت رابطه ۱ تعریف می شود.

$$L^\epsilon(x, y, f) = \begin{cases} 0 \\ |y - f(x)| - \epsilon \end{cases} \quad (1)$$

اگر $|y - f(x)| \leq \epsilon$ باشد آن گاه تابع زیان برابر صفر و در غیر این صورت برابر $|y - f(x)| - \epsilon$ خواهد بود. مجموعه داده های آموزشی به صورت $S = \{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$ و کلاس تابع به صورت $f(x) = \{w^T + b, w \in R^m, b \in R\}$ می باشد. در صورتی که داده ها از مقدار ϵ تخطی نمایند می بایست متناسب با مقدار تخطی، متغیر کمبود تعریف شود که در شکل ۲ نمونه ای از این تخطی نمایش داده شده است (Vapnik et al., 1997، غضنفری هاشمی و اعتماد شهیدی، ۱۳۹۱ و اسدپور و ستاری، ۱۳۹۴).

$$\text{Minimize } \frac{1}{2} \|w\|^2 + C \sum_{i=1}^N (\xi_i + \xi_i^*) \text{ Subject to } \begin{cases} f(x) - y_i \leq \epsilon + \xi_i \\ f(x) - y_i \leq \epsilon + \xi_i^* \\ \xi_i, \xi_i^* \geq 0 \end{cases} \quad (2)$$

1,2,3, ...,

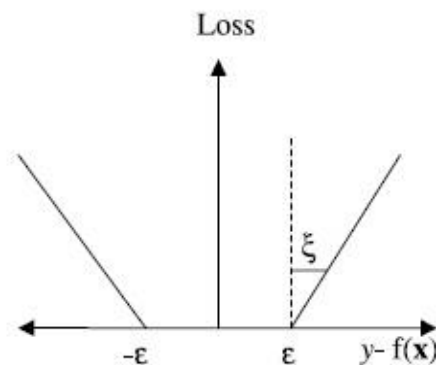


شکل ۲- میزان تخطی داده ها از مقدار مجاز ϵ

داده های مفروض، تعریف می شود. در واقع با یک تبدیل غیرخطی از فضای ورودی به فضایی با ابعاد بیش تر (حتی نامتناهی) می توان مسایل را به صورت خطی، تفکیک پذیر ساخت. در تحقیق حاضر از سه تابع کرنل مهم که عموماً در کاربردهای مهندسی متداول هستند استفاده شد (جدول ۲).

توابع کرنل، طبقه بندی کننده های غیرخطی را معرفی نمودند. بعدها با استفاده از نتایج کارهای ایشان، اصول اساسی ماشین بردار پشتیبان به وجود آمد و در نهایت وپنیک ماشین بردار پشتیبان را براساس رگرسیون گسترش داد (Vapnik., 1995). به طور کلی ماشین بردار پشتیبان برای طبقه بندی مسایل دو یا چند کلاسه به صورت خطی یا غیرخطی و رگرسیون مورد استفاده قرار می گیرد. در این روش، داده ها به دو دسته آموزش و آزمایش تقسیم می شوند.

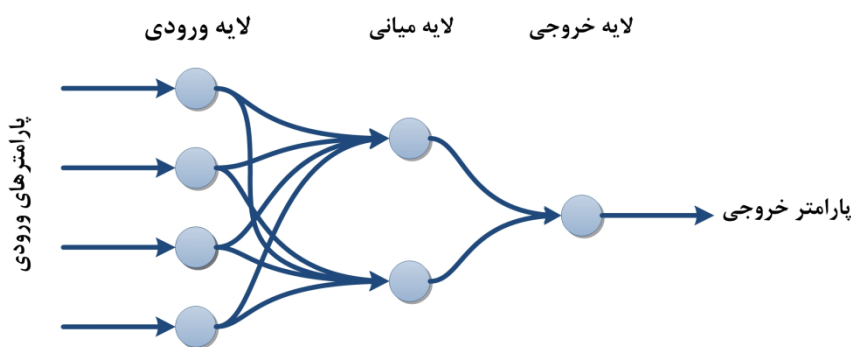
با استفاده از رگرسیون بر روی داده های برآزش داده شده یک ابرصفحه بدست می آید. خطای هر نقطه خاص را فاصله ی آن نقطه از ابرصفحه نشان می دهد. روش حداقل مربعات بهترین روشی است که تاکنون برای رگرسیون خطی پیشنهاد شده است. اما برای مسایل رگرسیون، این امکان وجود دارد که استفاده از برآورد کننده کمترین مربعات در حضور داده های پرت، بطور کامل، شدنی نباشد و در نتیجه رگرسیون عملکرد ضعیفی را از خود به نمایش بگذارد. بنابراین برای جلوگیری از عملکرد ضعیف باید یک برآورد کننده نیرومند که نسبت به تغییرات کوچک در مدل حساس نباشد را توسعه داد. در واقع، یک



در رابطه ۲، نرم بردار وزن و ξ_i^* و ξ_i متغیرهای کمبود کمکی هستند که در شکل ۲ نشان داده شده اند و پارامتر C ضریب تعادل پیچیدگی میان ماشین و تعداد نقاط تفکیک ناپذیر است که با سعی و خطا بدست می آید. از روش های متداول برای حل مسایل غیرخطی، استفاده از توابع کرنل می باشد که براساس ضرب داخلی

جدول ۲- توابع کرنل مورد استفاده در تحقیق حاضر

تابع ریاضی	توابع کرنل
$k(x_i, x_j) = (x_i^T x_j + 1)^p$	تابع کرنل چندجمله ای
$k(x_i, x_j) = e^{-\gamma x_i - x_j ^2}$	تابع کرنل با پایه شعاعی
$k(x_i, x_j) = \frac{1}{\left[1 + \left(2 \sqrt{\ x_i - x_j\ ^2} \cdot \sqrt{2^{1/\omega} \cdot \frac{1}{\sigma}} \right)^2 \right]^{2\omega}}$	تابع کرنل پیرسون



شکل ۳- نمایی از پرسپترون چندلایه (MLP)

n ، برای تابع f خواهد بود.

تابع غیرخطی f که انتقال اعداد از لایه‌ای به لایه دیگر را بر عهده دارد باید به لحاظ ریاضی هموار باشد، که از آن جمله می‌توان به توابع گوسی، سکانت هیپربولیک^۳، سیگموئید^۴ و تانژانت هیپربولیک^۵ اشاره کرد و از این میان تابع سیگموئید $S(x) = 1/(1 + \exp(-a))$ کاربرد بیشتری در مسایل مهندسی دارد. بنابراین در یک شبکه عصبی با n نرون ورودی، h نرون در لایه میانی و یک نرون خروجی، مقدار خروجی از شبکه عصبی به صورت رابطه‌ی ۴ خواهد بود.

$$Y = S_2 \left(\sum_{j=1}^n w_j \times \left(S_1(b_j + \sum_{i=1}^n w_{i,j} \cdot p_i) \right) \right) \quad (4)$$

معیارهای ارزیابی دقت

مقادیر خطای بین سه روش محاسباتی و داده‌های مشاهداتی توسط جذر میانگین مربعات خطا^۶ (RMSE)، میانگین خطای مطلق^۷ (MAE) و ضریب همبستگی^۸ (R) و با استفاده از روابط ۵ تا ۷ مورد بررسی قرار گرفت.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - y_i)^2} \quad (5)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |x_i - y_i| \quad (6)$$

$$R = \frac{(\sum_{i=1}^n x_i y_i - \frac{1}{n} \sum_{i=1}^n x_i \sum_{i=1}^n y_i)}{(\sum_{i=1}^n x_i^2 - \frac{1}{n} (\sum_{i=1}^n x_i)^2) (\sum_{i=1}^n y_i^2 - \frac{1}{n} (\sum_{i=1}^n y_i)^2)} \quad (7)$$

در روابط ۵ تا ۷، x_i و y_i مقادیر تخییر از تشت مشاهداتی و شبیه‌سازی شده و n تعداد مشاهدات می‌باشد. علاوه بر این، دیاگرام تیلور برای تحلیل دقت روش‌های مورد استفاده در شبیه‌سازی تخییر روزانه به کار گرفته شد (Taylor., 2001). دیاگرام تیلور، راه‌حلی گرافیکی برای ارزیابی دقت داده‌های شبیه‌سازی شده با به تصویر

شبکه عصبی مصنوعی^۱ (ANN)

شبکه‌های عصبی مصنوعی دارای الگوی پردازش اطلاعات هستند و ابزاری قدرتمند است که از شبیه‌سازی نرون‌های زیستی الهام گرفته شده‌است. بنابراین شبکه‌های عصبی مصنوعی مانند ساختار بیولوژیکی مغز انسان ویژگی‌هایی هم‌چون قدرت یادگیری، تعمیم‌دهی، پردازش موازی و تصمیم‌گیری را ترکیب می‌کند و می‌تواند مسایلی پیچیده با ماهیتی خطی یا غیرخطی را به‌طور کامل حل کند. به عبارت دیگر شبکه عصبی مصنوعی ساده شده‌ی تابع ریاضی مغز انسان است که از چند لایه و در هر لایه از چند گره (نرون) تشکیل شده‌است. مجموعه نرون‌های پایه به‌وجود آمده که لایه‌ی ورودی، یک یا چند لایه پنهان و یک لایه‌ی خروجی را تشکیل می‌دهند، معمولی‌ترین نوع شبکه عصبی مصنوعی است (شکل ۳). داده‌های ورودی به‌صورت لایه، لایه در خلال شبکه و در مسیری رو به جلو منتشر می‌شوند که به این نوع شبکه عصبی پرسپترون چندلایه^۲ (MLP) گفته می‌شود. تعداد نرون‌های لایه ورودی با تعداد عناصر بردار ورودی و تعداد نرون‌های لایه خروجی با تعداد عناصر بردار خروجی برابر می‌باشد (منه‌اج، ۱۳۸۴).

پیدا کردن تعداد نرون‌های لایه میانی و تحلیل دقیق آن‌ها در کل بسیار پیچیده است. اما با توجه به این که تعداد نرون‌های لایه مخفی تابعی از تعداد عناصر برداری ورودی و هم‌چنین بیشینه تعداد نواحی از فضای ورودی که به‌طور خطی از هم جدا پذیرند می‌باشد، می‌توان تعداد نرون‌های لایه مخفی به‌طور تجربی به‌دست آورد. خروجی نرون‌ها توسط معادله ۳ تعریف می‌گردد.

$$a = f(b_j + \sum_{i=1}^n p_i w_i) \quad (3)$$

که در این معادله p و a به‌ترتیب کمیت‌های ورودی و خروجی می‌باشند. کمیت w میزان تاثیر p روی a را تعیین می‌کند (وزن ارتباطی لایه‌ها). از سوی دیگر نیز مقدار ثابت 1 در جمله b ضرب شده و با مقدار $w \times p$ جمع می‌گردد. حاصل این جمع ورودی خالص

1- Artificial Neural Network
2- Multi-Layer Perceptron

3- hyperbolic secant
4- Sigmoid
5- Hyperbolic Tangent
6-Root Mean Square Error
7- Mean Absolute Error
8- correlation

های هواشناسی برای شبیه‌سازی دقیق‌تر تبخیر از تشت هفت سناریو مختلف، متشکل از داده‌های متفاوت هواشناسی، تعریف گردید (جدول ۳). سپس این سناریوها در روش‌های برنامه‌ریزی ژنتیک، رگرسیون بردار پشتیبان و شبکه‌ی عصبی مصنوعی برای شبیه‌سازی تبخیر در دو ایستگاه سینوپتیک تبریز و جلفا به کار گرفته شد که نتایج کلی محاسبات صورت گرفته برای سناریوهای تعریف شده برای روش‌های مذکور در جدول‌های ۴ و ۵ ارایه شده‌است.

کشیدن هم‌زمان پارامترهای آماری می‌باشد. در دیاگرام مذکور، هر نقطه بیانگر عملکرد روش متناظر بوده و هرچه نقاط متناظر روش‌ها به نقطه داده‌های مشاهداتی در مختصات قطبی نزدیک‌تر باشد، نشان دهنده‌ی دقت بالاتر و خطای کم‌تر آن روش می‌باشد (Gleckler et al., 2008).

نتایج و بحث

به منظور بررسی امکان استفاده از ترکیب‌های متفاوتی از داده-

جدول ۳- پارامترهای دخیل در هر سناریو در روش‌های ANN، SVR و GP

شماره سناریو	پارامترهای ورودی
۱	دما و رطوبت نسبی
۲	دما و سرعت باد
۳	دما و تابش خورشیدی
۴	دما، رطوبت نسبی و سرعت باد
۵	دما، رطوبت نسبی و تابش خورشیدی
۶	دما، سرعت باد و تابش خورشیدی
۷	دما، رطوبت نسبی، سرعت باد و تابش خورشیدی

جدول ۴- نتایج کلی محاسبات صورت گرفته برای سناریوهای تعریف شده برای روش‌های ANN، SVR و GP در شهر تبریز

شماره سناریو	GP			SVR			ANN			N
	MAE	RMSE	R	MAE	RMSE	R	MAE	RMSE	R	
۱	۱/۵۶	۲/۲۶	۰/۹۱	۱/۵۱	۲/۳۱	۰/۹۰	۱/۵۳	۲/۲۹	۰/۹۱	۶
۲	۱/۵۶	۲/۲۳	۰/۹۱	۱/۳۹	۲/۲۱	۰/۹۱	۱/۴۱	۲/۱۴	۰/۹۲	۶
۳	۱/۵۹	۲/۳۲	۰/۹۱	۱/۴۸	۲/۲۹	۰/۹۱	۱/۴۸	۲/۲۴	۰/۹۱	۱۰
۴	۱/۵۹	۲/۲۶	۰/۹۱	۱/۳۹	۲/۱۹	۰/۹۱	۱/۴۴	۲/۱۷	۰/۹۲	۳
۵	۱/۵۴	۲/۲۷	۰/۹۱	۱/۴۹	۲/۲۹	۰/۹۱	۱/۵۳	۲/۲۸	۰/۹۱	۷
۶	۱/۴۹	۲/۱۸	۰/۹۲	۱/۴۱	۲/۲۴	۰/۹۱	۱/۴۱	۲/۱۷	۰/۹۲	۴
۷	۱/۵۳	۲/۲۳	۰/۹۱	۱/۴۲	۲/۲۲	۰/۹۱	۱/۴۵	۲/۲۱	۰/۹۲	۵

جدول ۵- نتایج کلی محاسبات صورت گرفته برای سناریوهای تعریف شده برای روش‌های ANN، SVR و GP در شهر جلفا

شماره سناریو	GP			SVR			ANN			N
	MAE	RMSE	R	MAE	RMSE	R	MAE	RMSE	R	
۱	۲/۴۷	۳/۱۵	۰/۸۳	۲/۰۵	۲/۶۶	۰/۸۷	۲/۱۲	۲/۶۷	۰/۸۷	۲
۲	۲/۱۲	۲/۷۴	۰/۸۷	۱/۹۵	۲/۴۷	۰/۹۰	۱/۸۸	۲/۳۹	۰/۹۱	۴
۳	۲/۴۴	۳/۱۳	۰/۸۳	۲/۰۰	۲/۶۷	۰/۸۷	۲/۱۰	۲/۷۰	۰/۸۷	۱۰
۴	۲/۰۹	۲/۷۱	۰/۸۸	۱/۷۴	۲/۲۷	۰/۹۲	۱/۸۰	۲/۳۹	۰/۹۲	۶
۵	۲/۴۵	۳/۱۳	۰/۸۳	۱/۹۱	۲/۵۲	۰/۸۹	۲/۰۳	۲/۵۸	۰/۸۸	۷
۶	۲/۰۸	۲/۷۴	۰/۸۸	۱/۷۷	۲/۳۲	۰/۹۱	۱/۸۵	۲/۳۸	۰/۹۱	۱
۷	۲/۰۵	۲/۶۸	۰/۸۸	۱/۷۱	۲/۲۲	۰/۹۲	۱/۷۲	۲/۲۱	۰/۹۲	۱۰

نتایج حاصل از روش شبکه عصبی مصنوعی

عملکرد را در بین سناریوهای مختلف SVR داشته است. هم‌چنین، در ایستگاه جلفا سناریوی هفتم در روش رگرسیون بردار پشتیبان (SVR-7) با داشتن جذر میانگین مربعات خطا $2/21$ ، میانگین خطای مطلق $1/71$ و ضریب همبستگی $0/92$ به دلیل داشتن کم‌ترین مقدار خطا به‌عنوان سناریو برتر SVR معرفی شد، اما همان‌طور که اشاره شد مدل SVR-7 به داده‌های هواشناسی زیادی نیازمند است تا بتواند رابطه‌ای بین میزان تبخیر از تشت و داده‌های هواشناسی با دقت مناسب حاصل کند. هم‌چون روش شبکه‌ی عصبی مصنوعی، در این روش نیز مدل SVR-4 با جذر میانگین مربعات خطا $2/27$ ، میانگین خطای مطلق $1/74$ و ضریب همبستگی $0/92$ ، پس از SVR-7 بهترین نتایج را حاصل نمود. اگرچه خطای مدل SVR-4 نسبت به SVR-7 کمی بیش‌تر است، اما به دلیل استفاده از داده‌های هواشناسی کم‌تر می‌توان آن را به‌عنوان بهترین سناریوی رگرسیون بردار پشتیبان معرفی نمود.

نتایج حاصل از روش برنامه‌ریزی ژنتیک

با بررسی‌های به‌عمل آمده در مورد روش برنامه‌ریزی ژنتیک در ایستگاه تبریز، سناریو ششم (GP-6) با داشتن ضریب همبستگی $0/92$ ، میانگین خطای مطلق $1/49$ و جذر میانگین مربعات خطای $2/18$ خطای کم‌تری نسبت به سایر سناریوها داشته و بهترین عملکرد را ارائه داده است. در این روش، مدل GP-7 نیز اگرچه به پارامترهای هواشناسی بیش‌تری نیاز دارد، ولی با داشتن ضریب همبستگی $0/91$ ، میانگین خطای مطلق $1/53$ و جذر میانگین مربعات خطای $2/23$ پس از سناریو ششم دقیق‌ترین شبیه‌سازی‌ها را از مقدار تبخیر داشته است. اما مدل‌های GP-4 و GP-5 با توجه به پارامترهای آماری کم‌ترین دقت را داشته و برای شبیه‌سازی تبخیر از تشت با روش GP پیشنهاد نمی‌شوند. هم‌چنین بر اساس نتایج حاصله در ایستگاه جلفا، مدل GP-7 با استفاده از داده‌های دما، رطوبت نسبی، تابش خورشیدی و سرعت باد و دارا بودن کم‌ترین خطا با جذر میانگین مربعات خطا $2/68$ ، میانگین خطای مطلق $2/05$ و ضریب همبستگی $0/88$ به‌عنوان دقیق‌ترین روش در بین سناریوهای روش برنامه‌ریزی ژنتیک انتخاب شد. پس از مدل GP-7، سناریو چهارم (GP-4) با جذر میانگین مربعات خطا $2/71$ ، میانگین خطای مطلق $1/09$ و ضریب همبستگی $0/88$ و با اندکی خطای بیش‌تر نسبت به مدل GP-7 در مرتبه دوم قرار گرفت.

مقایسه نتایج به‌دست آمده از روش‌های مورد مطالعه

به‌طور کلی، بهترین نتایج برای شهر تبریز به‌ترتیب از سناریو دوم در روش شبکه عصبی مصنوعی (ANN-2)، سناریو ششم در روش

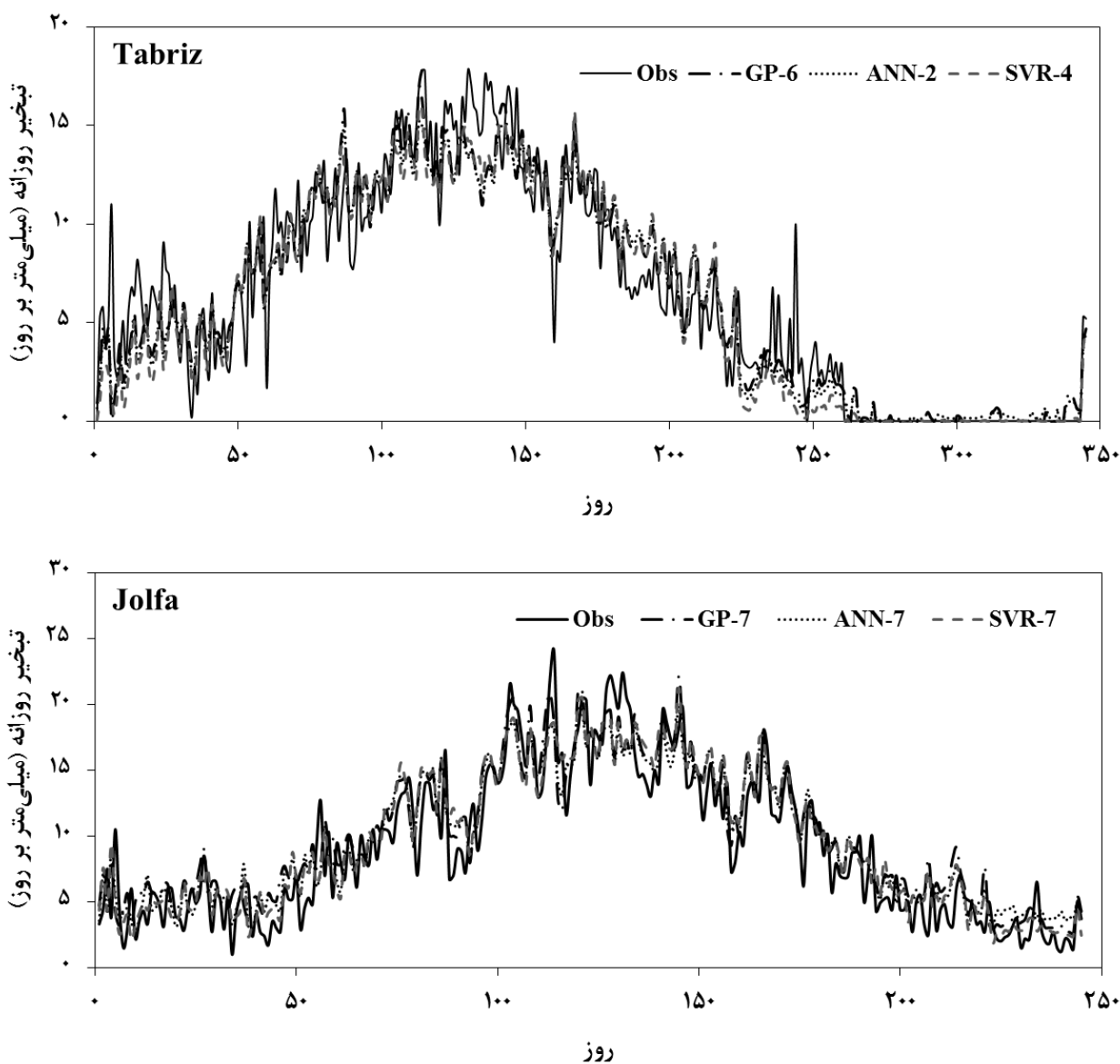
با توجه به نتایج به‌دست آمده در ایستگاه تبریز در روش شبکه‌ی عصبی مصنوعی، مدل ANN-2 با داشتن ضریب همبستگی $0/92$ ، میانگین خطای مطلق $1/41$ و جذر میانگین مربعات خطای $2/14$ عملکرد بهتری از خود نشان داد (جدول ۴). اگر چه این سناریو در مقایسه با مدل ANN-6 میانگین خطای مطلق بالاتری دارد، اما به دلیل کم‌بودن جذر میانگین مربعات خطا و هم‌چنین کم‌بودن پارامترهای ورودی می‌توان آن را به‌عنوان بهترین سناریو شبکه عصبی مصنوعی در شبیه‌سازی تبخیر در ایستگاه تبریز معرفی کرد. در مرتبه‌ی دوم نیز سناریو ششم (ANN-6) با داشتن ضریب همبستگی $0/92$ ، میانگین خطای مطلق $1/4$ و جذر میانگین مربعات خطای $2/17$ عملکرد قابل‌قبولی داشت. بر اساس نتایج بدست آمده در شهر جلفا سناریو هفتم در روش شبکه عصبی مصنوعی (ANN-7) با دارا بودن کم‌ترین جذر میانگین مربعات خطا $2/21$ ، کم‌ترین میانگین خطای مطلق $2/21$ و بیش‌ترین ضریب همبستگی $0/92$ بهترین نتیجه را در بین مدل‌های شبکه عصبی مصنوعی ارائه کرد. در رتبه‌ی بعدی، مدل ANN-4 با جذر میانگین مربعات خطا $2/39$ ، میانگین خطای مطلق $1/8$ و ضریب همبستگی $0/92$ شبیه‌سازی‌های به‌نسبت دقیقی را حاصل نمود (جدول ۵). گرچه مدل ANN-4 خطای بیش‌تری نسبت به مدل ANN-7 دارد، اما بدلیل نیاز کمی که به داده‌های هواشناسی دارد، می‌توان به‌عنوان روش بهینه در مواقع فقدان داده با قبول اندکی خطا مورد استفاده قرار گیرد.

نتایج حاصل از روش رگرسیون بردار پشتیبان

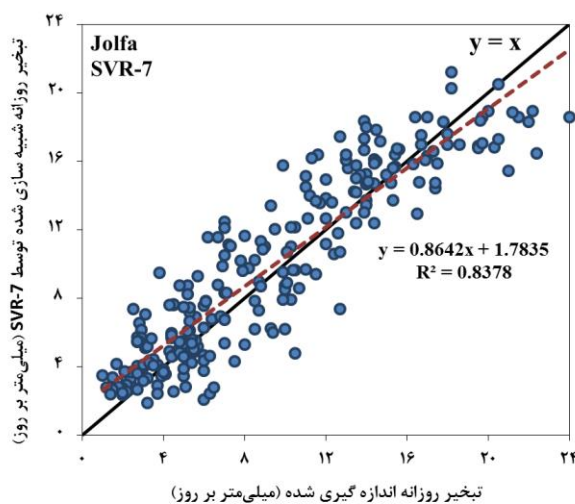
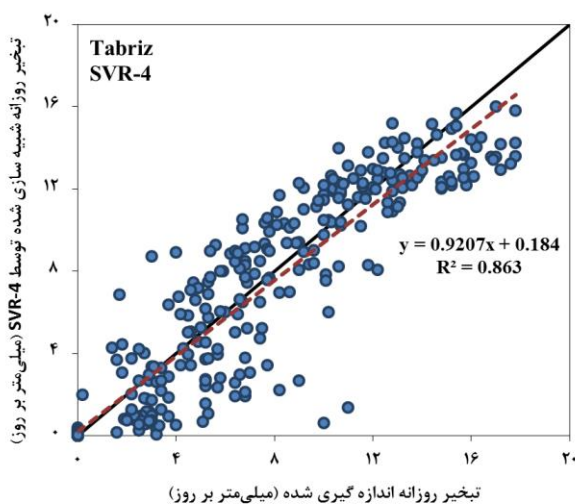
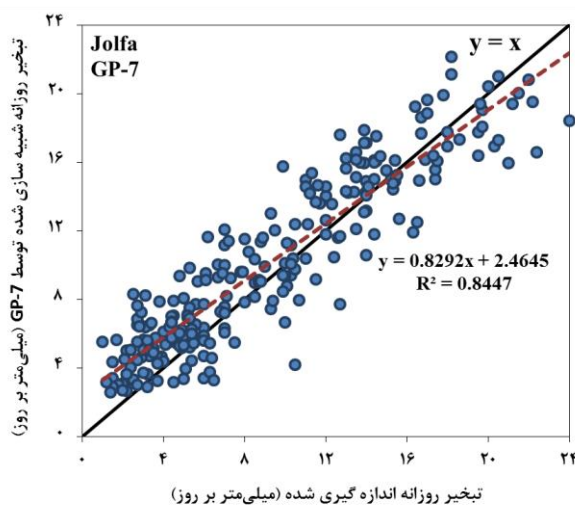
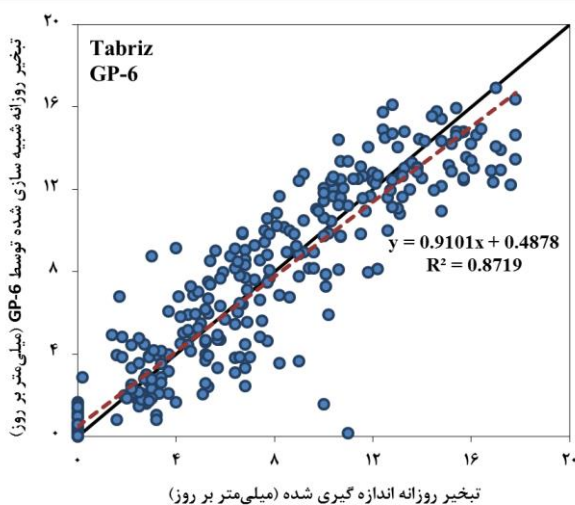
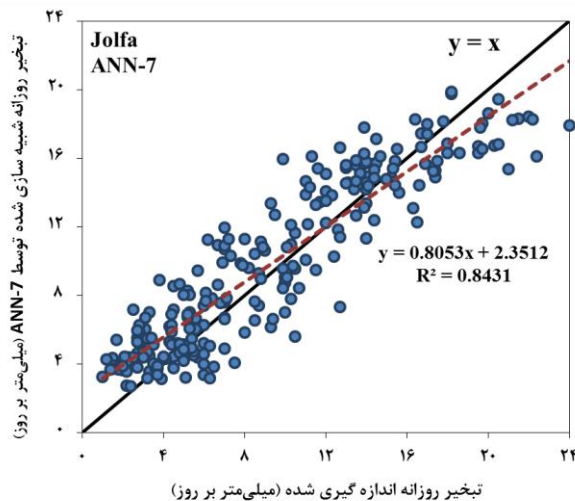
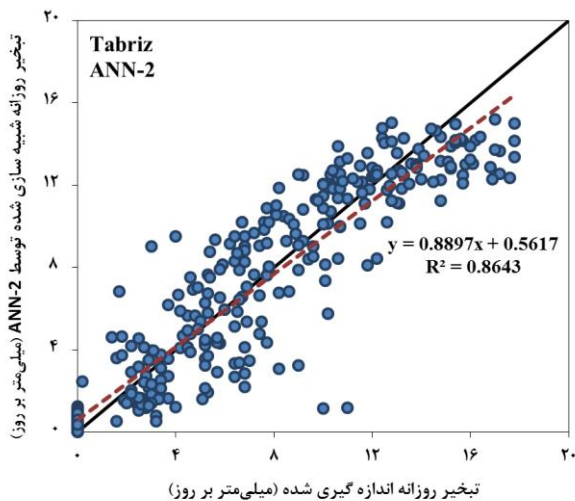
با توجه به محاسبات انجام گرفته تابع پیرسون در بین توابع کرنل در تمامی سناریوهای رگرسیون بردار پشتیبان بهترین نتایج را ارائه کرد. نتایج جدول ۴ در مورد ایستگاه تبریز نشان می‌دهد که در روش رگرسیون بردار پشتیبان، مدل SVR-4 با داشتن کم‌ترین خطا و دارا بودن پارامترهای آماری ضریب همبستگی $0/91$ ، میانگین خطای مطلق $1/39$ و جذر میانگین مربعات خطای $2/19$ بهترین عملکرد را در مقایسه با دیگر سناریوهای مختلف تعریف شده در مدل SVR داشته است. از طرفی، SVR-2 با توجه به این که تنها از دو پارامتر دما و سرعت باد استفاده می‌نماید، با داشتن ضریب همبستگی $0/91$ ، میانگین خطای مطلق $1/4$ و جذر میانگین مربعات خطای $2/21$ دقت به مراتب بیش‌تری نسبت به سایر سناریوهای رگرسیون بردار پشتیبان دارد. هم‌چنین به دلیل کم‌تر بودن پارامترهای مورد استفاده در این سناریو در مواقع کمبود داده با قبول مقداری خطا می‌توان از آن با اطمینان قابل‌قبولی استفاده کرد. مدل SVR-1 نیز با داشتن میانگین خطای مطلق $1/5$ و جذر میانگین مربعات خطای $2/31$ بدترین

هواشناسی ایستگاه‌های مورد مطالعه و در دوره زمانی خاصی بدست آمده و ممکن است نتایج حاصل شده در مناطق آب و هوایی مختلف دچار تغییراتی شود. بنابراین پیشنهاد می‌گردد پژوهشگران دقت روش‌های مورد مطالعه و دیگر روش‌های یادگیری ماشینی را برای شبیه‌سازی فرآیند تبخیر در پهنه‌های مختلف کشور مورد سنجش قرار دهند. نمودارهای تغییرات مقادیر تبخیر شبیه‌سازی شده برای سناریوهای برتر در هر سه روش (برنامه‌ریزی ژنتیک، رگرسیون بردار پشتیبان، شبکه عصبی مصنوعی) با مقادیر تبخیر مشاهداتی در مرحله‌ی صحت‌سنجی در شکل ۴ و نمودارهای پراکنش روش‌های مذکور در شکل ۵ نشان داده شده‌اند.

برنامه‌ریزی ژنتیک (GP-6) و سناریو چهارم در روش رگرسیون بردار پشتیبان (SVR-4) حاصل گردید. در بین روش‌های مورد مطالعه در این پژوهش برای شهر جلفا نیز بر اساس آن‌چه در جدول ۵ نیز موجود می‌باشد، با کم‌ترین میزان خطا به‌عنوان روش برتر انتخاب شد و مدل‌های SVR-7 و GP-7 در مرتبه‌های دوم و سوم قرار گرفتند که با قبول اندکی خطا می‌توان از این دو مدل نیز استفاده نمود. با نگاه کلی به نتایج و استفاده از تفاسیر فوق می‌توان چنین برداشت نمود که پارامترهای هواشناسی دما و سرعت باد مهم‌ترین نقش را در بالا بردن دقت شبیه‌سازی تبخیر از تشت دارا می‌باشند. لازم به ذکر است که نتایج به‌دست آمده بر اساس پارامترهای



شکل ۴- نمودارهای تغییرات زمانی مقادیر تبخیر با استفاده از بهترین مدل‌های مورد مطالعه



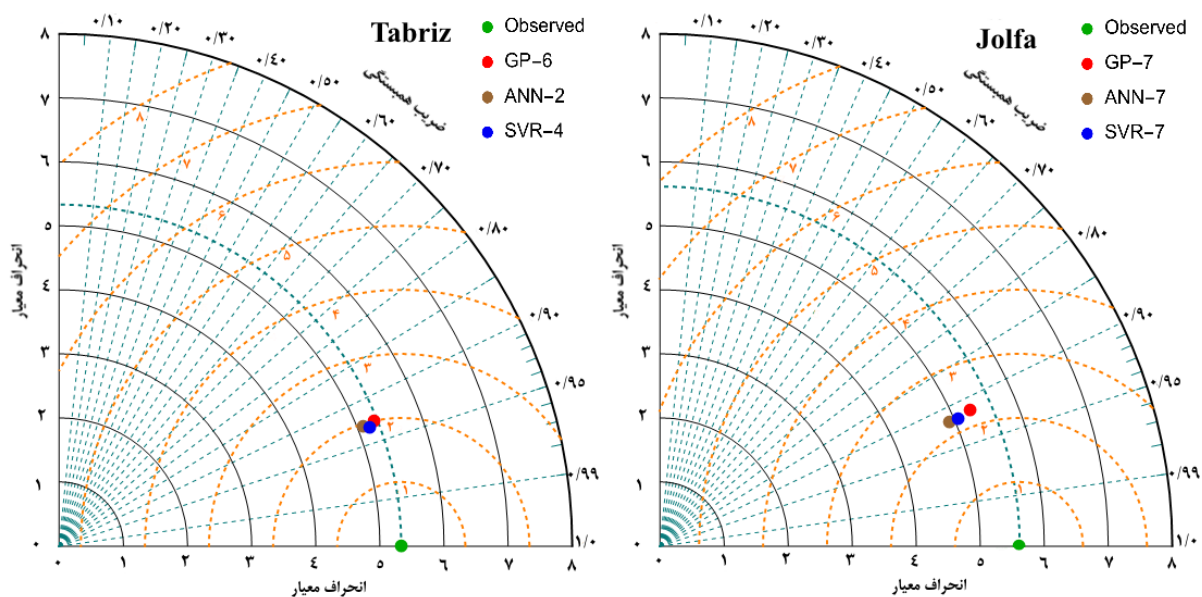
شکل ۵- نمودارهای پراکنش شبیه سازی تبخیر با استفاده از بهترین مدل های مورد مطالعه

۴ می توان چنین استنباط نمود که شبیه سازی های سناریو شماره ۲ برای شهر تبریز و شبیه سازی سناریو شماره ۷ در شهر جلفا در روش

روند اشاره شده در مورد بالابودن دقت سناریوهای برتر برای هر سه روش مورد مطالعه از شکل ۴ نیز قابل نتیجه گیری است. از شکل

بین داده‌های مشاهداتی و شبیه‌سازی شده توسط مدل‌های برتر مورد مطالعه رسم گردید (شکل ۶). لازم به ذکر است که در دیاگرام تیلور، فاصله طولی از مبدا مختصات نشان دهنده انحراف معیار، خط‌چین‌های شعاعی سبز رنگ نشان دهنده ضریب همبستگی و خط‌چین‌های قطاعی نارنجی رنگ نشان دهنده مقادیر جذر میانگین مربعات خطا می‌باشند که با بزرگ شدن قطاع دایره مقدار پارامتر مذکور بیش‌تر می‌شود. به عبارت دیگر، هر نقطه بر روی نمودار تیلور نشان دهنده هم‌زمان سه پارامتر انحراف معیار، ضریب همبستگی و جذر میانگین مربعات خطا می‌باشد.

ANN تطابق بیش‌تری با مقادیر مشاهداتی دارد. به عبارت دیگر، می‌توان مقدار تبخیر از تشت را در ایستگاه تبریز با داده‌های هواشناسی دما و سرعت باد، و در ایستگاه جلفا با داده‌های دما، سرعت باد، رطوبت و تابش خورشیدی با دقت قابل‌قبولی شبیه‌سازی نمود. هم‌چنین از شکل ۵ نیز مشاهده می‌گردد که پراکنش نقاط رسم شده حول محور نیم‌ساز برای مدل ANN-2 در شهر تبریز و ANN-7 در شهر جلفا کم‌تر بوده و مدل‌های مذکور دقت نسبتاً بالاتری از خود نشان داده‌اند. علاوه بر این، دیاگرام تیلور به منظور بررسی و تحلیل مقادیر انحراف معیار، ضریب همبستگی و جذر میانگین مربعات خطا



شکل ۶- دیاگرام تیلور روش‌های مورد مطالعه در شبیه‌سازی تبخیر و تعرق مرجع روزانه

پژوهش نشان دادند که در شهر تبریز سناریو دوم روش ANN و در شهر جلفا سناریو هفتم روش ANN کم‌ترین خطا را در مقایسه با سایر سناریوهای تعریف شده دارند و دقت به مراتب بالاتری را دارا می‌باشند. به عبارت دیگر می‌توان با استفاده از داده‌های هواشناسی دما و سرعت باد در ایستگاه تبریز و با داده‌های هواشناسی دما، سرعت باد، رطوبت نسبی و تابش خورشیدی در ایستگاه جلفا مقدار تبخیر را با دقت به نسبت بالایی شبیه‌سازی نمود. در نهایت پیشنهاد می‌گردد توانایی روش‌های مورد مطالعه در شبیه‌سازی تبخیر در اقلیم‌های مختلف توسط دیگر پژوهشگران مورد بحث و بررسی بیش‌تری قرار گیرد.

منابع

اسدپور اصل، ر.، ستاری، م.، ۱۳۹۴. پیش‌بینی عمق حفره آبستگي اطراف پایه پل با استفاده از روش رگرسیون ماشین بردار پشتیبان.

همان طوری که از شکل ۶ مشاهده می‌گردد، هر سه مدل ANN-2، GP-6 و SVR-4 تقریباً یکسانی در شبیه‌سازی مقدار تبخیر روزانه در ایستگاه تبریز دارند. هم‌چنین از دیاگرام تیلور مربوط به ایستگاه جلفا نیز می‌توان چنین نتیجه‌گیری نمود که مدل‌های ANN-7 و SVR-7 دقت به مراتب بالاتری نسبت به مدل GP-7 دارند.

نتیجه‌گیری

تبخیر به عنوان پارامتری اساسی در تحقیقات کشاورزی، هیدرولوژی، حفاظت آب و خاک از اهمیت خاصی برخوردار است. در این تحقیق، از روش‌های یادگیری ماشینی برنامه‌ریزی ژنتیک، رگرسیون بردار پشتیبان و شبکه عصبی مصنوعی برای شبیه‌سازی مقادیر روزانه تبخیر از تشت در دو ایستگاه هواشناسی تبریز و جلفا واقع در استان آذربایجان شرقی استفاده گردید. نتایج حاصل از این

- دهمین کنگره بین المللی مهندسی عمران. ایران.
- رحیمی خوب، ع. ۱۳۸۶. تدوین مدل شبکه عصبی مصنوعی تبخیر از تشت با حداقل داده هواشناسی. نهمین سمینار سراسری آبیاری و کاهش تبخیر، کرمان، ایران.
- زارع ابیانه، ح.، مقدم‌نیا، ع.، بیات ورکشی، م.، قاسمی، ع. ۱۳۸۹. تغییرات مکانی تبخیر از تشت و مقایسه آن با مدل‌های برآورد تبخیر در ایران. دانش آب و خاک. ۲۰: ۱-۱۱۳-۱۳۰.
- ستاری، م.ت.، اسمعیل‌زاده، ب. ۱۳۹۵. مقایسه نتایج مدل درختی M5 و برنامه‌ریزی ژنتیک با روش پنمن - مانتیث - فائو برای تخمین تبخیر - تعرق مرجع. مهندسی منابع آب. ۹: ۳۱-۱۱-۲۰.
- سلطانی، ع.، قربانی، م.ع.، فاخری‌فرد، ا.، دربندی، ص.، فرسادی‌زاده، د. ۱۳۸۹. برنامه‌ریزی ژنتیک و کاربرد آن در مدل‌سازی فرآیند بارش - رواناب. دانش آب و خاک جلد ۲۰: ۱-۶۱-۷۱.
- شاپان‌نژاد، م.، ساداتی‌نژاد، س.ج. ۱۳۸۷. تخمین تبخیر و تعرق پتانسیل روزانه گیاه مرجع چمن با استفاده از شبکه‌های عصبی مصنوعی. مرتع. ۲: ۱۸۲-۱۹۴.
- صمدیان‌فرد، س.، اسدی، ا. ۱۳۹۶. پیش‌بینی نمایه خشک‌سالی SPI به رگرسیون بردار پشتیبان و خطی چندگانه. حفاظت منابع آب و خاک. ۴: ۱-۱۶.
- غضنفری هاشمی، س.، اعتماد شهیدی، ا. ۱۳۹۱. پیش‌بینی عمق آبستگي اطراف پایه پل با استفاده از ماشین‌های بردار پشتیبان. عمران مدرس. ۲: ۲۳-۳۷.
- فلاح قالحری، غ.، شاکری، ف. ۱۳۹۴. کاربرد شبکه عصبی مصنوعی در پیش‌بینی بارش زمستانه. علوم و مهندسی آبخیزداری ایران. ۹: ۳۱-۹۹-۱۱۰.
- قره‌خانی، ا.، قهرمان، ن.، بذرافشان، ج. ۱۳۸۹. مطالعه روند تغییرات تبخیر از تشت در گستره اقلیمی ایران. پژوهش‌های آبخیزداری. ۹۸: ۸۵-۹۷.
- کارآموز، م.، عراقی‌نژاد، ش. ۱۳۹۳. هیدرولوژی پیشرفته. انتشارات دانشگاه صنعتی امیرکبیر. تهران. ایران.
- منه‌اج، م.ب. ۱۳۸۴. مبانی شبکه‌های عصبی (هوش محاسباتی). چاپ سوم. انتشارات دانشگاه صنعتی امیرکبیر.
- مهدوی، س.، رحیمی‌خوب، ع.، منتظر، ع.ا. ۱۳۹۰. مدل شبکه عصبی مصنوعی تبخیر ماهانه از تشت با استفاده از داده‌های هواشناسی - مطالعه موردی منطقه حاشیه دریای خزر. تحقیقات منابع آب ایران. ۲: ۷۱-۷۹.
- Boser, B.E., Guyon, I.M and Vapnik, V.N. 1992. A training algorithm for optimal margin classifiers. In D. Haussler, editor, 5th Annual ACM Workshop on COLT. Pittsburgh, PA. 144-152.
- Gleckler, P.J., Taylor, K.E and Doutriaux, C. 2008. Performance metrics for climate models. Journal of Geophysical Research: Atmospheres. 113.D6: 1-20.
- Koza, J.R. 1992. Genetic Programming: On the Programming of Computers by Means of Natural Selection. MIT Press. pp, 813.
- Taylor, K.E. 2001. Summarizing multiple aspects of model performance in a single diagram. Journal of Geophysical Research: Atmospheres. 106: 7183-7192.
- Terzi, O. 2010. Modeling of daily pan evaporation of lake Egirdir using data-driven techniques. International symposium on innovations in Intelligent systems and Applications. Istanbul. Turkey. 320-324.
- Vapnik, V. 1995. The Nature of Statistical Learning Theory. Springer, New York. pp, 314.
- Vapnik, V., Golowich, S., Smola, A. 1997. Support vector method for function approximation, regression estimation, and signal processing. In: Mozer, M., Jordan, M., Petsche, T. (Eds.), Advances in Neural Information Processing Systems. MIT Press, Cambridge. 9: 281-287.

Estimation of daily pan evaporation by using machine learning methods

S.Samadianfard^{1*}, S. Hashemi², M. Izadyar³

Received: Mar.14, 2018

Accepted: Jun.12, 2018

Abstract

Evaporation is one of the main processes in the hydrological cycle, and one of the most important factors in the related studies, namely agriculture, hydrology, aerology, exploitation of reservoirs, designing of irrigation and drainage systems, irrigation scheduling, and water resources management. Therefore, accurate estimation of evaporation rates has of the high importance in hydrology researches. In this regard, intelligent methods of Genetic Programming, Support Vector Regression, and Artificial Neural Networks have been used for evaluation and simulation of the pan evaporation rates over Tabriz and Jolfa synoptic stations, in the present research. For this purpose, meteorological data including evaporation, temperature, relative humidity, wind speed, and solar radiation during the period of 1993-2013 were used and the accuracy of studied methods were investigated by using statistical parameters of root mean square error, mean absolute error, correlation coefficient and Taylor diagram. The results showed that in the optimum case of Tabriz and Jolfa stations, genetic programming with error of 2.18 and 2.68, support vector regression with error of 2.19 and 2.22, and artificial neural network with error of 2.14 and 2.21, respectively, had reasonable performance in evaporation simulation. Conclusively, the second scenario of artificial neural network method through the input parameters of temperature and wind speed and the seventh scenario of artificial neural network method through the input parameters of temperature, humidity, wind speed, and solar radiation by having the best performance, were suggested as the accurate models with reasonable precision for simulating pan evaporation at Tabriz and Jolfa stations, respectively.

Keywords: Artificial neural network, Kernel functions, Statistical analysis, Water resources management.

1. Assistant Professor, Department of Water Engineering, Faculty of Agriculture, University of Tabriz, Tabriz, Iran
2. B.S student, Department of Water Engineering, Faculty of Agriculture, University of Tabriz, Tabriz, Iran
3. B.S student, Department of Water Engineering, Faculty of Agriculture, University of Tabriz, Tabriz, Iran
(*-Corresponding Author Email: s.samadian@tabrizu.ac.ir)