

ارزیابی مدل‌های شبکه بیزین و ماشین بردار پشتیبان در برآورد تبخیر-تعرق گیاه مرجع (مطالعه موردی: خرم‌آباد)

یاسر سبزواری^۱، علی حیدر نصراللهی^{۲*}، مجید شریفی پور^۳، بابک شاهی نژاد^۴

تاریخ دریافت: ۱۳۹۸/۸/۲۱ تاریخ پذیرش: ۱۳۹۸/۱۱/۲

چکیده

در سراسر دنیا مدل فائوپنمن مانیتث به‌عنوان یک روش مرجع، برای برآورد تبخیر-تعرق مرجع مورد استفاده قرار می‌گیرد. در این روش اطلاعات ورودی زیادی نیاز است که در خیلی از موارد دسترسی به این داده‌ها مشکل است، لذا جایگزینی مدل‌های ساده‌تر با ورودی‌های اولیه کم و دقت مناسب ضرورت می‌یابد. از اینرو هدف از این پژوهش بررسی دقت و قابلیت مدل‌های ماشین بردار پشتیبان و شبکه بیزین در برآورد تبخیر-تعرق مرجع و مقایسه با مدل فائوپنمن مانیتث می‌باشد. برای اطلاعات ورودی از اطلاعات ماهانه ایستگاه سینوپتیک خرم‌آباد شامل: بیشینه و کمینه درجه حرارت، بیشینه و شش الگوی ورودی برای مدل‌سازی تعیین گردید. ۷۰ درصد داده‌ها جهت آموزش و ۳۰ درصد داده‌ها جهت صحت‌سنجی مدل‌ها به کار گرفته شد. نتایج نشان داد الگوی ترکیبی شماره ۵ شامل: حداکثر درجه حرارت، سرعت باد، تابش خورشیدی، حداقل درجه حرارت و حداقل رطوبت نسبی در همه مدل‌ها دقیق‌ترین الگو است. این الگو در مرحله آزمون در شبکه بیزین، دارای $R^2=0/97$ و $RMSE=0/93$ و در ماشین بردار پشتیبان با هسته توابع پایه شعاعی، دارای $R^2=0/98$ و $RMSE=0/41$ بود. مقایسه عملکرد مدل‌ها نشان از برتری مدل ماشین بردار پشتیبان نسبت به مدل دیگر داشت به طوری که دارای AARE به میزان ۰/۵۲۵ و MR به میزان ۱/۰۰۵ بود.

واژه‌های کلیدی: تبخیر-تعرق مرجع، رگرسیون، خرم‌آباد، شبکه بیزین، ماشین بردار پشتیبان

مقدمه

لذا عموماً از روش‌های غیرمستقیم استفاده می‌شود. با استفاده از این روش‌ها ET_0 را می‌توان با استفاده از مدل‌های آب و هواشناسی که بر پایه معادلات فیزیکی، یا رابطه تجربی متغیرهای هواشناسی هستند، برآورد و شبیه‌سازی نمود. یکی از روش‌هایی که مبنای فیزیکی دارد، روش فائو-پنمن-مانیتث (PMF-56) است که ET_0 را به صورت روزانه و ماهانه برآورد می‌کند. سازمان خواروبار جهانی (FAO) این روش را به‌عنوان معادله استاندارد توصیه کرده است (Allen et al., 1998). این روش به تعدادی از متغیرهای اقلیمی نظیر: حداقل و حداکثر دمای هوا، تابش خورشیدی، رطوبت نسبی، ساعات آفتابی و بسیاری از پارامترهای دیگر نیاز دارد. با این وجود، این متغیرهای هواشناسی خصوصاً در کشورهای در حال توسعه، یا در دسترس نبوده یا معمولاً غیرقابل اعتماد هستند (Almorox et al., 2015)؛ بنابراین ارائه روشی که بتواند با داده‌های ورودی کمتر، دقت مناسبی را به همراه داشته باشد، خصوصاً در این مناطق الزامی است. در سال‌های اخیر، روش‌های هوش مصنوعی (AI^5) با داده‌های

برآورد دقیق تبخیر-تعرق و نیاز آبی گیاهان، عاملی اصلی در کنترل چندین فرآیند هیدرولوژیکی از جمله: برنامه‌ریزی و مدیریت منابع آب خصوصاً در مناطق خشک و نیمه‌خشک (Wen et al., 2015)، تعیین قیمت آب (Pour Ali Baba et al., 2013) و نیاز آبی آبیاری (Yassin et al., 2016) است. یکی از عوامل مؤثر بر نیاز آبی گیاهان تبخیر-تعرق مرجع (ET_0) بوده که یکی از مهم‌ترین اجزای چرخه هیدرولوژی و تابع متغیرهای اقلیمی از جمله: سرعت باد در نزدیکی سطح زمین، دمای هوا، تابش خورشیدی و رطوبت نسبی است. محاسبه ET_0 بصورت مستقیم با استفاده از لایسیمتر صورت می‌گیرد که معمولاً هزینه‌بر و زمان‌بر می‌باشد (Martí et al., 2015).

۱- دانشجوی کارشناسی ارشد مهندسی آبیاری و زهکشی، دانشکده کشاورزی، دانشگاه لرستان

۲، ۳ و ۴- استادیار گروه مهندسی آب، دانشکده کشاورزی، دانشگاه لرستان
* - نویسنده مسئول: (Email: nasrolahi.a@lu.ac.ir)

چهار مدل توازن انرژی سطحی تبخیر-تعرق را در دو اقلیم بهتر پیش‌بینی کرده است (Sun et al., 2019). براساس مطالعات صورت گرفته، مدیریت بهینه منابع آب و افزایش تولید محصولات کشاورزی، تضمین‌کننده امنیت غذایی در همه کشورها است. در اینصورت، تأثیر و اهمیت اجرای راهکارهایی که بهره‌وری آب را افزایش و هدر رفت آن را کاهش می‌دهد، ضروری است. یکی از این راهکارها، برآورد دقیق ET_0 است که نقص و کمبود اطلاعات هواشناسی، مشکلی تأثیرگذار در برآورد آن می‌باشد. بنابراین ET_0 بایستی با استفاده روش‌هایی که نیازمند اطلاعات ورودی کمتر و دقت قابل‌قبولی هستند، برآورد و مدل‌سازی شود. بنابراین هدف این پژوهش الف- بررسی و تعیین پارامترهای مؤثر در برآورد ET_0 ب- بررسی دقت مدل‌های رگرسیون چندمتغیره، ماشین بردار پشتیبان (SVM) و شبکه بیزین (BN) در برآورد متوسط ماهانه ET_0 ج- معرفی مدل برتر در برآورد ET_0 در شرایط آب و هوایی خرم‌آباد است.

مواد و روش‌ها

مطالعه موردی

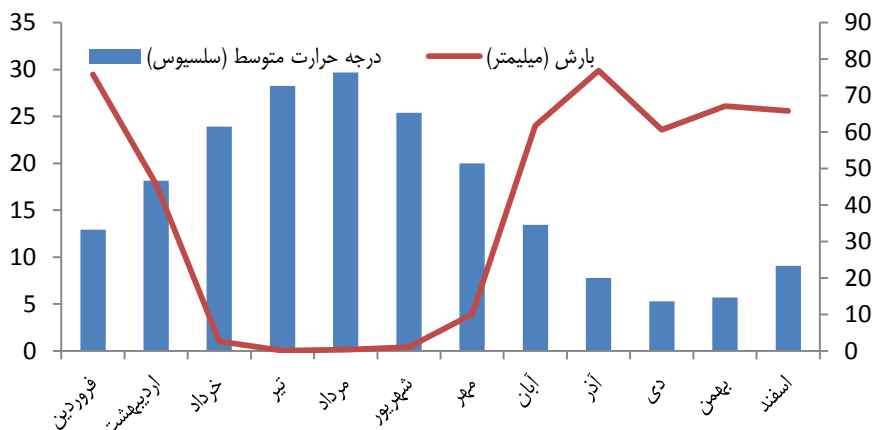
دشت خرم‌آباد، با وسعتی حدود ۲۵۰۰ کیلومترمربع در ۴۸ درجه و ۲۱ دقیقه طول شرقی و ۳۳ درجه و ۲۹ دقیقه عرض شمالی در مرکز استان لرستان واقع گردیده‌است. به‌طور کلی از نظر اقلیمی بر اساس ضریب دوماترین جز مناطق نیمه‌خشک و بر اساس کلیموگرام آمبرژه نیمه‌مرطوب سرد محسوب می‌شود. تغییرات متوسط درجه حرارت حداقل از $4/2^{\circ}C$ در زمستان تا $24/1^{\circ}C$ در تابستان و تغییرات متوسط درجه حرارت حداکثر از $6/1^{\circ}C$ در زمستان تا $42/9^{\circ}C$ در تابستان است. میانگین مجموع بارش سالیانه این منطقه حدود $466/5$ میلی‌متر است. نمودار شکل ۱ نشان‌دهنده متوسط ماهانه بارش و درجه حرارت ایستگاه خرم‌آباد در دوره زمانی مطالعاتی می‌باشد. آمار و اطلاعات هواشناسی مورد نیاز برای محاسبه‌ی تبخیر-تعرق گیاه مرجع در این پژوهش، از ایستگاه سینوپتیک خرم‌آباد با مختصات جغرافیایی طول 47° درجه و 17° دقیقه، عرض 33° درجه و 26° دقیقه و ارتفاع از سطح دریا $1147/8$ ، در یک بازه زمانی ۳۵ ساله (۱۳۹۵-۱۳۶۱) اخذ گردید. شکل ۲ نشان‌دهنده موقعیت دشت خرم‌آباد و ایستگاه سینوپتیک خرم‌آباد است. اطلاعات هواشناسی مورد استفاده شامل متوسط ماهانه: درجه حرارت حداکثر و درجه حرارت حداقل $(T_{max}$ و $T_{min})$ ، [C]، رطوبت نسبی حداکثر و رطوبت حداقل $(Rh_{min}$ و $Rh_{max})$ ، تابش خورشیدی $[R_s]$ $[MJ\ m^{-2}\ d^{-1}]$ و سرعت باد در ارتفاع دو متری (U_2) $[ms^{-1}]$ به‌منظور مقایسه ET_0 با روش فائونمن‌مانتیت می‌باشد. رابطه فائونمن‌مانتیت توسط آلن و همکاران به‌صورت رابطه ۱ ارائه شد (Allen et al., 1998).

هواشناسی محدود، به‌طور موفقیت‌آمیزی در شبیه‌سازی ET_0 مورد استفاده قرار گرفته‌اند. از جمله‌ی روش‌های هوش مصنوعی بکار گرفته‌شده برای شبیه‌سازی ET_0 می‌توان به ماشین بردار پشتیبان (Kisi and Cimen, 2009) و شبکه‌ی بیزین اشاره کرد. مقدم‌نیا و همکاران در پژوهشی با استفاده از تکنیک ماشین بردار پشتیبان، تبخیر را در زابل، جنوب شرق ایران به عنوان منطقه‌ای با اقلیم خشک شبیه‌سازی نمودند. نتایج نشان داد از بین توابع کرنل مختلف، تابع هسته پایه شعاعی، نسبت به سایر توابع برآورد بهتری از تبخیر داشته است (Moghaddamnia et al., 2008). گو و همکاران در تحقیقی تبخیر-تعرق مرجع روزانه را با استفاده از ماشین بردار پشتیبان برآورد نمودند. نتایج نشان داد که ماشین بردار پشتیبان می‌تواند داده‌های تبخیر-تعرق گیاه مرجع را تا ۹۰ درصد با مقادیر محاسبه شده‌ی روش فائو-پنمن-مانتیت تخمین بزند (Guo et al., 2011). مهدی‌زاده و همکاران با استفاده از مدل‌های رگرسیون چند متغیره تطبیقی برنامه‌ریزی بیان ژن و ماشین بردار پشتیبان، تبخیر-تعرق مرجع ماهانه را شبیه‌سازی کردند. نتایج نشان داد که روش رگرسیون چندمتغیره تطبیقی و ماشین بردار پشتیبان با هسته تابع پایه شعاعی به‌طور کلی بهتر از برنامه‌ریزی بیان ژن و ماشین بردار پشتیبان با توابع چندجمله‌ای مدل‌سازی را انجام می‌دهند. مقایسه دقت معادلات تجربی و روش‌های محاسبات نرم نشان داد که عملکرد رگرسیون چند متغیره تطبیقی و ماشین بردار پشتیبان با هسته تابع پایه شعاعی بهتر از معادلات تجربی است (Mehdizadeh et al., 2017). احمدی و همکاران (۱۳۹۵) در تحقیقی کاربرد سیستم‌های استنتاج فازی-عصبی تطبیقی و ماشین بردار پشتیبان در برآورد تبخیر-تعرق مرجع ماهانه شمال غرب کشور را مورد بررسی قرار دادند. نتایج حاصله نشان دهنده دقت بالا و خطای کم هر دو مدل در پیش‌بینی تبخیر-تعرق مرجع ماهانه بوده ولی کارایی مدل ماشین بردار پشتیبان کمی بهتر از مدل استنتاج فازی-عصبی تطبیقی بود. کیشی و علی‌ضمیر به مدل‌سازی تبخیر-تعرق مرجع با استفاده از یک روش اکتشافی پیوستگی مویک پرداختند و نتایج نشان داد که مدل‌های شامل چهار متغیر ورودی، دقت بهتری را نسبت به مدل‌های با ورودی کمتر ارائه می‌دهند. تشعشعات خورشیدی مؤثرترین متغیر در برآورد ET_0 بود (Kisi and Alizamir, 2018). عیسی‌زاده و همکاران (۱۳۹۶) در مطالعه‌ای به ارزیابی عملکرد شبکه عصبی مصنوعی (ANN) و ماشین بردار پشتیبان (SVM) در تخمین مقادیر روزانه تبخیر در ایستگاه‌های تبریز و مراغه پرداختند. نتایج نشان داد که هر دو مدل SVM و ANN عملکرد قابل‌قبولی در تخمین تبخیر-تعرق دارند. سان و همکاران در پژوهشی به بهبود شبیه‌سازی تبخیر-تعرق با استفاده از مدل میانگین شبکه بیزین با مدل توازن انرژی سطحی در چین پرداختند. نتایج نشان داد که مدل میانگین شبکه بیزین نسبت به

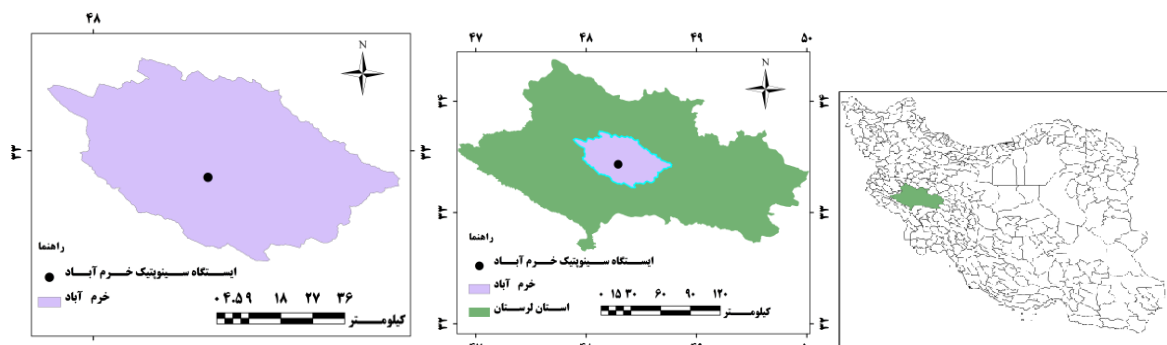
اشباع نسبت به دما، برحسب (kp/c) ؛ γ : ثابت سایکرومتری، برحسب (kp/c) ؛ $e_a - e_d$: کمبود فشار بخار در ارتفاع ۲ متری، برحسب (kp) ؛ U_2 : میانگین روزانه سرعت باد در ارتفاع ۲ متری از سطح زمین، برحسب T (m/s)؛ میانگین روزانه دمای هوا در ارتفاع ۲ متری از سطح زمین، برحسب (c).

$$ET_0 = \frac{[0.408(R_n - G) + \frac{[900\gamma]}{(T + 273)] \times U_2 \times (e_a - e_d)}{\Delta + \gamma \times (1 + 0.34U_2)} \quad (1)$$

در این رابطه: ET_0 : تبخیر-تعرق گیاه مرجع، برحسب R_n (mm/day): تشعشع خالص برحسب G (MJ/m²*day): شار گرمایی خاک، برحسب Δ (MJ/m²*day): شیب نمودار فشار بخار



شکل ۱- نمودار متوسط بارش و درجه حرارت ماهانه دوره زمانی ۱۳۶۱-۱۳۹۵



شکل ۲- موقعیت جغرافیایی دشت خرم‌آباد در لرستان و کشور

مدل رگرسیونی خطی-چند متغیره

با این روش می‌توان به‌طور هم‌زمان به تحلیل و بررسی چند متغیر مختلف پرداخت. رگرسیون چندمتغیره در واقع، رابطه بین چند متغیر پیش‌بین با متغیر پاسخ مورد نظر را بیان می‌کند. این گونه مدل‌ها دارای فرضیاتی هستند. فرض‌هایی که رگرسیون چندمتغیره را از رگرسیون ساده متمایز می‌نماید این است که:

- ۱) تعداد متغیرهای پیش‌بین (مستقل) در رگرسیون باید کمتر از تعداد مشاهدات باشد.
- ۲) همبستگی خطی کامل بین متغیرهای پیش‌بین و پاسخ وجود داشته‌باشد.

در این پژوهش، برای مدل‌های شبکه بیزین و ماشین بردار پشتیبان، T_{min} و T_{max} ، Rh_{min} ، Rh_{max} و R_s و U_2 به‌عنوان ورودی و ET_0 به‌عنوان خروجی به کار برده شد. اطلاعات مورد استفاده متشکل از دو دسته هستند. دسته اول داده‌های آموزش شامل ۷۰٪ داده‌ها (۲۹۵ ماه)، ۳۰٪ باقیمانده (۱۲۶ ماه)، داده‌های آزمون را دربردارد. جدول ۱، نشان‌دهنده پارامترهای آماری ماهانه داده‌های مورد استفاده می‌باشد. به‌منظور بررسی اثرگذاری متغیرهای مختلف هواشناسی بر ET_0 و تعیین الگوهای ورودی مدل‌سازی، از روش رگرسیون خطی چندمتغیره استفاده شد. ترکیب‌های ورودی مدل‌سازی در جدول ۲ ارائه شده است. بر این اساس ۶ الگوی ورودی برای ارزیابی ترکیب‌های مختلف متغیرهای هواشناسی تعیین شد.

جدول ۱- آمار توصیفی داده‌های مورد استفاده

متغیر	کشیدگی	چولگی	حداکثر	حداقل	انحراف معیار	میانگین
Tmax	-۱/۴۶	۰/۱۱	۴۱/۰۹	۶/۵۸	۱۰/۵۷	۲۴/۲۰
Tmin	-۱/۱۳	۰/۲۵	۲۱/۹۸	-۴/۲۳	۶/۹۸	۷/۸۶
Rhmax	-۱/۳۸	-۰/۴۸	۹۶/۳۰	۲۶/۵۲	۲۱/۱۴	۶۸/۴۳
Rhmin	-۱/۳۳	۰/۰۹	۵۸/۱۶	۵/۴۵	۱۴/۲۶	۲۷/۲۵
U ₂	۰/۰۹	-۰/۵۱	۸/۶۷	۱/۸۳	۱/۲۳	۵/۷۱
R _s	-۱/۱	۰/۱۶	۱۳/۰۱	۱/۳۰	۲/۴۰	۸/۰۲
ET ₀	-۱/۲۸	۰/۴۱	۱۵/۳۶	۰/۸۶	۴/۱۹	۶/۶۵
Tmax	-۱/۴۶	-۰/۰۵	۴۱/۰۹	۷/۴۱	۱۰/۲۱	۲۵/۹۷
Tmin	-۱/۱۸	۰/۰۷	۲۱/۹۸	-۳/۳۲	۶/۷۳	۹/۱۹
Rhmax	-۱/۵۴	-۰/۲۲	۹۷/۷۷	۲۶/۵۲	۲۰/۹۳	۶۴/۰۹
Rhmin	-۱/۱۴	۰/۳۶	۵۶/۹۶	۵/۴۵	۱۳/۳۰	۲۴/۱۰
U ₂	۰/۳۶	-۰/۳۹	۸/۶۷	۲/۲	۱/۰۱	۶/۱۵
R _s	-۱/۳	۰/۰۳	۱۲/۳۵	۳/۹۲	۲/۲۸	۸/۵۰
ET ₀	-۱/۴۳	۰/۱۹	۱۵/۳۶	۱/۷۴	۴/۱۳	۷/۵۴

جدول ۲- الگوهای ورودی مدل‌های ماشین بردار پشتیبان شبکه بیزین

ردیف	الگوی ورودی مدل
۱	ET ₀ =f{T _{max} }
۲	ET ₀ =f{T _{max} ,U ₂ }
۳	ET ₀ =f{T _{max} ,U ₂ ,R _s }
۴	ET ₀ =f{T _{max} ,U ₂ ,R _s ,T _{min} }
۵	ET ₀ =f{T _{max} ,U ₂ ,R _s ,T _{min} ,Rh _{min} }
۶	ET ₀ =f{T _{max} ,U ₂ ,R _s ,T _{min} ,Rh _{min} ,Rh _{max} }

در صورت نقض این دو فرض معادله رگرسیونی را نمی‌توان تخمین زد. مدل رگرسیونی به شکل ماتریسی را می‌توان به صورت معادله زیر نشان داد:

در صورت نقض این دو فرض معادله رگرسیونی را نمی‌توان تخمین زد. مدل رگرسیونی به شکل ماتریسی را می‌توان به صورت معادله زیر نشان داد:

$$Y = X\beta + e \quad (2)$$

که β ماتریس ضرایب رگرسیونی که ضرایب نامشخصی هستند که در حقیقت، مسئول برآورد پارامتر پاسخ هستند، e ماتریس خطای برازش و Y نیز ماتریس پاسخ است. با حل معادله (۲) برحسب β خواهیم داشت:

$$\beta = (X'X)^{-1}(X'Y) \quad (3)$$

که در رابطه بالا، X' ترانپوز ماتریس X است. برای محاسبه معکوس $(X'X)$ لازم است متغیرهای مستقل همبستگی زیادی نداشته باشند، زیرا در این صورت ماتریس $(X'X)$ را نمی‌توان معکوس کرد و باعث افزایش خطا در اثر گرد کردن داده‌ها و محاسبات می‌شود.

شبکه بیزین

شبکه بیزین، یک مدل گرافیکی احتمالاتی محسوب می‌شود

$$P(b|a) = P(a|b) \times \frac{P(b)}{P(a)} \quad (4)$$

در رابطه فوق $P(a)$: احتمال وقوع پیشامد a ، $P(b)$: احتمال وقوع پیشامد b ، $P(b|a)$: احتمال وقوع پیشامد b به شرطی که a اتفاق افتاده باشد و $P(a|b)$: احتمال وقوع پیشامد a به شرطی که b اتفاق افتاده باشد، هستند.

شبکه بیزین متشکل از گره‌ها، که متغیرهای مکانیزم مدیریتی هستند، بندها، که روابط منتج بین متغیرها را به وجود می‌آورند و احتمالات، است. شبکه بیزین ساختاری گرافیکی است که متغیرها به صورت گره و روابط علت و معلولی آن‌ها با استفاده از پیکان موسوم

هدفش حداقل سازی خطای ساختاری بوده و در نهایت منجر به یک جواب بهینه می شود (Vapnik, 1998). در مدل رگرسیونی SVR تابعی مرتبط با متغیر وابسته y که به تعدادی از متغیرهای مستقل x وابسته است، تخمین زده می شود. مانند دیگر مسائل رگرسیونی فرض می شود که رابطه میان متغیرهای وابسته و مستقل توسط یک تابع جبری مانند $f(x)$ به علاوه یک مقداری اختلال^۵ قابل محاسبه است. همانند رابطه زیر؛

$$f(x) = w^T \cdot \phi(x) + b \quad (6)$$

$$y = f(x) + noise \quad (7)$$

هدف، پیدا کردن فرم تابع $f(x)$ است. برای تحقق این کار، مدل SVR روی یک مجموعه ای از داده ها تحت عنوان مجموعه آموزش که شامل فرآیندی به منظور بهینه سازی تابع خطا است، آموزش داده می شود. در این تحقیق مدل های رگرسیونی SVR نوع دوم با نام ϵ -SVR به علت استفاده وسیع در مسائل رگرسیونی استفاده شد. در این مدل، تابع خطا به صورت زیر تعریف می شود:

$$\frac{1}{2} W^T W + C \sum_{i=1}^N \xi_i + C \sum_{i=1}^N \xi_i^* \quad (8)$$

تابع خطای رابطه ۸ با توجه به محدودیت های زیر باید حداقل شود (Hamel, 2011).

$$w^T \phi(x_i) + b - y_i \leq \epsilon + \xi_i^* \quad (9)$$

$y_i - w^T \phi(x_i) - b \leq \epsilon + \xi_i^*$ و $\xi_i \cdot \xi_i^* \geq 0 \quad i = 1, 2, 3, \dots, N$ در معادلات فوق، C ثابت گنجایش که به صورت عددی ثابت و مثبت بوده که به عنوان عامل جریمه در هنگام اتفاق افتادن خطای آموزشی مدل عمل می نماید، W بردار ضرایب، w^T ترانهاده بردار ضرایب، ξ_i^* و ξ_i متغیرهای کمبود به منظور تعیین کردن حدود بالا و پایین خطای آموزشی مرتبط با مقدار خطای مجاز ϵ ، b ثابت متغیرهای تابع رگرسیونی، N تعداد نمونه ها و ϕ تابع هسته^۶ می باشد. در نهایت می توان تابع SVR رگرسیونی را به شکل زیر بازنویسی نمود:

$$f(x_i) = \sum_{i=1}^N \bar{\alpha}_i \phi(x_i) \cdot \phi(x) + b \quad (10)$$

در معادله بالا $\bar{\alpha}_i$ میانگین ضرایب لاگرانژ است. در تحلیل مسائل غیرخطی، ماشین بردار پشتیبان ابعاد مسئله را از طریق توابع هسته تغییر می دهد. انتخاب هسته برای SVR به حجم داده های آموزشی و ابعاد بردار ویژگی، بستگی دارد. می توان از توابع مختلف هسته برای ساخت انواع مختلف SVR رگرسیونی استفاده نمود. در

به Edge مشخص می شود. چنانچه در گراف گره هایی باشد که قبل از آن ها، گره دیگری در گراف وجود داشته باشد، با توزیع احتمال شرطی تعریف می شوند، در غیر این صورت با احتمال آغازین بیان می شوند. پایین ترین بخش در شبکه بیزین از طریق قانون احتمال کل حاصل می شود. احتمالات مرتبط با بخش های بالایی شبکه بر اساس قانون بیز می باشند. در حالت کلی، اگر گره x دارای والد نباشد، گره غیرمشروط و در غیر این صورت گره مشروط است (Kuikka and Varis, 1997).

در مورد ساختار شبکه بیزین اگر مجموعه ای شامل n متغیر به صورت $\{X_1, X_2, X_3, \dots, X_n\}$ وجود داشته باشد که هر X_i یک متغیر تصادفی باشد و گره های والد این متغیر با $(X_i \text{ parent}(X_i))$ نشان داده شوند، احتمال هر رویداد مانند $X = (x_1, x_2, x_3, \dots, x_n)$ به صورت زیر محاسبه می شود.

$$P(x_1 + x_2 + \dots + x_n) = \prod P(x_i | \text{parents}(x_i)) \quad (5)$$

الگوریتم های یادگیری شبکه بیزین به الگوریتم های یادگیری بر مبنای محدودیت و بر مبنای امتیاز، طبقه بندی می شوند. دسته اول بر اساس استقلال شرطی و وابستگی بین متغیرها، توسط آزمون های آماری به دست می آید (مانند الگوریتم های PC^1 و NPT^2). در روش های یادگیری بر مبنای امتیاز نیز همه روابط ممکن بین گره ها ارزیابی شده و یک ساختار با بالاترین امتیاز به عنوان ساختار مورد نظر انتخاب می شود (Sadeghi Hesar et al., 2012). دو الگوریتم PC و NPT با توجه به سادگی، در شبکه های بیزین برای آموزش ساختار مدل سازی بیشتر مورد استفاده قرار می گیرند. در این تحقیق برای آموزش در شبکه های بیزین، از الگوریتم PC استفاده گردید. در این الگوریتم، مجموعه ای از استقلال و وابستگی عبارات شرطی توسط آزمون های آماری استخراج می شود که در آن استقلال شرطی بین هر جفت متغیر وجود نداشته و جهت کمان ها بر اساس استقلال شرطی تعیین می گردد. در این تحقیق برای آموزش پارامتری شبکه، با تنظیم دو متغیر سطح معنی داری^۳ و حداکثر اندازه مجاورت^۴ با توجه به تاثیر پارامترها بر یکدیگر استفاده شد. جهت اجرای شبکه بیزین از نرم افزار GeNie^۲ استفاده شد.

ماشین بردار پشتیبان

ماشین بردار پشتیبان از سیستم های یادگیری تحت نظارت محسوب می شود که برای طبقه بندی و رگرسیون مورد استفاده قرار می گیرد. این روش بر مبنای نظریه بهینه سازی مقید بنا شده است که

- 1- Path condition
- 2- Necessary Path Condition
- 3- Significance Level
- 4- Max Adjacency Size

5- Noise
6- Kernel

میانگین مقادیر مشاهداتی و محاسباتی می‌باشد. مقادیر آماره ریشه میانگین مربعات همواره مثبت بوده و بهترین حالت عملکرد زمانی است که مقدار به صفر نزدیک شود. مقدار مثبت آماره درصد خطای تخمین بیانگر آن است که مدل مورد بررسی تبخیر-تعرق مرجع را بیشتر برآورد کرده و مقدار منفی آن بیانگر آن است که مدل مقدار تبخیر-تعرق را کمتر برآورد می‌کند. شاخص نسبت میانگین به منظور بیش‌برآوردی یا کم‌برآوردی مدل‌ها نسبت به مقدار اندازه‌گیری شده، مورد استفاده قرار گرفت.

نتایج و بحث

رگرسیون چندمتغیره

جهت بررسی تأثیر پارامترهای ورودی مدل شامل: T_{min} , T_{max} , Rh_{min} , Rh_{max} , R_s و U_2 در برآورد ET_0 و تعیین الگوی ورودی مدل‌ها، تحلیل رگرسیونی چندمتغیره بین این پارامترها صورت گرفت که نتایج در جدول ۳ نشان داده شده است: در این جدول، مقدار ضریب همبستگی چندگانه (R)، ضریب تعیین (R^2) و ضریب تعیین تصحیح شده (R^2 Adjusted)، تجزیه واریانس و ضرایب رگرسیون نشان داده شده است. مقدار R برابر ۰/۹۷۶ و میزان R^2 تصحیح شده برابر با ۰/۹۵۲ است که رقم بالایی است و نشان می‌دهد پارامترهای پیش‌بینی توانسته‌اند ۹۵ درصد از واریانس متغیر پاسخ را پیش‌بینی کنند. در این مطالعه مقدار R^2 هم برابر ۰/۹۵۲ است که نشان از دقت بالای مدل مورد استفاده و کارایی آن دارد. با مشاهده‌ی سطح معنی‌داری حاصل شده در جدول می‌توان گفت چون سطح معنی‌داری همه پارامترها کمتر از ۰/۰۵ به دست آمده است، لذا همه این پارامترها بر متغیر ET_0 اثر معنی‌دار دارند و می‌توانند این متغیر را پیش‌بینی کنند. مقایسه ضریب استاندارد شده بدست آمده برای پارامترهای ورودی نشان می‌دهد که T_{max} با ضریب ۰/۶۰۴ دارای بیشترین تأثیر ET_0 است و پیش‌بینی کننده قوی تری نسبت به بقیه متغیرهاست. بعد از آن به ترتیب U_2 , R_s , T_{min} , Rh_{min} و Rh_{max} دارای بیشترین تأثیر در پیش‌بینی و برآورد ET_0 می‌باشند. مقدار t (چه مثبت و چه منفی) اگر بزرگ‌تر از ۱/۹۶ باشد و مقدار سطح معنی‌داری اگر کوچک‌تر از ۰/۰۱ باشد نشان می‌دهد متغیر پیش‌بین بر متغیر پاسخ تأثیر معنی‌دار دارد. با توجه به مقادیر مربوط به هر شش متغیر، همه‌ی متغیرها دارای تأثیر معنی‌دار هستند.

الگوی ورودی مدل‌ها برای پیش‌بینی تبخیر-تعرق مرجع در جدول ۲ نشان داده شده است. بر اساس این جدول، مدل‌سازی ET_0 با شش الگو انجام شد.

این تحقیق از این سه تابع هسته خطی، پایه شعاعی و چندجمله‌ای استفاده شد (Liu, 2011; Basak et al., 2007)، که روابط زیر دربردارنده روابط ریاضی آن‌ها می‌باشد.

$$K(x, x') = (x \cdot x') \quad \text{هسته خطی} \quad (11)$$

$$K(x, x') = ((x \cdot x') + 1)^d \quad \text{هسته چندجمله‌ای} \quad (12)$$

$$K(x, x') = \exp\left(-\frac{\|x - x'\|^2}{2\sigma^2}\right) \quad \text{هسته توابع پایه شعاعی} \quad (13)$$

ساده‌ترین نوع تابع هسته، هسته خطی است. هسته چندجمله‌ای، هسته‌ای است که برای مدل‌های غیرخطی مورد توجه می‌باشد و به همراه هسته پایه شعاعی در حال حاضر حائز اهمیت زیادی هستند. در روابط بالا d و σ به ترتیب؛ درجه هسته چند جمله‌ای و عرض هسته تابع پایه شعاعی است. جهت اجرای مدل ماشین بردار پشتیبان در محیط MATLAB 2019 کد نویسی انجام شده و عملکرد هسته‌های مذکور مورد بررسی قرار گرفت.

معیارهای ارزیابی

به منظور ارزیابی دقت و کارایی مدل‌ها، از معیارهای ضریب تبیین (R^2)، ریشه میانگین مربعات خطا ($RMSE^f$)، متوسط قدرمطلق خطای نسبی ($AARE^g$) (Legates and McCabe, 1999; Shiri et al., 2014) و نسبت میانگین (MR^h) به صورت زیر استفاده شد.

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (X_i - Y_i)^2} \quad (14)$$

$$R^2 = \frac{\left(\frac{\sum_{i=1}^N (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^N (X_i - \bar{X})^2 \sum_{i=1}^N (Y_i - \bar{Y})^2}} \right)^2 \quad \text{و} \quad 0 \leq R^2 \leq 1 \quad (15)$$

$$AARE = \frac{1}{n} \sum_{i=1}^n \frac{|xi - yi|}{yi} \quad (16)$$

$$MR = \frac{1}{n} \sum_{i=1}^n \frac{x_i}{y_i} \quad (17)$$

در روابط بالا، X_i و Y_i به ترتیب مقادیر محاسباتی و مشاهداتی در گام زمانی i ام، N تعداد گام‌های زمانی، \bar{X} و \bar{Y} نیز به ترتیب

- 1- Liner
- 2- Polynomial
- 3- Radial Basis Function
- 4- Root Mean Square Error
- 5- Average Absolute Relative Error
- 6- Mean Ratio

جدول ۳- مقدار ضرایب دقت مدل و پارامترهای رگرسیون چندمتغیره

خطای معیار شبیه سازی	ضریب تعیین تصحیح شده	ضریب تعیین	ضریب همبستگی چندگانه		
+ / ۹۷۷	+ / ۹۵۲	+ / ۹۵۲	+ / ۹۷۶		
سطح معنی داری	t	ضرایب استاندارد شده	ضرایب استاندارد نشده	B	خطای معیار
		Beta			
۰ / ۰۰۱	-۳ / ۳۵۷			(Constant)	-۰ / ۳۸۲
.	۶۵ / ۸۲۷	۰ / ۶۰۴		T _{max}	۰ / ۲۵
.	۷ / ۸۷۶	۰ / ۰۵۵		T _{min}	-۰ / ۳۳
.	-۴۲ / ۰۳۲	-۰ / ۱۸۶		Rh _{max}	-۰ / ۰۳۴
.	-۷ / ۴۳۴	-۰ / ۰۲۸		Rh _{min}	-۰ / ۰۰۹
.	۸۴ / ۷۹۶	۰ / ۲۱۲		U ₂	۰ / ۳۹۹
.	۲۹ / ۱۸۹	۰ / ۱۰۸		R _S	۰ / ۱۳۳

مدل سازی تبخیر- تعرق مرجع به وسیله شبکه بیزین

در مدل سازی تبخیر- تعرق مرجع به وسیله شبکه بیزین، عملکرد مدل تحت ۶ ساختار ورودی متفاوت مورد بررسی قرار گرفت. نتایج عملکرد مدل در همه ساختارها در جدول ۴ ارائه شده است. با توجه به این جدول، ساختارهای ترکیبی نسبت به ساختارهای ساده دارای عملکرد بهتری هستند. از بین ۶ ساختار متفاوت، ساختار ترکیبی

شماره ۵ با $R^2 = ۰/۹۷$ و $RMSE = ۱/۰۹$ (mm/day) در مرحله آموزش و $R^2 = ۰/۹۷$ و $RMSE = ۰/۹۳$ (mm/day) در مرحله آزمون بهترین عملکرد را نسبت به سایر ساختارها داشته و توانسته است میزان تبخیر-تعرق مرجع منطقه مورد مطالعه را با دقت مناسبی شبیه سازی نماید.

جدول ۴- نتایج حاصل از مدل شبکه بیزین

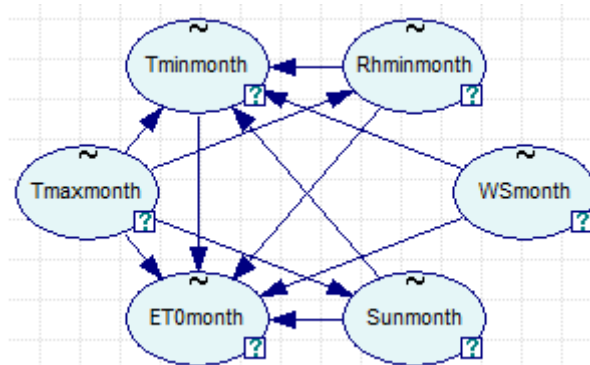
آموزش		آزمون		الگو	ردیف
RMSE	R ²	RMSE	R ²		
۲/۸۵	۰/۹۳	۲/۸۳	۰/۹۳	M1	۱
۵/۰۳	۰/۹۶	۴/۹۴	۰/۹۵	M2	۲
۵/۷۲	۰/۹۶	۵/۵۹	۰/۹۶	M3	۳
۴/۳۹	۰/۹۶	۴/۲۶	۰/۹۶	M4	۴
۰/۹۳	۰/۹۷	۱/۰۹	۰/۹۷	M5	۵
۲/۷۶	۰/۹۷	۲/۸۷	۰/۹۷	M6	۶

در شکل ۳ ساختار شبکه بیزین حاصل از الگوی منتخب نشان داده شده است. هدف از روش شبکه بیزین یافتن ارتباط تبخیر-تعرق مرجع و پارامترهای مؤثر بر آن می باشد، که منجر به ارائه رابطه ۱۸ شد.

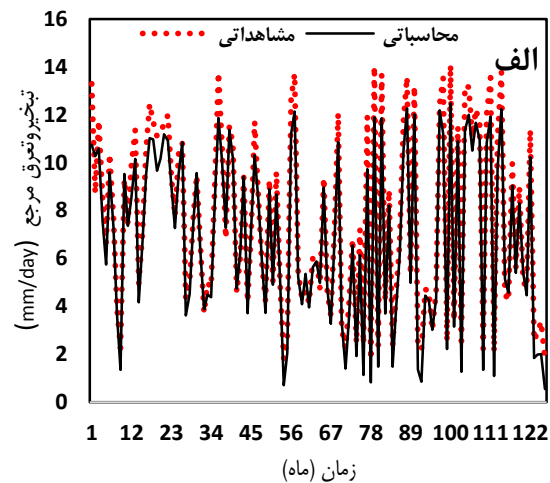
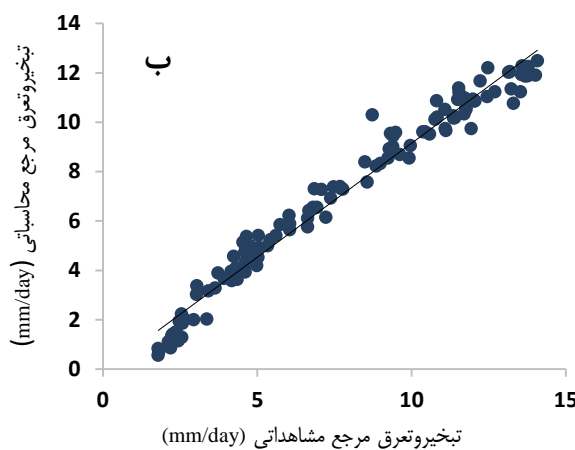
$$ET_0 = 0.0306667 * T_{max} + 0.279022 * T_{min} - 0.0638165 * Rh_{min} + 0.290493 * U_2 + 0.373788 * R_s + Normal(0.801277, 0.73961) \quad (18)$$

نمودار حاصل از ساختار منتخب از شبکه بیزین در شکل ۴ نشان داده شده است. در شکل ۴- الف مقادیر تبخیر-تعرق مرجع مشاهداتی و تبخیر-تعرق مرجع محاسباتی مرحله آموزش مدل نسبت به زمان نمایش داده شده است. شکل ۴- ب نمودار پراکنش مقادیر مشاهداتی با مقادیر محاسباتی مرحله آموزش مدل نسبت به نیمساز اول را نشان

می دهد. همان طور که در این شکل نشان داده شده است بیشتر مقادیر در اطراف خط نیمساز پراکنده شده اند که بیانگر نزدیک بودن مقادیر محاسباتی به مقادیر مشاهداتی است. اما در خصوص مقادیر حداکثر، مدل عملکرد تقریباً نامناسبی داشته و می توان اظهار کرد که مدل شبکه بیزین در تخمین مقادیر تبخیر-تعرق مرجع ضعیف عمل کرده و مقادیر را کمتر از مقدار واقعی تخمین زده است که در شکل قابل ملاحظه است.



شکل ۳- ساختار شبکه بیزین حاصل از الگوی منتخب برای شبیه‌سازی



شکل ۴- الف- نمودار مقادیر محاسباتی و پیش‌بینی شده نسبت به زمان، ب- نمودار پراکنش میان مقادیر محاسباتی و پیش‌بینی شده شبکه بیزین در مرحله آزمون

مشابه دست یافتند. آن‌ها در پژوهش خود نشان دادند تابع هسته پایه شعاعی با الگوی شامل درجه حرارت، سرعت باد، تشعشع خورشیدی و رطوبت نسبی که نسبت به سایر الگوها رطوبت نسبی وارد مدل‌سازی شد، با $R^2 = 0.97$ و $RMSE = 0.36$ (mm/day) در مرحله آزمون دارای بیشترین دقت بود (Mehdizadeh et al., 2017). شایان ذکر است که در برخی از الگوها نتایج هسته چند جمله‌ای نیز از دقت خوبی برخوردار است؛ اما در کل نتایج ماشین بردار تابع هسته پایه شعاعی عملکرد و دقت بیشتری داشت که با نتایج Mehdizadeh et al. (2017) و (Moghaddamnia et al., 2009) مطابقت دارد. در شکل ۵ نمودار حاصل از مدل منتخب را برای داده‌های بخش آزمون نشان می‌دهد. با توجه به شکل ۵، مدل ماشین بردار پشتیبان دارای عملکرد مناسبی بوده است، به طوری که مدل مقادیر پیش‌بینی شده را نزدیک به مقدار واقعی‌شان تخمین زده‌است. در تحلیل این نتایج می‌توان بیان داشت که ماشین بردار پشتیبان مدلی با دقت قابل قبول

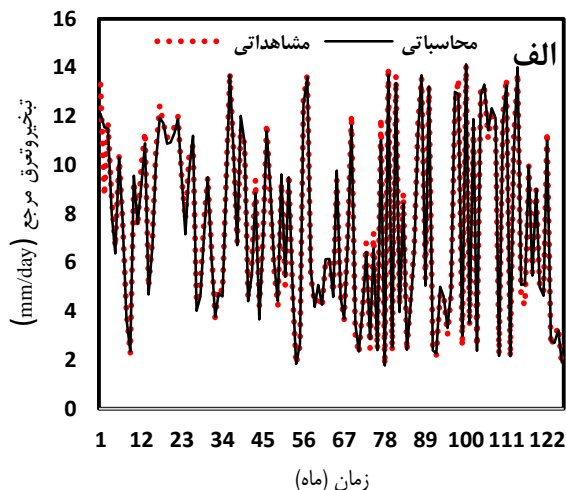
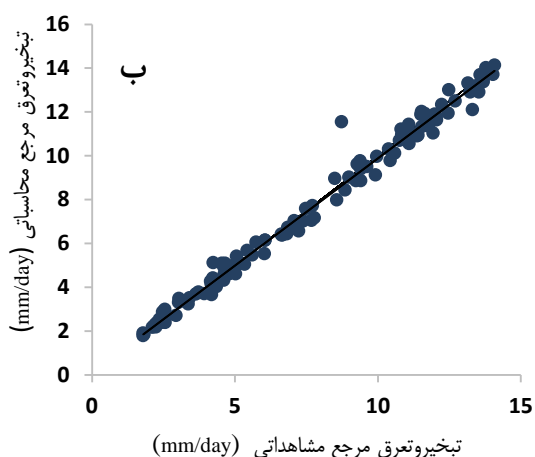
مدل‌سازی تبخیر-تعرق مرجع به وسیله ماشین بردار پشتیبان، در شبیه‌سازی تبخیر-تعرق مرجع به وسیله ماشین بردار پشتیبان، عملکرد مدل به ازای توابع مختلف هسته اعم از هسته چندجمله‌ای، خطی و توابع پایه شعاعی و با استفاده از ۶ ساختار مختلف ورودی، بررسی و ارزیابی شد. جدول ۵ نتایج حاصل از ساختارها به ازای هسته‌های مذکور که در مجموع ۱۸ الگو می‌باشند، را دربردارد. با توجه به جدول، ساختار مدل ترکیبی شماره ۵ و ۶ با هر دو تابع هسته پایه شعاعی (RBF) و هسته چندجمله‌ای (POLY) دارای بهترین عملکرد بوده‌اند. می‌توان بیان کرد چون در این دو الگو متغیر رطوبت نسبی وارد مدل شده است، دقت مدل‌سازی بهبود یافته است. الگوی شماره ۵ با هسته توابع پایه شعاعی با بیش‌ترین ضریب تبیین $R^2 = 0.98$ و کم‌ترین ریشه میانگین مربعات خطا $RMSE = 0.41$ (mm/day) در مرحله آزمون جواب بهینه‌تری نسبت به سایر ساختارها داشته است. مهدی‌زاده و همکاران هم در پژوهش خود به نتیجه

می‌شود پارامتر با سرعت و دقت بالاتری نسبت به سایر هسته‌ها تخمین زده شود که این از خصوصیات بارز توابع پایه شعاعی به شمار می‌رود.

بر مبنای نظریه بهینه‌سازی مقید است که بر مبنای کمینه‌سازی خطای ساختاری استوار است. بنابراین استفاده از روش‌های یادگیری تحت نظارت با استفاده از هسته توابع پایه شعاعی در پیش‌بینی، باعث

جدول ۵- نتایج حاصل از مدل ماشین بردار پشتیبان

آزمون		آموزش		تابع	الگو	ردیف
RMSE	R ²	RMSE	R ²			
0/99	0/93	0/83	0/96	RBF	M1	۱
0/97	0/93	0/88	0/95	POLY		
0/96	0/93	1/0۴	0/93	LINE		
0/59	0/97	0/50	0/98	RBF	M2	۲
0/60	0/97	0/51	0/98	POLY		
0/7۸	0/96	0/8۸	0/95	LINE		
0/55	0/9۸	0/4۹	0/98	RBF	M3	۳
0/55	0/9۸	0/4۹	0/98	POLY		
0/74	0/96	0/83	0/96	LINE		
0/53	0/98	0/48	0/98	RBF	M4	۴
0/52	0/98	0/42	0/9۹	POLY		
0/73	0/96	0/79	0/96	LINE		
0/4۱	0/9۸	0/3۷	0/99	RBF	M5	۵
0/42	0/98	0/37	0/99	POLY		
0/6۷	0/97	0/73	0/96	LINE		
0/4۲	0/9۸	0/37	0/9۹	RBF	M6	۶
0/4۲	0/9۹	0/37	0/99	POLY		
0/63	0/97	0/65	0/97	LINE		



شکل ۵- الف- نمودار مقادیر محاسباتی و پیش‌بینی شده نسبت به زمان، ب- نمودار پراکنش میان مقادیر محاسباتی و پیش‌بینی شده ماشین بردار پشتیبان در مرحله آزمون

در شرایط آب و هوایی خرم‌آباد شبیه‌سازی کنند. همچنین از بین مدل‌های بکاررفته مدل ماشین بردار پشتیبان با هسته توابع پایه شعاعی از بیش‌ترین دقت $R^2 = 0/98$ و کم‌ترین ریشه میانگین مربعات خطا $RMSE = 0/41$ (mm/day) در مرحله صحت‌سنجی دارا می‌باشد. مهدی‌زاده و همکاران در ارزیابی مدل‌های ماشین بردار

ارزیابی و مقایسه عملکرد مدل‌ها

به منظور مقایسه عملکرد مدل‌ها، جواب بهینه هر یک از مدل‌های مورد استفاده در این پژوهش در جدول ۶ آورده شده‌است. همانطور که در این جدول مشاهده می‌شود، مدل‌ها بر اساس معیارهای ارزیابی توانسته‌اند با دقت قابل قبولی تبخیر-تعرق مرجع را

$R^2 = 0/33$ (mm/day) و RMSE = ۰/۳۳ در مرحله آموزش، دارای بیشترین دقت بوده است (Mehdizadeh et al., 2017).

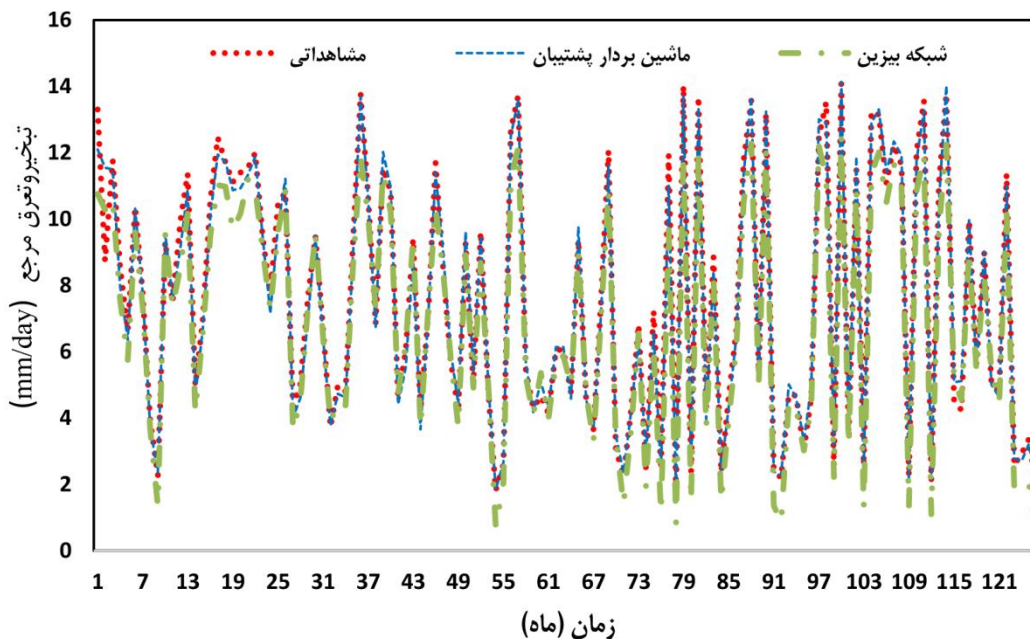
پشتیبان، برنامه‌ریزی بیان ژن و رگرسیون چند متغیره تطبیقی، نشان دادند که مدل ماشین بردار پشتیبان با هسته پایه شعاعی با $R^2 = 0/97$ و RMSE = ۰/۳۶ (mm/day) در مرحله آزمون و ۰/۹۸

جدول ۶- نتایج عملکرد مدل‌های شبکه بیزین و ماشین بردار پشتیبان

آزمون		آموزش		مدل	
RMSE	R^2	RMSE	R^2		
۰/۹۳	۰/۹۷	۱/۰۸	۰/۹۶	شبکه بیزین	
0/41	0/98	0/37	0/99	ماشین بردار پشتیبان	

ارائه دهد. در تبیین نتایج بالا می‌توان بیان نمود که ماشین بردار پشتیبان بر خلاف سایر شبکه‌های عصبی که خطای طبقه‌بندی را کمینه می‌کنند، ریسک عملیاتی را به‌عنوان تابع هدف در نظر گرفته و مقدار بهینه آن را حساب می‌کند. بنابراین با بکارگیری آن با استفاده از توابع پایه شعاعی در پیش بینی، باعث می‌شود تخمین پارامتر، دقت بالاتر و خطای کمتری نسبت به سایر هسته‌ها داشته باشد.

در شکل ۶ نتایج بهینه به‌دست آمده از مدل نسبت به زمان نشان داده شده‌است. با توجه به شکل ۶ مشاهده می‌شود که ماشین بردار پشتیبان در برآورد مقادیر کمینه و بیشینه نسبت به شبکه بیزین عملکرد بهتری داشته و توانسته است مقادیر تخمینی را نزدیک به مقادیر مشاهداتی پیش‌بینی نماید. در مقابل شبکه بیزین در تخمین مقادیر کمینه عملکرد خوبی نداشته و نتوانسته است دقت مناسبی را



شکل ۶- نمودار مقادیر محاسباتی و مشاهداتی مدل‌های مختلف نسبت به زمان

شاخص MR، مدل ماشین بردار پشتیبان با هسته پایه شعاعی ۰/۰۰۵ کم‌ترین بیش‌برآوردی و مدل شبکه بیزین با مقدار شاخص ۰/۱۸۳ بیش‌ترین بیش‌برآوردی را نسبت به مقادیر مشاهداتی داشته‌است.

در پایان به منظور مقایسه عملکرد دو مدل، از شاخص‌های MR و AARE استفاده شد که نتایج آن در جدول ۷ آورده شده است. بر اساس AARE، مدل ماشین بردار پشتیبان با هسته چندجمله‌ای کم‌ترین میزان خطا (۰/۰۵۲۵) را داشته‌است. همچنین بر اساس

جدول ۷- شاخص‌های خطای برآورد مدل‌های شبکه بیزین و ماشین بردار پشتیبان

شاخص ارزیابی	شبکه بیزین	ماشین بردار پشتیبان
AARE	۰/۲۴۵۴	۰/۰۵۲۵
mr	۱/۱۸۳	۱/۰۰۵

نتیجه گیری

در این پژوهش عملکرد مدل‌های ماشین بردار پشتیبان و شبکه بیزین جهت برآورد تبخیر-تعرق مرجع با استفاده از داده‌های ماهانه ایستگاه خرم‌آباد طی دوره زمانی ۱۳۶۱ تا ۱۳۹۵ مورد بررسی قرار گرفت. تبخیر-تعرق مرجع پیش‌بینی شده در مدل‌های مذکور، بر اساس شاخص‌های ارزیابی و تبخیر-تعرق مرجع مشاهداتی (روش فائو-پنمن-مانتیث) مقایسه و ارزیابی شد. اجرای مدل‌ها با ساختارهای ورودی مختلف نشان داد که در هر دو مدل مورد بررسی، ساختارهای ترکیبی نسبت به سایر ساختارها از عملکرد بهتری داشته‌اند که نشان‌دهنده انتخاب مناسب پارامترهای ورودی و مؤثر بودن آن‌ها در مدل‌سازی بوده‌است. از سوی دیگر، افزایش تعداد متغیرهای مؤثر در ورودی باعث گسترش حافظه مدل برای تخمین مقادیر خروجی است که تعداد داده‌ها جهت آموزش شبکه بیشتر شده و شبکه به خوبی تعمیم داده‌شود. لازم به ذکر است که مدل‌ها در ساختارهایی با متغیرهای کمتر نیز جواب‌های خوبی ارائه نمودند که می‌تواند در ایستگاه‌هایی با نواقص آماری در برآورد تبخیر-تعرق مرجع استفاده شوند. نتایج حاصل از شاخص‌های ارزیابی نشان داد مدل ماشین بردار پشتیبان دارای دقت بهتر و خطای کمتری نسبت به مدل شبکه بیزین است که در تحلیل این موضوع می‌توان گفت که دقت بالای ماشین بردار پشتیبان ناشی از حداقل‌سازی خطای ساختاری است. استفاده از روش‌های یادگیری تحت نظارت با توابع پایه شعاعی باعث می‌شود تخمین پارامتر از دقت و سرعت بالا و خطای کمتری نسبت به سایر مدل‌ها برخوردار شود و این یکی از ویژگی‌های ممتاز توابع پایه شعاعی به شمار می‌رود.

منابع

- احمدی، ف.، آیشم، س.، خلیلی، ک. و بهمنش، ج. ۱۳۹۵. کاربرد سیستم استنتاج فازی-عصبی تطبیقی و ماشین بردار پشتیبان برای برآورد تبخیر-تعرق مرجع ماهانه شمال غرب کشور. نشریه آب و خاک (علوم و صنایع کشاورزی). جلد ۳۰، شماره ۱، فروردین-اردیبهشت ۱۳۹۵، ۲۶۰-۲۷۴.
- عیسی‌زاده، م.، شیرزاد، م. و رضایی بنفشه، م. ۱۳۹۶. ارزیابی عملکرد شبکه عصبی مصنوعی (ANN) و ماشین بردار پشتیبان (SVM) در تخمین مقادیر روزانه تبخیر (مطالعه موردی: ایستگاه هواشناسی تبریز و مراغه). پژوهش‌های جغرافیای طبیعی، دوره ۴۹، شماره ۱، بهار ۱۳۹۶، ۱۵۱-۱۶۸.
- Allen, R.G., Pereira, L.S., Raes, D. and Smith, M. 1998. Crop evapotranspiration-Guidelines for computing crop water requirements-FAO Irrigation and drainage paper 56. Fao, Rome, 300(9), p.D05109.
- Almorox, J., Quej, V.H. and Martí, P. 2015. Global performance ranking of temperature-based approaches for evapotranspiration estimation considering Köppen climate classes. Journal of Hydrology, 528, pp.514-522.
- Basak, D., Pal, S. and Patranabis, D.C. 2007. Support vector regression. Neural Information Processing-Letters and Reviews, 11(10), pp.203-224.
- Davies, P.E. 2007. Bayesian Decision Networks for Management of High Conservation Assets (National Water Initiative-Australian Government Water Fund. Report 6/6 Report to the Conservation of Freshwater Ecosystem Values Project, Water Resources Division, Department of Primary Industries and Water).
- Guo, X., Sun, X. and Ma, J. 2011. Prediction of daily crop reference evapotranspiration (ET₀) values through a least-squares support vector machine model. Hydrology Research, 42(4), pp.268-274.
- Hamel, L.H. 2011. Knowledge discovery with support vector machines (Vol. 3). John Wiley & Sons.
- Kisi, O. and Alizamir, M. 2018. Modelling reference evapotranspiration using a new wavelet conjunction heuristic method: Wavelet extreme learning machine vs wavelet neural networks. Agricultural and Forest Meteorology, 263, pp.41-48.
- KIŞI, O. and Cimen, M. 2009. Evapotranspiration modelling using support vector machines/Modélisation de l'évapotranspiration à l'aide de 'support vector machines'. Hydrological Sciences Journal, 54(5), pp.918-928.
- Kuikka, S. and Varis, O. 1997. Uncertainties of climatic change impacts in Finnish watersheds: a Bayesian network analysis of expert knowledge. Boreal Environment Research, 2, pp.109-109.
- Legates, D.R. & McCabe, G.J. 1999. Evaluating the use of goodness-of-fit measures in hydrologic and hydroclimatic model validation. Water Resources Research 35 (1), 233-241.
- Lerner, U., Parr, R., Koller, D. & Biswas, B. 2000.

- Monthly rainfall forecasting using Bayesian belief networks. *International Research Journal of Applied and Basic Sciences*, 3:2226-223130.
- Shiri, J., Nazemi, A.H., Sadraddini, A.A., Landaras, G., Kisi, O., Fard, A.F. and Marti, P. 2014. Comparison of heuristic and empirical approaches for estimating reference evapotranspiration from limited inputs in Iran. *Computers and Electronics in Agriculture*, 108, pp 230-241.
- Sun, H., Yang, Y., Wu, R., Gui, D., Xue, J., Liu, Y. and Yan, D. 2019. Improving Estimation of Cropland Evapotranspiration by the Bayesian Model Averaging Method with Surface Energy Balance Models. *Atmosphere*, 10(4), p.188.
- Vapnik, V. and Vapnik, V. 1998. *Statistical learning theory*.
- Wen, X., Si, J., He, Z., Wu, J., Shao, H. and Yu, H. 2015. Support-vector-machine-based models for modeling daily reference evapotranspiration with limited climatic data in extreme arid regions. *Water Resources Management*, 29(9), pp.3195-3209.
- Yassin, M.A., Alazba, A.A. and Mattar, M.A. 2016. Artificial neural networks versus gene expression programming for estimating reference evapotranspiration in arid climate. *Agricultural Water Management*, 163, pp.110-124.
- Bayesian fault detection and diagnosis in dynamic systems. In *Aaai/iaai*. 531-537.
- Liu, G.Q. 2011. Comparison of regression and ARIMA models with neural network models to forecast the daily streamflow of White Clay Creek. PhD thesis. University of Delaware. 545 pages.
- Martí, P., González-Altozano, P., López-Urrea, R., Mancha, L. A., & Shiri, J. 2015. Modeling reference evapotranspiration with calculated targets. Assessment and implications. *Agricultural Water Management*, 149, 81-90.
- Mehdizadeh, S., Behmanesh, J., & Khalili, K. 2017. Using MARS, SVM, GEP and empirical equations for estimation of monthly mean reference evapotranspiration. *Computers and Electronics in Agriculture*, 139, 103-114.
- Moghaddamnia, A., Ghafari, M., Piri, J. and Han, D. 2009. Evaporation estimation using support vector machines technique. *International Journal of Engineering and Applied Sciences*, 5(7), pp.415-423.
- Pour-Ali Baba, A., Shiri, J., Kisi, O., Fard, A. F., Kim, S., & Amini, R. 2013. Estimating daily reference evapotranspiration using available and estimated climatic data by adaptive neuro-fuzzy inference system (ANFIS) and artificial neural network (ANN). *Hydrology Research*, 44(1), 131-146.
- SadeghiHesar, A., Tabatabaee, H., & Jalali, M. 2012.

Evaluation of Bayesian Network and Support Vector Machine Models in Estimation of Reference Evapotranspiration (Case Study: Khorramabad)

Y. Sabzevari¹, A. Nasrolahi^{2*}, M. Sharifipour³, B. Shahinejad⁴

Recived: Nov.12, 2019

Accepted: Jan.22, 2020

Absract

Around the world, the Penman-Monteithe-FAO model is used as a reference method to estimate reference evapotranspiration. This method requires a lot of input data, which in many cases are difficult to access, so it is necessary to replace simpler models with low inputs and good accuracy. Therefore, the purpose of this study was to evaluate the accuracy and capability of Bayesian Network and Support Vector Machine models in estimating reference evapotranspiration and comparing it with the Penman-Monteithe-FAO model. For input data, monthly data of Khoramabad synoptic station including: maximum and minimum temperature, maximum and minimum relative humidity, solar radiation and wind speed in period 1990-2016 (420 months) were used. Based on the effect of input parameters on output, six input patterns were determined for modeling. 70% of data were used for training and 30% for model validation. The results showed that pattern number 5 includes: maximum Temperature, wind speed, solar radiation, minimum temperature and minimum relative humidity has the best accuracy in all models. This pattern in test phase, has $R^2 = 0.97$, RMSE = 0.93 and $R^2 = 0.98$, RMSE = 0.41 respectively in the Bayesian network and support Vector Machine with radial basis functions kernel. Comparison of the models showed that the support vector machine has more accuracy with AARE of 0.0525 and MR of 0.005.

Keywords: Bayesian Network, Khorramabad, Reference Evapotranspiration, Regression, Support Vector Machine

1- M.Sc. Student, Department of Water Engineering, Faculty of Agriculture and Natural Resources, Lorestan University
2, 3 and 4- Assistant Professor, Department of Water Engineering, Faculty of Agriculture and Natural Resources, Lorestan University

(*- Corresponding Author Email: nasrolahi.a@lu.ac.ir)