

ارزیابی عملکرد شبکه عصبی مصنوعی (ANN) و ماشین بردار پشتیبان (SVM) در تخمین مقادیر روزانه تبخیر (مطالعه موردی: ایستگاه‌های هواشناسی تبریز و مراغه)

محمد عیسی‌زاده* - دانشجوی دکتری، مهندسی منابع آب، دانشکده کشاورزی، دانشگاه تبریز
منیر شبرزاد - دانشجوی کارشناسی ارشد سنجش از دور و GIS، دانشکده جغرافیا و برنامه‌ریزی، دانشگاه تبریز
مجید رضایی بنفشه - دانشیار، گروه آب‌وهواشناسی، دانشکده جغرافیا و برنامه‌ریزی، دانشگاه تبریز

پذیرش مقاله: ۱۳۹۵/۱/۱۷ تأیید نهایی: ۱۳۹۵/۹/۳

چکیده

تبخیر مؤلفه‌ای اساسی در چرخه هیدرولوژی است و نقش مهمی در مدیریت منابع آب دارد. در این تحقیق عملکرد مدل‌های شبکه عصبی مصنوعی (ANN) و ماشین بردار پشتیبان (SVM) در تخمین تبخیر روزانه ارزیابی شده است. داده‌های روزانه هواشناسی میانگین دما، سرعت باد، فشار هوا، رطوبت نسبی، بارش، دمای نقطه شبنم، و ساعت آفتابی ایستگاه‌های سینوپتیک تبریز و مراغه، به منزله ورودی مدل‌های ANN و SVM، برای تخمین تبخیر روزانه استفاده شد. نخست ده ترکیب مختلف از هفت ورودی و سپس ورودی‌های منفرد به منظور تخمین تبخیر به کار گرفته شدند. نتایج مدل‌های استفاده‌شده نشان داد که هر دو مدل ANN و SVM عملکرد قابل قبولی در تخمین تبخیر دارند. ارزیابی نتایج استفاده از ورودی‌های تکی نشان داد که به ترتیب کاربرد پارامترهای میانگین دما و ساعت آفتابی - نسبت به پارامترهای دیگر - نتایج بهتری در تخمین تبخیر هر یک از ایستگاه‌ها داشته است. بررسی‌های این تحقیق نشان می‌دهد که اگرچه تفاوت معنی‌داری بین نتایج سه تابع کرنل ماشین بردار پشتیبان وجود ندارد، تابع کرنل پایه شعاعی در مقایسه با توابع کرنل دیگر از دقت زیاد و عملکرد بهتری در تخمین تبخیر روزانه برخوردار است.

کلیدواژه‌ها: تبریز، تخمین تبخیر، شبکه عصبی مصنوعی، ماشین بردار پشتیبان، مراغه.

مقدمه

برآورد صحیح تبخیر در برنامه‌ریزی‌های منابع آب ضروری است. از تشت تبخیر - به سبب سهولت تفسیر داده‌های آن در سراسر دنیا - همچون شاخصی برای تعیین تبخیر دریاچه‌ها و مخازن استفاده می‌شود (زارع ایبانه و همکاران، ۱۳۹۰). یکی از پیش‌نیازهای مدیریت بهینه آب در حوضه آبخیز تخمین دقیق مؤلفه‌های بیلان آب است. تبخیر - تعرق پتانسیل یکی از عوامل تأثیرگذار در بیلان آب به‌شمار می‌رود. برآورد دقیق تبخیر و تعرق در مطالعاتی از قبیل تغییر اقلیم جهانی، تکامل محیطی، و کنترل منابع آب نقش مهمی بازی می‌کند (لیو و همکاران، ۲۰۱۰). با توجه به میزان اندک نزولات جوی و محدودیت منابع آب در ایران، مدیریت صحیح منابع آب از اهمیت زیادی برخوردار است و کاربرد آن به دقت

فراوانی نیاز دارد. یکی از مسائل مهم در تبخیر^۱ ثابت مقدار تبخیر است؛ به طوری که مقدار اندازه‌گیری شده بیانگر کل تبخیر باشد. روش‌های مستقیم و غیرمستقیم فراوانی برای برآورد شدت تبخیر وجود دارد. روش تشت تبخیر جزو روش‌های مستقیم است که معمولاً از آن برای برآورد تبخیر دریاچه‌ها و مخازن استفاده می‌شود (کیم و همکاران، ۲۰۰۸). در زمینه برآورد مقدار تبخیر تا کنون پژوهش‌های بسیاری انجام گرفته است؛ به مواردی اشاره می‌شود:

بروتن و همکاران (۲۰۰۰) تبخیر روزانه از تشت تبخیر را با استفاده از شبکه عصبی مصنوعی، رگرسیون خطی، و روش پرستلی تیلور تخمین زدند. آن‌ها در تحقیق خود از داده‌های بارندگی، دما، رطوبت نسبی، تابش خورشیدی، و سرعت باد به عنوان ورودی استفاده کردند. نتایج مطالعه آن‌ها بیانگر آن است که میزان تبخیر برآوردشده با روش شبکه عصبی مصنوعی خطای کمتری نسبت به دو روش دیگر دارد. تیزی و ارول کسکین (۲۰۰۵)، با به‌کارگیری داده‌های ایستگاه هواشناسی نزدیک دریاچه‌ای در غرب ترکیه، میزان تبخیر روزانه از تشت را با استفاده از مدل شبکه عصبی مصنوعی برآورد کردند. نتایج تحقیق آن‌ها نشان داد بهترین ترکیب ورودی شامل داده‌های دمای هوا، دمای سطح آب، تابش خورشیدی، و فشار هواست. پیری و همکاران (۲۰۰۹) مدل شبکه عصبی مصنوعی را برای تخمین میزان تبخیر در مناطقی با اقلیم گرم و خشک (ایستگاه چاه نیمه زابل در ایران) به‌کار بردند. آن‌ها در تحقیقشان از داده‌های روزانه دمای هوا، سرعت باد، کمبود فشار، بخار اشباع، و رطوبت نسبی در بازه آماری ۱۹۹۵ - ۲۰۰۶ استفاده کردند. نتایج عملکرد شبکه عصبی مصنوعی در مقایسه با مدل‌های تجربی بسیار مطلوب‌تر است. مقدم‌نیا و همکاران (۲۰۰۹) بین تکنیک‌های شبکه عصبی مصنوعی و فازی - عصبی در برآورد تبخیر از تشت یک ایستگاه هواشناسی منطقه سیستم مقایسه‌ای انجام دادند. نتایج تحقیقات آن‌ها نشان‌دهنده کارایی خوب روش ANN در برآورد تبخیر از تشت است. کیسی (۲۰۰۹) عملکرد روش‌های رگرسیون چندمتغیره، ANN، و معادله تجربی استفان - استیوارت را برای مدل‌سازی تبخیر ماهانه سه ناحیه ایالات کالیفرنیا مقایسه کرد. نتایج ارزیابی روش‌های مختلف نشان داد که شبکه‌های عصبی RBF و MLP در مقایسه با روش‌های رگرسیون خطی چندمتغیره و معادله تجربی از کارایی و دقت بیشتری برخوردارند.

طبری و همکاران (۲۰۱۰)، در استان همدان، با استفاده از شبکه عصبی مصنوعی و رگرسیون غیرخطی، تبخیر روزانه از تشت را تخمین زدند. آن‌ها در مطالعه خود از داده‌های بارندگی، دمای هوا، رطوبت نسبی، تابش خورشیدی، و سرعت باد به منزله ورودی استفاده کردند. نتایج مطالعه آن‌ها بیانگر آن است که میزان تبخیر برآوردشده به روش شبکه عصبی مصنوعی خطای کمتری دارد. نورانی و سیاح‌فرد (۲۰۱۲)، با استفاده از سه روش شبکه عصبی پرسپترون چندلایه، شبکه عصبی پایه شعاعی^۱، و شبکه المان^۲ به تخمین میزان تبخیر روزانه ایستگاه‌های هواشناسی تبریز و ارومیه اقدام کردند. نتایج مطالعه آن‌ها نشان می‌دهد که مدل MLP در برآورد تبخیر بهتر از دو مدل دیگر عمل می‌کند. تزل و بویوکیلدیز (۲۰۱۵) قابلیت شبکه عصبی پرسپترون چندلایه، شبکه تابع پایه شعاعی، و رگرسیون بردار پشتیبان را در برآورد تبخیر ماهانه ایستگاه هواشناسی بیشه‌پیر بررسی کردند. دما، رطوبت نسبی، سرعت باد، و بارش به عنوان ورودی و تبخیر به عنوان خروجی در نظر گرفته شدند. هر دو مدل شبکه عصبی مصنوعی و رگرسیون بردار پشتیبان

1. radial basis neural network (RBNN)

2. elman network

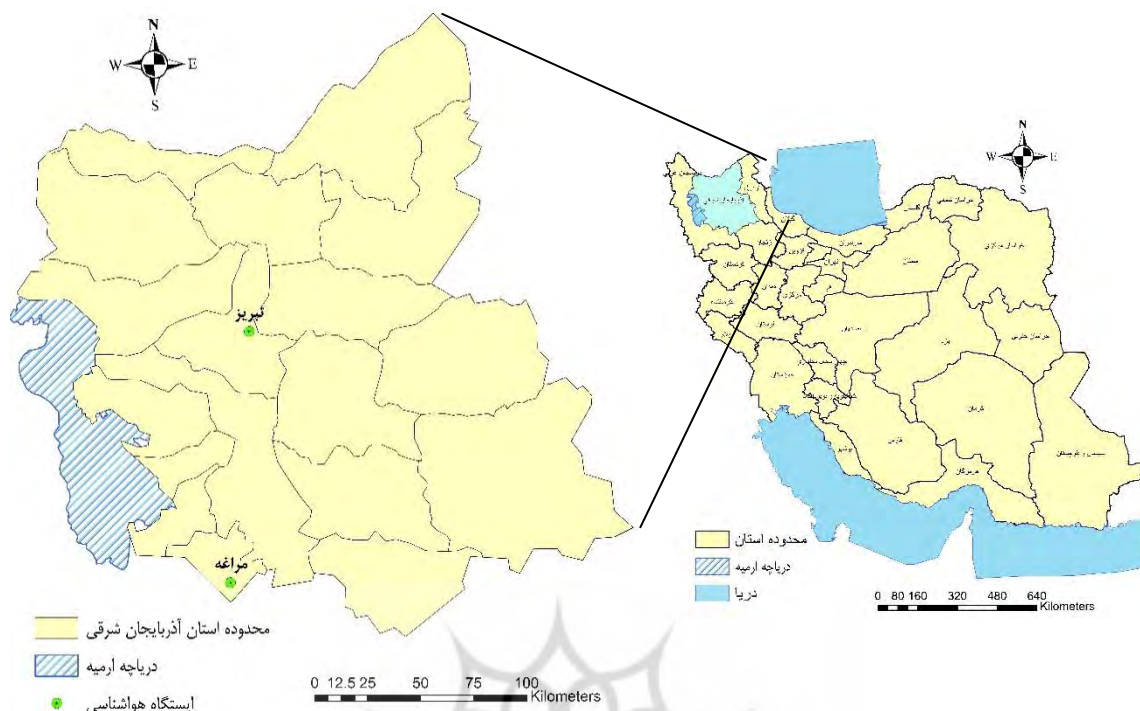
نتایج مشابهی به دست داده‌اند، ولی بهترین عملکرد متعلق به شبکه عصبی مصنوعی با $R^2=0.905$ است. کیسی (۲۰۱۶)، به منظور مدل‌سازی تبخیر ایستگاه مرسین و آنتالیا در کشور ترکیه، میزان صحت و دقت سه روش ماشین بردار پشتیبان، رگرسیون چندمتغیره تطبیقی، و مدل درختی را بررسی کرد. در مرحله اول میزان تبخیر با استفاده از داده‌های محلی برآورد شد، که نتایج حاکی از عملکرد بهتر مدل ماشین بردار پشتیبان نسبت به دو مدل دیگر است. در مرحله دوم میزان تبخیر بدون داده‌های محلی برآورد شد، که با توجه به نتایج مدل رگرسیون چندمتغیره تطبیقی، نسبت به دو مدل دیگر دقت بیشتری دارد. بامری (۱۳۹۳) با استفاده از دو مدل شبکه عصبی مصنوعی و رگرسیون خطی میزان تبخیر از تشت را در ایستگاه‌های زابل، زهک، زاهدان، خاش، ایرانشهر، چابهار، و سراوان به صورت روزانه برآورد کرد. در این مطالعه بهترین ترکیب برای ورودی شبکه برای هر هفت ایستگاه به طور یکسان میانگین دما، میانگین رطوبت نسبی، سرعت باد، و ساعت آفتابی انتخاب شد. نتایج برآورد تبخیر هر یک از ایستگاه‌ها حاکی از برتری مدل شبکه عصبی مصنوعی نسبت به مدل رگرسیون خطی است.

با توجه به پیشینه پژوهش، نگارندگان، در تحقیق حاضر، سعی می‌کنند تا کارایی مدل شبکه عصبی مصنوعی را در کنار مدل ماشین بردار پشتیبان برای تخمین مقدار تبخیر روزانه از تشت ایستگاه‌های سینوپتیک تبریز و مراغه ارزیابی کنند و دقت این مدل‌ها را در این تخمین نسبت به یکدیگر بررسی کنند. بدین منظور، عملکرد مدل‌های یادشده به ازای هفده ترکیب مختلف ورودی بررسی شده است. اولین ویژگی ترکیب‌های ورودی این تحقیق جامع بودن متغیرهای ورودی است و دومین ویژگی استفاده از ترکیب‌های تک‌متغیره است. استفاده از ترکیب‌های تک‌متغیره در ارزش‌گذاری هر یک از متغیرهای ورودی (در زمان فقدان اطلاعات کافی) بسیار حائز اهمیت است. همچنین، برای تخمین دقیق‌تر تبخیر، از مدل شبکه عصبی مصنوعی به ازای ۱ تا ۲۰ نورون در لایه پنهان و مدل ماشین بردار پشتیبان به ازای سه تابع کرنل مختلف استفاده شده است.

مواد و روش‌ها

مشخصات ایستگاه هواشناسی و داده‌های مورد استفاده

در این پژوهش از داده‌های ایستگاه‌های سینوپتیک تبریز و مراغه، که به ترتیب در ارتفاع‌های ۱۳۶۴ و ۱۳۴۴ متر از سطح دریا واقع‌اند، استفاده شد (شکل ۱). مقدار میانگین دمای هوا در ایستگاه‌ها به ترتیب برابر با ۱۲/۲۱ و ۱۲/۸۲ درجه سانتی‌گراد است و میزان بارش سالانه ایستگاه‌ها به ترتیب برابر ۲۸۵ و ۳۰۵ میلی‌متر است. آمار مورد استفاده در این تحقیق شامل تبخیر از تشت (E)، دمای میانگین روزانه (T)، میانگین فشار سطح ایستگاه (P)، میانگین سرعت باد (U)، میانگین رطوبت (H)، بارش (R)، دمای نقطه شبنم (Td)، و ساعت آفتابی (n) طی دوره آماری ۱۳۷۱ تا ۱۳۹۱ است. به سبب وجود داده‌های مفقودی و پرت زیاد در بعضی ماه‌ها، در این تحقیق فقط از داده‌های روزانه خرداد، تیر، مرداد، شهریور، و مهر استفاده شده است. ذکر این نکته لازم است که داده‌های مفقودی و پرت این ماه‌ها با روش رگرسیون خطی تکمیل شده‌اند. در این تحقیق از ده ترکیب مختلف ورودی (جدول ۱) برای ارزیابی مدل‌های شبکه عصبی مصنوعی و ماشین بردار پشتیبان استفاده شده است.



شکل ۱. منطقه و ایستگاه‌های هواشناسی مورد مطالعه

جدول ۱. ترکیب‌های مختلف ورودی، برای تخمین مقدار تبخیر روزانه از تشت

شماره ترکیب	متغیرهای ورودی	خروجی
۱	T, P, U, H, R, Td, n	E
۲	T, P, U, H, R, Td	E
۳	T, P, U, H, R, n	E
۴	T, P, U, H, Td, n	E
۵	T, P, U, R, Td, n	E
۶	T, P, H, R, Td, n	E
۷	T, U, H, R, Td, n	E
۸	P, U, H, R, Td, n	E
۹	T, P, U, H, Td	E
۱۰	P, U, H, R, n	E

با توجه به آمار موجود، از ۷۵ درصد از داده‌ها برای آموزش و از ۲۵ درصد دیگر برای صحت‌سنجی مدل‌های شبکه

عصبی مصنوعی و ماشین بردار پشتیبان استفاده شد.

ماشین بردار پشتیبان (SVM)

اولین کاربرد این روش در مسائل آب را دیباک و همکاران (۲۰۰۱) برای شبیه‌سازی بارش رواناب ارائه کردند. 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 (اسکندری و همکاران، ۱۳۹۰).

در مدل رگرسیون SVM 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^۱ (خطای مجاز) مشخص شود.

$$f(x) = W^T \cdot \varphi(x) + b \quad (۱)$$

$$y = f(x) + noise \quad (۲)$$

چنانچه W بردار ضرایب و b ثابت مشخصه‌های تابع رگرسیونی و φ نیز تابع کرنل باشد، آنگاه هدف پیدا کردن فرم تابعی برای $f(x)$ است. این مهم با واسنجی مدل SVM به وسیله مجموعه‌ای از نمونه‌ها (مجموعه واسنجی) محقق می‌شود.

این روند شامل بهینه‌سازی متوالی تابع خطاست. بسته به تعریف این تابع خطا، دو نوع مدل SVM تعریف می‌شود: SVM رگرسیونی نوع ۱ (به عنوان SVM- ϵ رگرسیونی نیز شناخته می‌شود) و SVM رگرسیونی نوع ۲ (به عنوان SVM- ν رگرسیونی شناخته می‌شود). شایان ذکر است در این تحقیق از مدل SVM- ϵ رگرسیونی، به سبب کاربرد گسترده آن در مطالعات رگرسیونی، برای پیش‌بینی جریان رودخانه استفاده شده است (اسکندری و همکاران، ۱۳۹۰). بنابراین، برای محاسبه W و b لازم است تابع خطا (رابطه ۳) در مدل SVM- ϵ با در نظر گرفتن شرایط مندرج (قیود) در رابطه ۴ بهینه شود.

$$\frac{1}{2} W^T w + C \sum_{i=1}^N \epsilon_i + C \sum_{i=1}^N \epsilon_i^* \quad (۳)$$

$$y_i - W^T \cdot \varphi(x_i) - b \leq \epsilon + \epsilon_i \quad W^T \cdot \varphi(x_i) + b - y_i \leq \epsilon + \epsilon_i^* \quad \epsilon_i, \epsilon_i^* \geq 0, i = 1, 2, \dots, N \quad (۴)$$

در معادلات بالا، C عددی صحیح و مثبت است که عامل تعیین جریمه در هنگام رخ دادن خطای واسنجی مدل است. φ تابع کرنل^۲، N تعداد نمونه‌ها، و دو مشخصه^۳ ϵ_i و ϵ_i^* متغیرهای کمبود^۳ هستند که حد بالا و پایین خطای آموزش مرتبط با مقدار خطای مجاز را مشخص می‌کنند (اسکندری و همکاران، ۱۳۹۰).

در مسائل پیش‌بینی می‌شود که داده‌ها درون بازه مرزی قرار گیرند. حال اگر داده‌ای خارج از بازه قرار گرفت، آنگاه یک خطا معادل ϵ_i و ϵ_i^* وجود خواهد داشت. ذکر این نکته نیز لازم است که مدل SVM مشکلات ناشی از

1. noise
2. kernel function
3. slack variable

کم‌تخمینی^۱ و فوق برآزشی را با کمینه‌کردن هم‌زمان دو ترم $WT.W/2$ و خطای آموزشی، یعنی $\sum_{i=1}^N (\xi_i + \xi_i^*)$ ، را در

رابطه ۳ حل می‌کند. بنابراین، با معرفی دو ضریب لاگرانژ α_i و α_i^* مسئله بهینه‌سازی با حداکثرسازی عددی تابع درجه دوم زیر (رابطه ۵) با شرایط رابطه ۶ حل خواهد شد.

$$\sum_{i=1}^N y_i (\alpha_i - \alpha_i^*) - \varepsilon \sum_{i=1}^N (\alpha_i + \alpha_i^*) - \frac{\Delta}{2} \sum_{i,j=1}^N (\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*) \phi(x_i)^T \cdot \phi(x_j) \quad (5)$$

$$\sum_{i=1}^N (\alpha_i + \alpha_i^*) = \dots \leq \alpha_i \leq C \quad \dots \leq \alpha_i^* \leq C \quad i = 1, 2, \dots, N \quad (6)$$

تابع هدف رابطه ۵ تابع محدب است، بنابراین، جواب رابطه ۵ یکتا و بهینه خواهد بود. پس از تعریف ضرایب لاگرانژ در رابطه ۳، مشخصه‌های w و b در مدل SVM رگرسیونی با استفاده از شرایط تئوری کراش-کوهن-تاکر^۲ محاسبه

می‌شود (فلتچر، ۱۹۸۷) که در آن $W = \sum_{i=1}^N (\alpha_i + \alpha_i^*) \phi(x_i)$ است.

در نتیجه برای مدل SVM رگرسیونی خواهیم داشت:

$$W = \sum_{i=1}^N (\alpha_i + \alpha_i^*) \phi(x_i)^T \cdot \phi(x) + b \quad (7)$$

باید توجه کرد که ترم‌های لاگرانژ $(\alpha_i + \alpha_i^*)$ می‌تواند صفر و یا غیر صفر باشد. بنابراین، فقط مجموعه داده‌هایی که ضرایب $\overline{\alpha_i}$ آن‌ها غیر صفر است در معادله رگرسیون نهایی وارد می‌شوند؛ این مجموعه داده‌ها به عنوان بردارهای پشتیبان شناخته می‌شوند. به طور ساده، بردارهای پشتیبان آن داده‌هایی هستند که به ساختارشدن تابع رگرسیونی کمک می‌کنند. در میان بردارهای یادشده آن‌هایی که مقدار $|\overline{\alpha_i}|$ آن‌ها کمتر از C باشد بردارهای پشتیبان حاشیه‌ای^۳ نامیده می‌شوند. هنگامی که مقدار $|\overline{\alpha_i}|$ بردارهای پشتیبان برابر مقدار C باشد، به عنوان بردار پشتیبان خطا^۴ یا بردار پشتیبان کراندار شناخته می‌شود. بردارهای پشتیبان حاشیه‌ای در حاشیه مرز غیرحساس یافت می‌شوند؛ در حالی که بردارهای پشتیبان خطا خارج از بازه‌اند. در نهایت، تابع SVM رگرسیونی را می‌توان به شکل رابطه ۸ بازنویسی کرد:

$$f(x) = \sum_{i=1}^N \overline{\alpha_i} \phi(x_i)^T \cdot \phi(x) + b \quad (8)$$

در رابطه ۸، $\overline{\alpha_i}$ میانگین ضرایب لاگرانژ است. محاسبه $\phi(x)$ در فضای مشخصه آن ممکن است بسیار پیچیده باشد. برای حل این مشکل روند معمول در مدل SVM رگرسیونی انتخاب یک تابع کرنل است (اسکندری و همکاران، ۱۳۹۰). از ماشین بردار پشتیبان معمولاً با سه تابع کرنل پایه شعاعی (RBF)^۵، چندجمله‌ای درجه d ، و خطی استفاده می‌شود که روابط مورد استفاده در هر یک از آن‌ها به ترتیب در زیر آورده شده است (کاوزوگلو و کولکسین، ۲۰۰۹؛ بافینگ و همکاران، ۲۰۰۸).

1. under fitting
2. Karush- Kuhn-Tucker
3. margin support vector
4. error support vector
5. radial basis function

$$k(x, x_i) = \exp\left(-\|x - x_i\|^2 / \sigma^2\right) \quad (9)$$

$$k(x, x_i) = (t + \langle x, x_i \rangle)^d \quad (10)$$

$$k(x, x_i) = x \cdot x_i \quad (11)$$

با توجه به اینکه از پُرکاربردترین توابع کرنل، کرنل خطی، پایه شعاعی، و چندجمله‌ای است (باساک و همکاران، ۲۰۰۷؛ لیو، ۲۰۱۱؛ وپنیک و کروونیکیس، ۱۹۹۱)، در این تحقیق نیز از این سه تابع کرنل استفاده شده است. ذکر این نکته لازم است که فرایند محاسبات ماشین بردار پشتیبان بر اساس کدنویسی در محیط متلب انجام شد و پارامترهای توابع کرنل از طریق سعی و خطا بهینه گردید.

شبکه عصبی مصنوعی با آموزش لونبرگ - مارکوارت

ANN یک سیستم پردازشگر توزیعی اطلاعات است که خصوصیات عملکرد آن شبیه به ساختار عصبی مغز انسان است و گسترش آن بر اساس قواعد زیر است:

- پردازش اطلاعات در عناصر منفردی به نام گره انجام می‌شود و این گره‌ها در دسته‌هایی به نام «لایه» مرتب می‌شوند؛

- سیگنال‌ها بین گره‌ها از طریق اتصالات منتقل می‌شوند؛

- هر اتصال وزنی دارد که نشان‌دهنده نیروی اتصال آن است؛

- هر گره مجموعه وزنی سیگنال‌ها را با به‌کاربردن یک تبدیل غیرخطی- که تابع محرک^۱ نامیده می‌شود- به سیگنال خروجی تبدیل می‌کند (زیلاند و همکاران، ۱۹۹۹؛ ASCE، ۲۰۰۰).

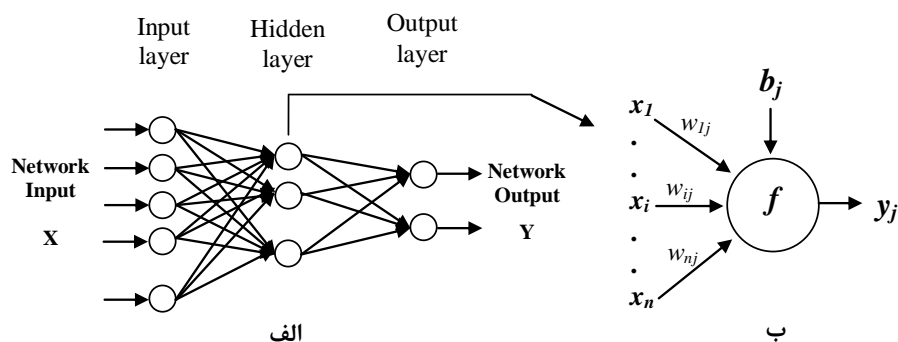
می‌توان گفت که حدود ۹۰ درصد از مدل‌های شبکه عصبی- که در مسائل هیدرولوژی به کار گرفته شده‌اند- از نوع شبکه‌های عصبی پیشرو^۲ هستند (نجفی و همکاران، ۱۳۹۴) و مدل MLP با الگوریتم آموزش لونبرگ- مارکوارت (LM)^۳ از مهم‌ترین آن‌ها به‌شمار می‌رود. معماری معمول این شبکه متشکل از سه بخش لایه ورودی، پنهان، و خروجی است. تعداد نورون‌های موجود در لایه ورودی و خروجی بستگی به نوع مسئله دارد، ولی تعداد گره‌های لایه پنهان با سعی و خطا به‌دست می‌آید.

در شبکه‌های پیشرو، اتصالات میان گره‌هایی وجود دارد که در لایه‌های متفاوت قرار دارند. همچنین، ورودی‌ها در لایه ورودی شبکه نمایش داده می‌شوند و عمل تحریک از ورودی به سمت خروجی شبکه شروع می‌شود (علیزاده، ۱۳۹۰). شکل ۲ ساختار کلی شبکه عصبی مصنوعی را نشان می‌دهد.

1. activation function

2. feed forward neural networks

3. Levenberg-Marquardt



شکل ۲. الف) شمایی کلی از شبکه‌های عصبی پیشرو سه لایه؛ ب) ساختار گره زام از این شبکه (ASCE، ۲۰۰۰)

در شکل ۲ ب، $X = (x_1, x_2, \dots, x_n)$ بردار ورودی‌ها، w_{ij} وزن اتصالی از آامین گره به آامین گره در لایه بعد، b_j بایاس مربوط به گره زام، y_j ها خروجی‌های هر لایه، و Y خروجی نهایی ANN است. همچنین، f تابع محرک هر نورون است که اطلاعات ورودی از نورون‌های لایه قبلی را به نورون‌های لایه بعدی نگاشت می‌کند. در شبکه‌های پیشرو، اتصالات میان گره‌هایی وجود دارد که در لایه‌های متفاوت قرار دارند. همچنین، ورودی‌ها در لایه ورودی شبکه نمایش داده می‌شوند و عمل تحریک از ورودی به سمت خروجی شبکه شروع می‌شود (ASCE، ۲۰۰۰).

ذکر این نکته لازم است که در تحقیق حاضر از تعداد نورون‌های ۱ تا ۲۰ برای لایه پنهان شبکه استفاده شد تا از این طریق بتوان تأثیر ساختارهای مختلف شبکه بر عملکرد آن در شبیه‌سازی تبخیر را ارزیابی کرد. از آنجا که متداول‌ترین نوع توابع محرک تابع تانژانت سیگموئید است، در این مطالعه از آن برای نگاشت اطلاعات از لایه ورودی به لایه پنهان و از تابع محرک خطی برای نگاشت اطلاعات از لایه پنهان به لایه خروجی استفاده شد. همچنین، به علت تفاوت در واحد داده‌های ورودی، کلیه داده‌های ورودی قبل از مدل‌سازی استاندارد شد. سپس، در فرایند محاسبات وارد شدند. ذکر این نکته لازم است که فرایند محاسبات شبکه عصبی مصنوعی بر اساس کدنویسی در محیط متلب انجام گرفته است.

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 کاربرد در علوم مهندسی آب ضرایب همبستگی (CC)^۱ و ریشه میانگین مربعات خطا (RMSE)^۲ است، که به ترتیب در رابطه‌های ۱۲ و ۱۳ آورده شده‌اند. دقیق‌ترین مدل، با توجه به این معیارها، مدلی خواهد بود که مقدار این سه معیار برای آن به ترتیب نزدیک به یک و صفر باشد.

$$CC = \frac{\sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^N (x_i - \bar{x})^2 \sum_{i=1}^N (y_i - \bar{y})^2}} \quad (12)$$

1. correlation coefficient

2. root mean square error

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - y_i)^2} \quad (13)$$

در رابطه‌های ۱۲ و ۱۳ x_i و \bar{x} به ترتیب مقادیر مشاهداتی تبخیر و میانگین مقادیر مشاهداتی، y_i مقادیر محاسباتی تبخیر، \bar{y} میانگین مقادیر محاسباتی، و N تعداد داده‌هاست (میسرا و همکاران، ۲۰۰۹).

یافته‌های پژوهش

پس از انجام دادن محاسبات تخمین تبخیر از تست با استفاده از مدل‌های شبکه عصبی مصنوعی و ماشین بردار پشتیبان به ازای ده ترکیب ذکر شده در جدول ۱، دو آماره CC و $RMSE$ برای هر یک از ترکیب‌ها محاسبه شد. نتایج مربوط به هر یک از مدل‌ها در زیر نشان داده شده است:

مدل شبکه عصبی مصنوعی

جدول ۲ نتایج به دست آمده از شبیه‌سازی‌های شبکه عصبی مصنوعی به ازای ده ترکیب ارائه شده (جدول ۱) را نشان می‌دهد.

جدول ۲. نتایج تخمین تبخیر با استفاده از شبکه عصبی مصنوعی به ازای ده ترکیب مختلف ورودی

ایستگاه	شماره ترکیب	تعداد نورون بهینه	آموزش		صحت‌سنجی	
			CC	RMSE(mm)	CC	RMSE(mm)
تبریز	۱	۸	۰/۷۹	۲/۱۸	۰/۷۸	۲/۱۲
تبریز	۲	۱۰	۰/۷۸	۲/۲۲	۰/۷۶	۲/۲۲
تبریز	۳	۱۳	۰/۷۸	۲/۲۱	۰/۷۷	۲/۲۱
تبریز	۴	۱۳	۰/۷۹	۲/۱۷	۰/۷۷	۲/۱۷
تبریز	۵	۶	۰/۷۸	۲/۳۱	۰/۷۷	۲/۳۱
تبریز	۶	۹	۰/۷۶	۲/۳۱	۰/۷۴	۲/۳۱
تبریز	۷	۳	۰/۷۸	۲/۲۳	۰/۷۷	۲/۲۳
تبریز	۸	۳	۰/۷۸	۲/۲۱	۰/۷۸	۲/۲۱
تبریز	۹	۶	۰/۷۸	۲/۲۳	۰/۷۶	۲/۲۳
تبریز	۱۰	۱۵	۰/۷۳	۲/۴۳	۰/۷۲	۲/۴۳
مراغه	۱	۱۱	۰/۶۵	۱/۳۹	۰/۶۰	۱/۶۲
مراغه	۲	۱۳	۰/۶۲	۱/۴۴	۰/۵۸	۱/۶۴
مراغه	۳	۵	۰/۶۲	۱/۴۴	۰/۵۹	۱/۶۴

ادامه جدول ۲. نتایج تخمین تبخیر با استفاده از شبکه عصبی مصنوعی به ازای ده ترکیب مختلف ورودی

ایستگاه	شماره ترکیب	تعداد نرون پهنه	آموزش		صحت‌سنجی	
			CC	RMSE(mm)	CC	RMSE(mm)
مراغه	۴	۱۷	۰/۶۵	۱/۴۰	۰/۵۹	۱/۶۳
مراغه	۵	۵	۰/۶۳	۱/۴۳	۰/۵۹	۱/۶۴
مراغه	۶	۵	۰/۶۳	۱/۴۳	۰/۶۰	۱/۶۱
مراغه	۷	۱۰	۰/۶۲	۱/۴۴	۰/۵۹	۱/۶۳
مراغه	۸	۸	۰/۶۲	۱/۵۶	۰/۵۸	۱/۶۵
مراغه	۹	۱۱	۰/۶۰	۱/۴۶	۰/۵۸	۱/۶۵
مراغه	۱۰	۷	۰/۵۹	۱/۴۸	۰/۵۷	۱/۶۶

با توجه به جدول ۲، در حالت کلی مدل شماره یک- که شامل هر هفت ورودی در نظر گرفته شده است- دارای دقت بیشتری در تخمین مقادیر تبخیر از تشت ایستگاه تبریز و مراغه است. بر اساس نتایج ترکیب‌های ۱ تا ۸، می‌توان نتیجه گرفت که حذف یکی از متغیرهای ترکیب ۱ تأثیر چندانی در نتایج خروجی ندارد. در واقع، شش پارامتر ورودی دیگر تأثیر پارامتر حذف‌شده را جبران می‌کنند. نتایج ترکیب‌های نه و ده نشان می‌دهد که میانگین دما و دمای نقطه شبنم نسبت به بارش و ساعت آفتابی تأثیر بیشتری در تخمین تبخیر از تشت دارند. با توجه به اینکه میانگین تبخیر از تشت ایستگاه‌های تبریز و مراغه در ماه‌های مورد بررسی تقریباً برابر با ۱۰/۶۴ و ۱۱/۱۷ است، خطای نزدیک به دو میلی‌متر حاکی از دقت خوب مدل شبکه عصبی مصنوعی در تخمین تبخیر از تشت است.

در مرحله بعدی تبخیر از تشت به عنوان خروجی و هر یک از هفت متغیر- میانگین دما، میانگین فشار سطح ایستگاه، میانگین سرعت باد، میانگین رطوبت، بارش، دمای نقطه شبنم، و ساعت آفتابی- به صورت منفرد به عنوان ورودی در نظر گرفته شدند. جدول ۳ نتایج حاصل از تخمین تبخیر از تشت در دوره آموزش و صحت‌سنجی به ازای هر یک از هفت متغیر ورودی را نشان می‌دهد.

نتایج جدول ۳ اولویت استفاده از تک‌متغیرهای ورودی برای تخمین مقدار تبخیر هر یک از ایستگاه‌های تبریز و مراغه را نشان می‌دهد. بنابراین، استفاده از متغیرهای هواشناسی میانگین دمای روزانه، میانگین فشار سطح ایستگاه، ساعت آفتابی، سرعت باد، میانگین رطوبت، دمای نقطه شبنم، و بارش به ترتیب اولویت بیشتری نسبت به یکدیگر برای تخمین مقدار تبخیر روزانه از تشت ایستگاه تبریز دارند. در ایستگاه مراغه نیز این اولویت‌ها به ترتیب شامل میانگین دمای روزانه، ساعت آفتابی، میانگین رطوبت، بارش، میانگین فشار سطح ایستگاه، دمای نقطه شبنم، و سرعت باد است.

جدول ۳. نتایج تخمین تبخیر با استفاده از شبکه عصبی مصنوعی به ازای هر یک از متغیرهای ورودی

ایستگاه	متغیر ورودی	تعداد نورون بهینه	آموزش		صحت سنجی	
			CC	RMSE(mm)	CC	RMSE(mm)
تبریز	میانگین دما	۴	۰/۷۲	۲/۴۶	۰/۷۰	۲/۴۲
تبریز	میانگین فشار سطح ایستگاه	۱	۰/۵۱	۳/۰۵	۰/۵۳	۲/۸۶
تبریز	میانگین سرعت باد	۱	۰/۴۶	۳/۱۵	۰/۴۴	۳/۰۳
تبریز	میانگین رطوبت	۲	۰/۴۴	۳/۱۸	۰/۳۹	۳/۱۱
تبریز	بارش	۲	۰/۱۷	۳/۵۰	۰/۰۹	۳/۳۷
تبریز	دمای نقطه شبنم	۵	۰/۲۷	۳/۴۲	۰/۳۲	۳/۲۰
تبریز	ساعت آفتابی	۱۸	۰/۴۹	۳/۰۹	۰/۴۵	۳/۰۲
مراغه	میانگین دما	۴	۰/۵۲	۱/۵۶	۰/۵۰	۱/۷۵
مراغه	میانگین فشار سطح ایستگاه	۶	۰/۳۸	۱/۷۰	۰/۲۶	۱/۹۶
مراغه	میانگین سرعت باد	۸	۰/۱۴	۱/۸۲	۰/۰۹	۲/۰۱
مراغه	میانگین رطوبت	۱۵	۰/۴۴	۱/۶۵	۰/۴۴	۱/۸۲
مراغه	بارش	۴	۰/۳۱	۱/۷۴	۰/۳۳	۱/۹۱
مراغه	دمای نقطه شبنم	۲	۰/۱۰	۱/۸۲	۰/۰۹	۲/۰۱
مراغه	ساعت آفتابی	۱۰	۰/۴۹	۱/۶۰	۰/۴۵	۱/۸۱

نورانی و سیاح فرد (۲۰۱۲)، بعد از بررسی حساسیت هر یک از ورودی‌های به کاررفته در تخمین تبخیر روزانه از تشت ایستگاه‌های هواشناسی تبریز و ارومیه، میانگین دمای هوا را مؤثرترین ورودی معرفی کردند که با نتایج این تحقیق تطابق کامل دارد. بنابراین، در صورت کمبود داده در ایستگاه هواشناسی، می‌توان به ترتیب از سایر پارامترهای اشاره شده برای تخمین تبخیر استفاده کرد. مقایسه نتایج جدول‌های ۲ و ۳ نشان می‌دهد که میزان تبخیر را می‌توان با استفاده از یک ورودی نیز بسیار خوب تخمین زد و اضافه کردن ورودی‌ها تا مقدار تقریباً مشخصی باعث بهبود نتایج می‌شود. یکی از مهم‌ترین دلایل این موضوع وجود همبستگی بین ورودی‌ها با تبخیر از تشت ایستگاه‌هاست.

مدل ماشین بردار پشتیبان

پس از تخمین تبخیر از تشت به ازای ترکیب‌های ورودی جدول ۱، خلاصه نتایج در جدول ۴ آورده شده است. ذکر این نکته لازم است که، برای تخمین تبخیر از تشت، هر یک از سه تابع کرنل پایه شعاعی، چندجمله‌ای، و خطی به ازای کلیه ترکیب‌ها ارزیابی شده است و در جدول نتایج فقط تابع کرنلی که بهترین نتیجه را به دست داده ذکر شده است.

جدول ۴. نتایج تخمین تبخیر با استفاده از ماشین بردار پشتیبان به ازای ده ترکیب مختلف ورودی

ایستگاه	شماره ترکیب	تابع کرنل منتخب	آموزش		صحت‌سنجی	
			CC	RMSE(mm)	CC	RMSE(mm)
تبریز	۱	پایه شعاعی	۰٫۸۰	۲٫۱۰	۰٫۷۶	۲٫۳۲
تبریز	۲	پایه شعاعی	۰٫۷۸	۲٫۱۸	۰٫۷۸	۲٫۲۷
تبریز	۳	چندجمله‌ای	۰٫۷۹	۲٫۱۸	۰٫۷۷	۲٫۱۸
تبریز	۴	پایه شعاعی	۰٫۷۹	۲٫۱۵	۰٫۷۷	۲٫۲۰
تبریز	۵	چندجمله‌ای	۰٫۷۹	۲٫۱۵	۰٫۷۷	۲٫۲۷
تبریز	۶	پایه شعاعی	۰٫۷۷	۲٫۲۳	۰٫۷۳	۲٫۳۹
تبریز	۷	چندجمله‌ای	۰٫۷۹	۲٫۱۷	۰٫۷۷	۲٫۲۳
تبریز	۸	چندجمله‌ای	۰٫۷۹	۲٫۱۶	۰٫۷۸	۲٫۱۷
تبریز	۹	چندجمله‌ای	۰٫۷۷	۲٫۲۱	۰٫۷۹	۲٫۱۹
تبریز	۱۰	پایه شعاعی	۰٫۷۳	۲٫۴۱	۰٫۷۲	۲٫۴۲
مراغه	۱	پایه شعاعی	۰٫۶۶	۱٫۳۸	۰٫۶۲	۱٫۵۹
مراغه	۲	پایه شعاعی	۰٫۶۴	۱٫۴۶	۰٫۵۶	۱٫۵۰
مراغه	۳	پایه شعاعی	۰٫۶۷	۱٫۳۸	۰٫۵۶	۱٫۶۰
مراغه	۴	پایه شعاعی	۰٫۶۷	۱٫۴۲	۰٫۶۱	۱٫۴۴
مراغه	۵	پایه شعاعی	۰٫۶۷	۱٫۳۹	۰٫۶۲	۱٫۴۹
مراغه	۶	پایه شعاعی	۰٫۶۵	۱٫۴۴	۰٫۶۲	۱٫۴۸
مراغه	۷	پایه شعاعی	۰٫۶۴	۱٫۴۸	۰٫۶۰	۱٫۴۳
مراغه	۸	پایه شعاعی	۰٫۶۷	۱٫۳۹	۰٫۵۷	۱٫۵۹
مراغه	۹	پایه شعاعی	۰٫۶۳	۱٫۴۱	۰٫۵۹	۱٫۶۷
مراغه	۱۰	پایه شعاعی	۰٫۶۲	۱٫۴۵	۰٫۶۲	۱٫۵۵

با توجه به نتایج جدول ۴، ماشین بردار پشتیبان با تابع کرنل چندجمله‌ای و به ازای پارامترهای ورودی ترکیب هشتم با مقادیر RMSE و CC به ترتیب برابر با ۲٫۱۷ و ۰٫۷۸ بهترین تخمین را برای تبخیر از تشت تبریز انجام داده است. پس از ترکیب هشتم، ترکیب سوم ورودی به ازای تابع کرنل چندجمله‌ای، نسبت به سایر ترکیب‌ها، دارای نتایج بهتری است. بهترین تخمین تبخیر ایستگاه مراغه نیز به ترتیب به ازای ترکیب‌های هفتم و چهارم انجام گرفته است. نتایج ترکیب‌های ۱ تا ۸ در تخمین تبخیر هر دو ایستگاه نشان می‌دهد که حذف یکی از پارامترهای ترکیب اول (شامل همه هفت ورودی اصلی) تأثیر چندانی در تخمین تبخیر ندارد. بنابراین، در صورت فقدان یکی از پارامترهای هفتگانه

ورودی این تحقیق، تخمین تبخیر از تشت را با مشکل جدی روبه‌رو نمی‌کند. همچنین، با توجه به نزدیک‌بودن نتایج تخمین‌ها، نمی‌توان یک ترکیب را به عنوان ترکیب مطلقاً برتر انتخاب کرد و این مسئله در ایستگاه‌های مختلف می‌تواند متغیر باشد. بنابراین، در هر ایستگاه بایستی به صورت جداگانه بهترین ترکیب و مدل برای تخمین تبخیر ارزیابی شود. همان طور که در جدول ۴ ملاحظه می‌شود، توابع کرنل چندجمله‌ای و پایه شعاعی پُرکاربردترین توابع‌اند و حاکی از قابلیت این دو تابع کرنل است. البته، ذکر این نکته لازم است که در کلیه موارد نتایج این دو تابع بسیار نزدیک به یکدیگر بوده است.

در مرحله بعدی، تبخیر از تشت به عنوان خروجی در نظر گرفته شد و هر یک از هفت متغیر- میانگین دما، میانگین فشار سطح ایستگاه، میانگین سرعت باد، میانگین رطوبت، بارش، دمای نقطه شبنم، و ساعت آفتابی- به صورت منفرد به عنوان ورودی در نظر گرفته شدند. جدول ۵ نتایج حاصل از تخمین تبخیر از تشت در دوره آموزش و صحت‌سنجی به ازای هر یک از هفت متغیر ورودی را نشان می‌دهد.

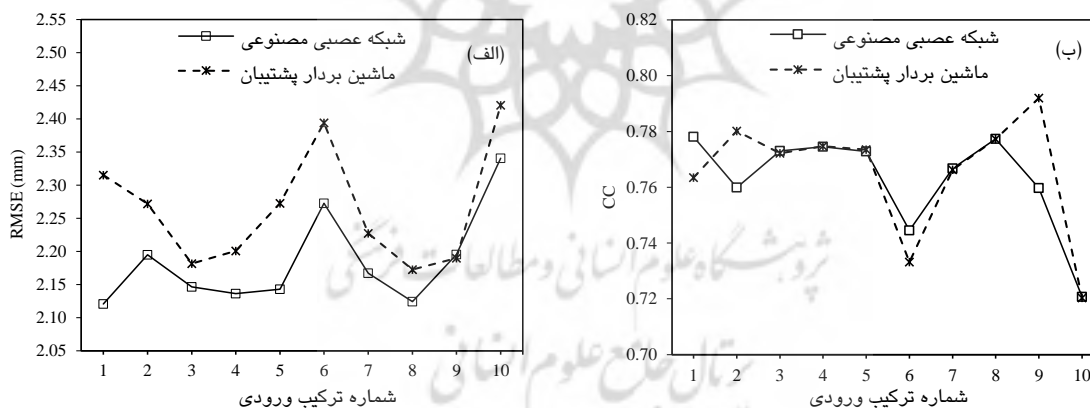
جدول ۵. نتایج تخمین تبخیر با استفاده از ماشین بردار پشتیبان به ازای هر یک از متغیرهای ورودی

ایستگاه	متغیر ورودی	تابع کرنل منتخب	آموزش		صحت‌سنجی	
			CC	RMSE(mm)	CC	RMSE(mm)
تبریز	میانگین دما	پایه شعاعی	۰/۷۲	۲/۴۲	۰/۷۰	۲/۵۳
تبریز	میانگین فشار سطح ایستگاه	پایه شعاعی	۰/۵۱	۲/۹۷	۰/۵۳	۳/۱۰
تبریز	میانگین سرعت باد	چندجمله‌ای	۰/۴۳	۳/۱۳	۰/۵۲	۳/۰۹
تبریز	میانگین رطوبت	خطی	۰/۴۲	۳/۱۸	۰/۴۲	۳/۱۷
تبریز	بارش	پایه شعاعی	۰/۱۷	۳/۴۷	۰/۱۰	۳/۴۵
تبریز	دمای نقطه شبنم	پایه شعاعی	۰/۳۱	۳/۳۵	۰/۲۵	۳/۳۶
تبریز	ساعت آفتابی	پایه شعاعی	۰/۴۸	۳/۰۹	۰/۵۰	۳/۰۲
مراغه	میانگین دما	چندجمله‌ای	۰/۵۲	۱/۶۰	۰/۵۱	۱/۶۶
مراغه	میانگین فشار سطح ایستگاه	پایه شعاعی	۰/۳۶	۱/۷۴	۰/۳۰	۱/۸۳
مراغه	میانگین سرعت باد	پایه شعاعی	۰/۱۳	۱/۸۴	۰/۱۲	۱/۹۳
مراغه	میانگین رطوبت	پایه شعاعی	۰/۴۳	۱/۶۷	۰/۴۲	۱/۸۰
مراغه	بارش	پایه شعاعی	۰/۳۴	۱/۷۷	۰/۳۳	۱/۷۸
مراغه	دمای نقطه شبنم	چندجمله‌ای	۰/۱۱	۱/۸۹	۰/۰۷	۱/۸۳
مراغه	ساعت آفتابی	چندجمله‌ای	۰/۴۶	۱/۶۶	۰/۵۰	۱/۶۵

نتایج جدول ۵ نشان می‌دهد که استفاده از میانگین دمای روزانه (به صورت منفرد) نسبت به سایر پارامترهای یادشده در تخمین میزان تبخیر هر دو ایستگاه نسبت به سایر پارامترها مؤثرتر است. بنابراین، در صورت کمبود داده در ایستگاه هواشناسی، به ترتیب اولویت می‌توان از سایر پارامتر برای تخمین تبخیر استفاده کرد. ساعت آفتابی نیز پس از دمای متوسط مهم‌ترین پارامتر در تخمین تبخیر از تحت ایستگاه‌های تبریز و مراغه است. همان طور که جدول ۵ نشان می‌دهد، در این بخش نیز تابع کرنل پایه شعاعی نسبت به دو تابع کرنل دیگر عملکرد بهتری داشته است. مقایسه نتایج جدول‌های ۴ و ۵ نشان می‌دهد که میزان تبخیر را می‌توان با استفاده از یک ورودی نیز بسیار خوب تخمین زد و اضافه کردن ورودی‌ها تا مقدار تقریباً مشخصی باعث بهبود نتایج می‌شود. یکی از مهم‌ترین دلایل این موضوع وجود همبستگی بین ورودی‌ها با تبخیر از تحت ایستگاه‌هاست.

مقایسه عملکرد مدل‌های تخمینگر

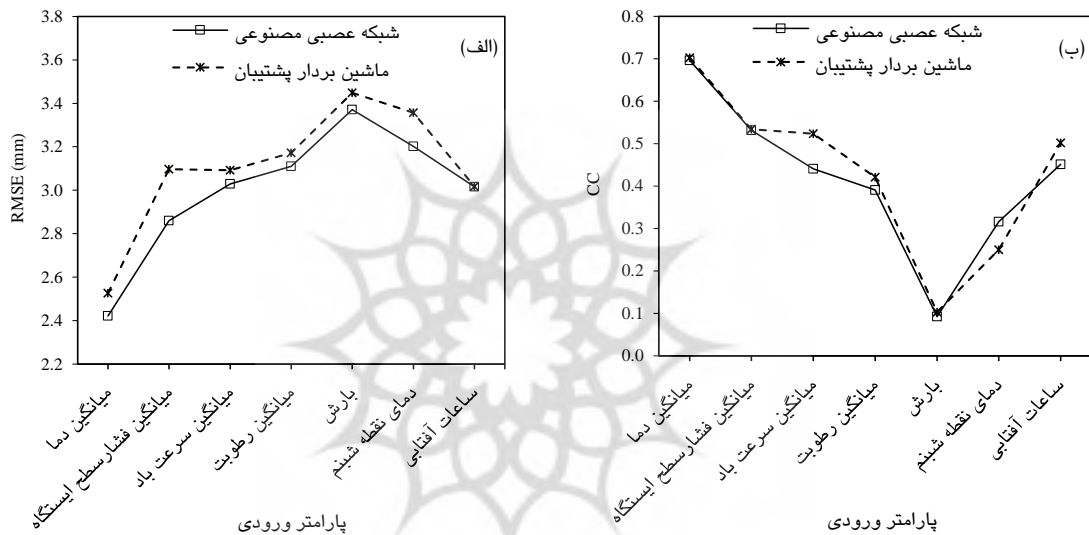
دو مدل شبکه عصبی مصنوعی و ماشین بردار پشتیبان در دو بخش مقایسه می‌شود: بخش اول مقایسه عملکرد دو مدل در ده ترکیب ورودی جدول ۱ است. از باب نمونه، مقادیر RMSE و CC دوره صحت‌سنجی تخمین تبخیر ایستگاه تبریز با استفاده از هر یک از مدل‌ها و به ازای هر ده ترکیب ورودی در شکل ۳ آورده شده است.



شکل ۳. الف) مقایسه مقادیر RMSE و ب) مدل‌های شبکه عصبی مصنوعی و ماشین بردار پشتیبان به ازای ده ترکیب ورودی

با توجه به نمودارهای فوق، در تخمین تبخیر ایستگاه تبریز، به غیر از ترکیب نه، در سایر ترکیب‌ها میزان خطای مدل شبکه عصبی مصنوعی کمتر از ماشین بردار پشتیبان است. میزان کاهش RMSE مدل شبکه عصبی مصنوعی نسبت به ماشین بردار پشتیبان بین ۰/۱۹ تا ۰/۴ میلی‌متر در روز و میزان کاهش RMSE مدل ماشین بردار پشتیبان نسبت به شبکه عصبی مصنوعی در ترکیب نه، برابر ۰/۰۱ میلی‌متر در روز است. نمودار ۲ الف نشان می‌دهد که شبکه عصبی مصنوعی در تخمین تبخیر ایستگاه تبریز قابلیت بیشتری نسبت به ماشین بردار پشتیبان داراست. مقایسه مقادیر CC نیز نشان می‌دهد که در دو مورد مقدار CC ماشین بردار پشتیبان بهتر از CC شبکه عصبی مصنوعی است و در دو مورد دیگر برعکس بوده است و در سایر موارد نیز مقادیر CC دو مدل بسیار نزدیک به یکدیگرند. در حالت کلی، مدل

ماشین بردار پشتیبان مقدار CC را کمی بهتر از مدل شبکه عصبی مصنوعی تخمین می‌زند. این روند در تخمین تبخیر ایستگاه مراغه متفاوت بوده است؛ به طوری که در جدول‌های ۲ و ۴ نشان داده شد، عملکرد ماشین بردار پشتیبان نسبت به شبکه عصبی مصنوعی بهتر بوده است. با توجه به نتایج به دست آمده، می‌توان نتیجه گرفت که تعیین مدل و ترکیب برتر بایستی بر اساس سعی و خطا انجام گیرد و به طور قطع نمی‌توان مدل یا ترکیبی را برای تخمین تبخیر کلیه ایستگاه‌ها معرفی کرد. عیسی‌زاده و همکاران نیز در مطالعه‌شان بر این موضوع تأکید کرده‌اند. بخش دوم مقایسه عملکرد دو مدل به ازای هفت ترکیب ورودی (هر یک از پارامترها به صورت جداگانه) است. از باب نمونه، مقادیر RMSE و CC هر یک از مدل‌ها در تخمین تبخیر از تشت ایستگاه تبریز به ازای هر یک از این ورودی‌ها در شکل ۴ آورده شده است.



شکل ۴. مقایسه مقادیر RMSE (الف) و CC (ب) مدل‌های شبکه عصبی مصنوعی و ماشین بردار پشتیبان به ازای هر یک از متغیرهای ورودی

با توجه به شکل ۴ الف، مدل شبکه عصبی مصنوعی به ازای تک‌ورودی‌های نشان داده شده، نسبت به مدل ماشین بردار پشتیبان، خطای کمتری در تخمین مقدار تبخیر از تشت داشته است. البته، هر دو مدل به ازای ورودی ساعت آفتابی خطای یکسانی داشته‌اند. مقدار ضریب همبستگی شبکه عصبی مصنوعی به ازای ورودی دمای نقطه شبنم ۰/۰۷ بیشتر از ماشین بردار پشتیبان بوده است و در سایر موارد میزان ضریب همبستگی ماشین بردار پشتیبان بیشتر یا برابر با شبکه عصبی مصنوعی است. در حالت کلی، ماشین بردار پشتیبان نتایج بهتری برای ضریب همبستگی به دست آورده است. از آنجا که معیار اصلی مقایسه مدل‌ها مقدار خطاست، با توجه به نتایج ذکر شده، مدل شبکه عصبی مصنوعی نسبت به مدل ماشین بردار پشتیبان عملکرد بهتری در تخمین مقدار تبخیر از تشت ایستگاه تبریز داشته است. نتایج به دست داده شده در جدول ۵، مدل ماشین بردار پشتیبان را برتر از مدل شبکه عصبی مصنوعی در تخمین تبخیر از تشت ایستگاه مراغه نشان می‌دهد. بنابراین، همان طور که در بخش قبل توضیح داده شد، برای تعیین ترکیب و مدل برتر در تخمین تبخیر هر یک از ایستگاه‌ها بایستی از سعی و خطا استفاده کرد. البته، ذکر این نکته لازم است که اختلاف مدل‌ها به ازای ترکیب‌های مشابه ناچیز است، بنابراین، مدل و ترکیب برتر می‌تواند با تغییر ایستگاه مورد ارزیابی تغییر کند.

نتیجه‌گیری

در این تحقیق از مدل‌های شبکه‌ی عصبی مصنوعی و ماشین بردار پشتیبان برای تخمین تبخیر روزانه از تشت ایستگاه‌های سینوپتیک تبریز و مراغه استفاده شد. برای تخمین هرچه بهتر تبخیر، هفده ترکیب ورودی مختلف به عنوان ورودی مدل‌ها در نظر گرفته شد و نتایج زیر حاصل شد.

نتایج نشان داد شبکه‌ی عصبی مصنوعی به ازای ترکیب اول- که شامل دمای میانگین روزانه، میانگین فشار سطح ایستگاه، میانگین سرعت باد، میانگین رطوبت، بارش روزانه، دمای نقطه‌ی شبنم، و ساعت آفتابی است- تخمین دقیق‌تری از تبخیر از تشت ایستگاه‌های تبریز و مراغه ارائه می‌دهد. در صورت فقدان داده‌ی کافی برای تخمین تبخیر، به‌ترتیب استفاده از تک‌ورودی‌ها، میانگین دمای روزانه و ساعت آفتابی- نسبت به سایر پارامترهای یادشده- در تخمین میزان تبخیر ایستگاه‌های تبریز و مراغه تأثیر بیشتری داشتند. مدل ماشین بردار پشتیبان به‌ترتیب به ازای ترکیب هشتم و هفتم بهترین تخمین را برای تبخیر از تشت هر یک از ایستگاه‌های تبریز و مراغه انجام داد. نتایج نشان داد، با حذف یکی از ورودی‌های ترکیب اصلی، تغییر چندانی در دقت مدل ایجاد نمی‌شود. عملکرد بهتر توابع کرنل نیز به‌ترتیب به تابع پایه‌ی شعاعی، چندجمله‌ای، و خطی اختصاص داشته است. در مواردی که داده‌ی مناسب و کافی برای تخمین تبخیر وجود ندارد، ماشین بردار پشتیبان دمای متوسط روزانه را مؤثرترین ورودی در نظر می‌گیرد.

بر اساس نتایج مطالعات بروتن و همکاران (۲۰۰۰)، مقدم‌نیا و همکاران (۲۰۰۹)، کیسی (۲۰۰۹، ۲۰۱۶)، طبری و همکاران (۲۰۱۰)، تزل و بویوکیلدیز (۲۰۱۵)، و بامری (۱۳۹۳)، مدل‌های شبکه‌ی عصبی مصنوعی و ماشین بردار، نسبت به سایر مدل‌های مورد مقایسه، عملکرد بهتری داشتند؛ این نتایج با نتایج مطالعه‌ی حاضر مبنی بر دقت زیاد و کارایی خوب این مدل‌ها مطابقت دارد.

نتایج ترکیب‌های ۱ تا ۸ در تخمین تبخیر هر دو ایستگاه نشان داد که حذف یکی از پارامترهای ترکیب اول (شامل همه‌ی هفت ورودی اصلی) تأثیر چندانی در تخمین تبخیر ندارد. بنابراین، در صورت فقدان یکی از پارامترهای هفتگانه‌ی ورودی این تحقیق، تخمین تبخیر از تشت را با مشکل جدی روبه‌رو نمی‌کند. با توجه به نتایج این مطالعه، عملکرد مدل‌ها بسیار نزدیک به یکدیگر بود؛ این نتیجه در مطالعه‌ی تزل و بویوکیلدیز (۲۰۱۵) نیز تأیید شده است. همچنین، با توجه به نزدیک بودن نتایج تخمین‌ها، نمی‌توان یک ترکیب را به عنوان ترکیب مطلقاً برتر انتخاب کرد و این مسئله در ایستگاه‌های مختلف می‌تواند متغیر باشد. بنابراین، در هر ایستگاه بایستی به صورت جداگانه بهترین ترکیب و مدل برای تخمین تبخیر تعیین شود. مقایسه‌ی نتایج جدول‌های ۲، ۳، ۴، و ۵ نشان می‌دهد که میزان تبخیر را می‌توان با استفاده از یک ورودی نیز بسیار خوب تخمین زد و اضافه کردن سایر ورودی‌ها در بهبود نتایج نقش کمتری دارد. یکی از مهم‌ترین دلایل این موضوع وجود همبستگی بین ورودی‌ها با تبخیر از تشت ایستگاه‌های تبریز و مراغه است. در حالت کلی، عملکرد هر دو مدل شبکه‌ی عصبی مصنوعی و ماشین بردار پشتیبان در تخمین تبخیر روزانه از تشت دقیق و کارآمد ارزیابی شده است.

منابع

- اسکندری، ع.؛ نوری، ر.؛ معراجی، ح. و کیقادی، ا. (۱۳۹۰). توسعه مدلی مناسب بر مبنای شبکه عصبی مصنوعی و ماشین بردار پشتیبان برای پیش‌بینی به هنگام اکسیژن‌خواهی بیوشیمیایی ۵ روزه، *محیط‌شناسی*، ۳۸(۶۱): ۷۱-۸۲.
- بامری، م. (۱۳۹۳). برآورد تبخیر استان سیستان و بلوچستان به روش رگرسیون خطی و شبکه عصبی مصنوعی، پایان‌نامه کارشناسی ارشد مهندسی آب، دانشکده آب و خاک دانشگاه زابل.
- زارع ایبانه، ح.؛ نوری، ع.؛ لیاقت، ع.؛ نوری، ح. و کریمی، و. (۱۳۹۰). مقایسه روش پنمن مانیتیت فائو و تشت تبخیر کلاس A با داده‌های لایسیمیتری در برآورد تبخیر و تعرق گیاه برنج در منطقه آمل، *پژوهش‌های جغرافیایی طبیعی*، ۷۶: ۷۱-۸۳.
- علیزاده، ا. (۱۳۹۰). *اصول هیدرولوژی کاربردی*، چ ۳۳، مشهد: انتشارات دانشگاه امام رضا.
- عیسی‌زاده، م. (۱۳۹۴). تخمین جریان رودخانه زرنه‌رود با استفاده از مدل‌های هیبریدی فراکوشی، پایان‌نامه کارشناسی ارشد مهندسی آب، دانشکده کشاورزی، دانشگاه تبریز.
- نجفی، ا.؛ صفاری، ا.؛ قنواتی، ع. و کرم، ا. (۱۳۹۴). شبیه‌سازی و تحلیل دبی‌های حداکثر لحظه‌ای با استفاده از شبکه عصبی مصنوعی (مطالعه موردی: ایستگاه‌های هیدرومتری هفت حوض، سولقان، قلاک، و مقصودبیک در کلان‌شهر تهران)، *پژوهش‌های ژئومورفولوژی کمی*، ۹۰-۱۰۳: (۱)۴.
- Alizadeh, A. (2011). *The principle of applied hydrology*, Imam Reza Publication, Mashhad (In Persian).
- ASCE Task Committee on Application of Artificial Neural Networks in Hydrology (2000). Artificial neural networks in hydrology, I: preliminary concepts, *Journal of Hydrologic Engineering*, 5(2): 115-123.
- Bamri, M. (2014). Evaporation estimation of Sistan and Baluchestan province using linear regression method and artificial neural network, Water Engineering master's thesis, Supervisor doctor Parviz Haghghatjoo, water and soil Faculty of the University of Zabol (In Persian).
- Baofeng, G.; Gunn, S.R.; Damper, R.I. and Nelson, J.D.B. (2008). Customizing kernel Functions for SVM-based hyperspectral image classification, *IEEE Transactions on Image Processing*, 17(4): 622-629.
- Basak, D.; Pal, S. and Patranabis, D.C. (2007). Support vector regression, *Neural Inf. Process*, 11: 203-225.
- Bruton, J.M.; McClendon, R.W. and Hoogenboom, G. (2000). Estimating daily pan evaporation with artificial neural network, *Trans. ASAE*, 43(2): 492-496.
- Dibike, Y.; Velickov, S.; Solomatine, D. and Abbott, M. (2001). Model induction with of support vector machines: Introduction and applications, *Journal of Computing in Civil Engineering*, 15(3): 208-216.
- Eskandari, A.; Nouri, R.; Meraji, H. and Kiaghadi, A. (2012). Development of appropriate model based on artificial neural network and support vector machine for forecasting 5-Days Biochemical Oxygen Demand (BOD5), *Journal of Ecology*, 61: 71-82 (In Persian).
- Fletcher, R. (1987). *Practical methods of optimization*, Wiley, New York. 456p.
- Kavzoglu, T. and Colkesen, I. (2009). A kernel functions analysis for support vector machines for land cover classification, *International Journal of Applied Earth Observation and Geoinformation*, 11(5): 352-359.
- Kim, S. and Kim, HS. (2008). Neural networks and genetic algorithm approach for nonlinear evaporation and evapotranspiration modeling, *Journal of Hydrology*, 351: 299-317.
- Kisi, O. (2009). Modeling monthly evaporation using two different neural computing, *Techniques Irrigation Science*, 27(5): 417-430.
- Kisi, O. (2016). Pan evaporation modeling using least square support vector machine, multivariate adaptive regression splines and M5 model tree, *Journal of Hydrology*, 528: 312-320.
- Liu, G.Q. (2011). Comparison of regression and ARIMA models with Neural Network models to forecast the daily stream flow, PhD thesis, University of Delaware, 545p.
- Liu, S.; Bai, J.; Jia, Z.; Jia, L.; Zhou, H. and Lu, L. (2010). Estimation of evapotranspiration in the Mu Us Sandland of China, *Hydrology and Earth System sciences*, 14: 573-584.

- Misra, D.; Oommen, T.; Agarwal, A. and Mishra, S.K. (2009). Application and analysis of support vector machine based simulation for runoff and sediment yield, *Journal of Bio Systems Engineering*, 103 (9): 527-535.
- Moghaddamnia, A.; Ghafari Gousheh, M.; Piri, J.; Amin, S. and Han, D. (2009). Evaporation estimation using artificial neural networks and adaptive neuro-fuzzy inference system techniques, *Advances in Water Resources*, 32(1): 88-97.
- Najafi, A.; Safari, A.; Ghanavati, A. And Karam, A. (2015). Simulation and analysis of maximum instantaneous flows using artificial neural network (Case study: hydrometric stations of Haft Hoz, Soghalan, Ghalak and Maghsod Beg in Tehran, *Quantitative Geomorphological Researches*, 4(1): 90-103 (In Persian).
- Nourani, V. and Sayyah Fard, M. (2012). Sensitivity analysis of the artificial neural network outputs in simulation of the evaporation process at different climatologic regimes, *Advances in Engineering Software*, 47: 127-146.
- Piri, J.; Amin, S.; Moghaddamnia, A.; Keshavarz, A.; Han, D. and Remesan, R. (2009). Daily pan evaporation modeling in a hot and dry climate, *Journal of Hydrologic Engineering*, 14(8): 803-811.
- Tabari, H.; Marofi, S. and Sabziparvar, A.A. (2010). Estimation of daily pan evaporation using artificial neural network and multivariate non-linear regression, *Irrigation Sciences*, 28: 399-406.
- Terzi, O. and Erol Keskn, M. (2005). Modeling of daily pan evaporation, *Journal of Applied Sciences*, 5(2): 368-372.
- Tezel, G. and Buyukyildiz, M. (2015). *Monthly evaporation forecasting using artificial neural networks and support vector machines*, Theoretical and Applied Climatology.
- Vapnik, V. and Chervonenkis, A. (1991). The necessary and sufficient conditions for consistency in the empirical risk minimization method, *Pattern Recognition and Image Analysis*, 1(3): 283-305.
- Zare Abyaneh, H.; Nori, H.; Liyaghat, A.; Nori, A. and Karimi, V. (2011). Compare Penman-Monteith method and pan class A by lysimeter data to estimate evapotranspiration of rice plant in the Amol region, *Physical Geography Research Quarterly*, 76: 71-83 (In Persian).
- Zealand, C.M.; Burn, D.H. and Simonovic, S.P. (1999). Short term stream flow forecasting using artificial neural networks, *Journal of Hydrology*, 214: 32-48.