

## مدل‌سازی تبخیر از سطح آزاد آب با استفاده از ماشین بردار پشتیبان و حداقل مربعات ماشین بردار پشتیبان

معصومه فراستی<sup>۱</sup>، سیدمرتضی سیدیان<sup>۲</sup>، کیمیا داب<sup>۳</sup>

تاریخ ارسال: ۱۳۹۸/۰۲/۲۹

تاریخ پذیرش: ۱۳۹۹/۰۷/۲۳

مقاله علمی پژوهشی می باشد

### چکیده

تبخیر یکی از فرایندهای مهم و تأثیرگذار در چرخه آبی است. تشت تبخیر به علت سهولت تفسیر داده‌های آن در سراسر دنیا به‌عنوان شاخصی برای تعیین تبخیر از دریاچه‌ها و مخازن استفاده می‌شود. بنابراین با ثبت درست مقدار تبخیر از تشت می‌توان تبخیر و تعرق گیاه مرجع را تخمین زد. روابط تجربی ارائه‌شده برای تخمین تبخیر از سطوح آزاد با در نظر گرفتن پارامترهای هواشناسی به‌عنوان ورودی، دارای تنوع زیاد است. دقت روابط تجربی در مناطق مختلف متفاوت است و در هر منطقه نیاز به واسنجی دارد. همچنین از دقت بالایی برخوردار نبوده و دسترسی به تمام پارامترهای ورودی مشکل و یا اندازه‌گیری آن‌ها محتاج صرف هزینه و زمان زیادی می‌باشد. هدف از این تحقیق ارزیابی کارایی مدل‌های ماشین بردار پشتیبان و حداقل مربعات ماشین بردار پشتیبان جهت تخمین تبخیر از سطح آزاد آب در استان گلستان می‌باشد. در این تحقیق از داده‌های هواشناسی روزانه سه ایستگاه سینوپتیک (کلاله، گرگان و بندر ترکمن) به مدت ۱۷ سال (۱۳۹۳-۱۳۷۶) استفاده شد. نتایج نشان داد بین الگوهای ورودی به مدل‌های SVM و LSSVM، الگوی ۱۶ پارامترهای ورودی رطوبت نسبی کمینه، رطوبت نسبی بیشینه، سرعت باد و ساعات آفتابی دارای بیشترین  $R^2$  و کمترین RMSE و MBE بود. مدل LSSVM در ایستگاه بندر ترکمن دارای بهترین پیش‌بینی نسبت به دو ایستگاه دیگر بوده است. همچنین در همه ایستگاه‌های مورد مطالعه مدل LSSVM دارای  $R^2$  بیشتر و RMSE و MBE کمتری نسبت به مدل SVM بوده است.

واژه‌های کلیدی: تبخیر از سطح آزاد آب، مدل SVM، مدل LSSVM، تشت تبخیر

<sup>۱</sup>. دانشیار گروه مرتع و آبخیزداری، دانشکده کشاورزی، دانشگاه گنبد کاووس، گنبد، ایران، ایمیل نویسنده مسئول: Farasati2760@gmail.com (نویسنده مسئول)

<sup>۲</sup>. استادیار گروه مرتع و آبخیزداری، دانشکده کشاورزی، دانشگاه گنبد کاووس، گنبد، ایران، s.m.seyedian@gmail.com

<sup>۳</sup>. دانشجوی کارشناسی، دانشکده کشاورزی، دانشگاه گنبد کاووس، گنبد، ایران، daab@gmail.com

## مقدمه

تبخیر و تعرق یکی از مؤلفه‌های اصلی بیلان آب در یک منطقه و از جمله عوامل اثرگذار جهت برنامه‌ریزی دقیق آبیاری می‌باشد. لذا، برآورد دقیق نیاز آبی گیاه، تأثیر شایانی بر کاهش معضل بحران آب، به‌ویژه در مناطق خشک و نیمه‌خشک خواهد داشت. برآورد صحیح تبخیر در برنامه‌ریزی‌های منابع آب ضروری است. تبخیر از سطح آزاد به دلیل سهولت تفسیر داده‌های آن در سراسر دنیا، برای تعیین تبخیر دریاچه‌ها و مخازن استفاده می‌شود (زارع ایبانه، ۱۳۹۰). یکی از پیش‌نیازهای مدیریت بهینه آب در حوضه آبخیز تخمین دقیق مؤلفه‌های بیلان آب است. برآورد دقیق تبخیر و تعرق در مطالعاتی از قبیل تغییر اقلیم جهانی، تکامل محیطی و کنترل منابع آب نقش مهمی بازی می‌کند (Liu et al., 2010). 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 به‌دقت فراوانی نیاز دارد. یکی از رایج‌ترین روش‌های غیرمستقیم، استفاده مقادیر برآورد شده‌ی تبخیر تعرق گیاه مرجع- با استفاده از معادله‌ی فائوپنمن-مانتیت اصلاح‌شده می‌باشد. در این روش، پس از تعیین نیاز آبی گیاه مرجع، نیاز آبی گیاه موردنیاز بر اساس ضرایب گیاهی به دست می‌آید (Allen et al., 1998). تعداد داده‌های موردنیاز در روش پنمن-مانتیت به‌رغم دقت قابل‌قبول آن، همواره ذهن محققان را به سمت استفاده از روش‌های ساده‌تر و درعین‌حال دقیق معطوف ساخته است. نتایج پژوهش‌های پیشین حاکی از این واقعیت است که مدل‌های خطی، به دلیل تغییرپذیری زیاد مؤلفه‌ی تبخیر تعرق عملکرد- مناسبی در برآورد آن نداشته و باید از مدل‌های غیرخطی استفاده نمود که خود نیازمند برآورد پارامترهای زیادی خواهد بود (wang et al., 2016). در این راستا، مدل‌های هوشمند با توجه به توانایی در حل پدیده‌های غیرخطی و پیچیده می‌تواند مفید واقع شود. از جمله‌ی این مدل‌ها می‌توان به مدل‌های ماشین بردار پشتیبان (SVM) و LSSVM اشاره نمود.

SVM یکی از روش‌های یادگیری با ناظر است که از آن برای طبقه‌بندی و رگرسیون استفاده می‌شود. رویکرد

SVM به این صورت است که در فاز آموزش، سعی می‌شود که مرز تصمیم‌گیری به‌گونه‌ای انتخاب گردد که حداقل فاصله آن با هر یک از دسته‌های موردنظر ماکزیمم گردد. این نوع انتخاب باعث می‌شود که تصمیم‌گیری در عمل، شرایط نویزی را به‌خوبی تحمل کند و پاسخدهی خوبی داشته باشد. SVM روشی برای طبقه‌بندی دوتایی در فضای ویژگی‌های دلخواه است و از این‌رو روشی مناسب برای مسائل پیش‌بینی بشمار می‌رود (pai, 2007).

از مدل خبره SVM بیشتر در مسائل هیدرولوژیکی و منابع آب استفاده شده است. مدل SVM با استفاده از مسئله بهینه‌سازی دوگانه برای پیش‌بینی رواناب روزانه از رودخانه بختیاری با استفاده از داده‌های هواشناسی و بارندگی استفاده گردید و نتایج آن با مدل ANN و ANN-GA مقایسه شد. روش SVM منجر به مقادیر RMSE کمتری نسبت به دو روش دیگر شد و روشی سریع‌تر و با دقت بالاتر بود (behzad., 2009). SVM به‌عنوان یک روش موفق در مطالعات هیدرولوژیکی مانند پیش‌بینی رواناب (۹) ، پیش‌بینی سیلاب (liong and sivapragasam, 2002) ; (Yu et al., 2006) و کنترل آب زیرزمینی (Asefa, 2004) استفاده گردید. در زمینه تبخیر - تعرق نیز توانایی SVM در اصلاح دقت برآورد تبخیر روزانه در مخازن چاه نیمه زابل موردبررسی قرار گرفت (Moghadamnia, 2006). اسلامیان و همکاران (Eslamian, 2009) برای برآورد تبخیر از گلخانه از مدل SVM و ANN استفاده و نتایج دو مدل را با یکدیگر مقایسه کردند. مقادیر ضرایب همبستگی برای ANN و SVM به ترتیب ۰/۹۲ و ۰/۹۶ و مقادیر راندمان به ترتیب ۰/۸۳ و ۰/۹۱ به دست آمد که این مقادیر بیانگر توانایی دو مدل SVM و ANN در برآورد تبخیر در گلخانه می‌باشد. (Kisi and Cimen, 2010) و Cimen در تحقیقاتی که عملکرد مدل SVM و روابط تجربی هارگریوز، ریتچ و تورک را در پیش‌بینی تبخیر و تعرق روزانه مورد مقایسه قرار دادند. به این نتیجه رسیدند که مدل SVM با بهره‌گیری از ورودی‌های درجه حرارت روزانه، تابش خورشیدی، رطوبت نسبی و سرعت باد از دقت بسیار بالایی در پیش‌بینی تبخیر و تعرق مرجع روزانه برخوردار است. رضایی و همکاران (۲۰۱۳) به ارزیابی کارایی مدل حداقل



گرفت. نتایج تحقیق نشان داد که روش عصبی فازی نسبت به بقیه روش‌ها، از دقت بیشتری برای برآورد تبخیر از تشت برخوردار است (شادمانی، ۲۰۱۱).

Wang و همکاران (2016) به مدل‌سازی تبخیر از تشت با استفاده از شش روش پرسپترون چندلایه، ژنتیک فازی، دو روش رگرسیون و حداقل مربعات ماشین بردار پشتیبان در اقلیم‌های مختلف چین پرداختند. نتایج نشان داد که مدل‌های مورد مطالعه نسبت به روش‌های رگرسیون تبخیر تعرق را با دقت بیشتری تعیین کردند.

بررسی روند تبخیر و تعرق پتانسیل ماهانه در استان سیستان و بلوچستان در قالب ۱۲ الگوی زمانی تأخیری با استفاده از برنامه‌ریزی ژنتیک برای در یک دوره ۴۳ ساله مدل‌سازی گردید. نتایج نشان داد که تمامی مدل‌های این تحقیق در الگوی ۱۳ تأیید شده است. Guo و همکاران (۲۰۱۱) در تحقیقی تبخیر و تعرق گیاه مرجع به صورت روزانه با استفاده از ماشین بردار پشتیبان برآورد کردند. نتایج تحقیق نشان داد که مدل ماشین بردار پشتیبان می‌تواند داده‌های تبخیر و تعرق گیاه مرجع را تا ۵۳ درصد با مقادیر محاسبه شده با روش پنمن مانیتیت تقریب بزند.

تحقیقاتی در دو منطقه از شمال کالیفرنیا از مدل‌های رگرسیون، ماشین بردار پشتیبان، شبکه‌های عصبی و مدل‌های تجربی برای برآورد تبخیر و تعرق روزانه گیاه مرجع در استفاده گردید نتایج نشان داد که مدل ماشین بردار پشتیبان در هر دو قسمت کالیفرنیا از دقت بالاتری نسبت به بقیه مدل‌ها برخوردار است (Guo, 2011; Kisi and Zounemat, 2014).

در تحقیقی در جنوب شرق ترکیه برای پیش‌بینی تبخیر ماهانه از شبکه عصبی مصنوعی و ماشین بردار پشتیبان استفاده شد. نتایج نشان داد که هر دو مدل شبکه‌های عصبی و ماشین بردار پشتیبان از دقت بالا و نزدیک به هم برخوردار هستند (Kisi and Zounemat, 2014). در تحقیقی در هند برای برآورد تبخیر از تشتک، از مدل‌های شبکه‌های عصبی مصنوعی، نرو فازی، ماشین بردار پشتیبان، منطق فازی و مدل‌های تجربی هارگریوز سامانی و استفان - استوارت استفاده گردید. نتایج تحقیق نشان داد که مدل‌های ماشین بردار پشتیبان و منطق فازی از دقت

مربعات ماشین بردار پشتیبان در پیش‌بینی سطح ایستابی پرداختند. نتایج چهار ترکیب، نشان داد که عملکرد مدل LSSVM نسبت به مدل ANN برتری دارد. رضایی و همکاران (۱۳۹۲) شبکه پایش سطح آب زیرزمینی را با استفاده از مدل حداقل مربعات ماشین بردار پشتیبان طراحی کردند. نتایج حاصل از این تحقیق نشان می‌دهد که ترکیب برتر مدل LSSVM دربرگیرنده شاخص‌های عملکرد است. Wang و همکاران (2016) به مدل‌سازی تشت از تبخیر با استفاده از شش روش محاسبات در اقلیم‌های مختلف در چین پرداختند. با توجه به پیشینه پژوهش، در این تحقیق سعی شد کارایی مدل‌های ماشین بردار پشتیبان (SVM) و حداقل مربعات بردار پشتیبان (LSSVM) در تخمین مقدار تبخیر از سطح آزاد برای ایستگاه‌های کلاله، بندر ترکمن و گرگان مورد بررسی قرار گرفت. بدین منظور عملکرد مدل‌های ذکر شده به ازای هدفه ترکیب مختلف ورودی بررسی شد.

احمدی و همکاران (2017)، به بررسی کاربرد سیستم‌های استنتاج فازی-عصبی و ماشین بردار پشتیبان برای برآورد تبخیر تعرق ماهانه پرداختند. نتایج نشان داد که مدل SVM کمی بهتر از مدل Anfis بود. همچنین زمانی که از حافظه سری زمانی تبخیر و تعرق برای ورودی مدل‌ها استفاده گردید، نسبت به حالتی که از متغیرهای اقلیمی به‌عنوان ورودی استفاده شد، دقت کمتر بود.

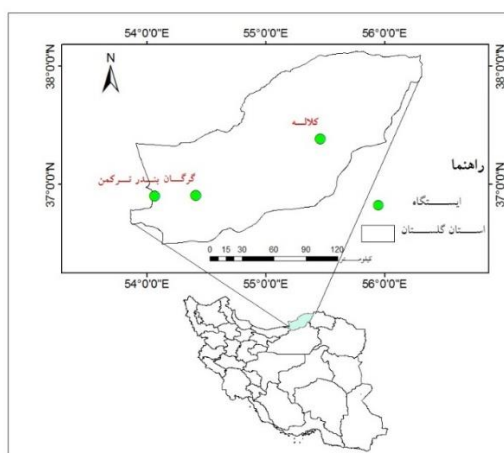
نتایج پژوهشی در ترکیه نشان داد که روش‌های فرا ابتکاری قابلیت بالایی در برآورد میزان تبخیر تعرق دارند (Kisi and Zounemat, 2014). زیرا این مدل‌ها اولاً نیاز به رابطه ریاضی برای پدیده‌های پیچیده مورد بررسی ندارند، ثانیاً چون با مدل‌های بهینه‌سازی تلفیق شده‌اند، ورودی‌های آن‌ها تحلیل حساسیت شده و ساختار بهینه آن‌ها به‌صورت خودکار استخراج می‌شود (شایان نژاد، ۲۰۰۶). به همین دلیل، پژوهش‌های متعددی به ارزیابی مدل‌های هوش مصنوعی در برآورد تبخیر تعرق پتانسیل پرداخته‌اند. در تحقیقی با استفاده از داده‌های مشاهده شده تشت کلاس A، کاربرد روش‌های شبکه‌های عصبی مصنوعی، سیستم استنتاج فازی عصبی و همچنین روش تجربی استیفان استوارت، جهت برآورد تبخیر روزانه کرمان مورد بررسی قرار

برای اجرای مدل‌های SVM و LSSVM استفاده شد. پارامترهای موردنیاز جهت تعیین تبخیر شامل روز، ماه، سال، درجه حرارت حداقل و حداکثر و میانگین، رطوبت نسبی حداقل، حداکثر و میانگین و ساعات آفتابی بود. جدول ۱، میانگین متغیرهای مورد استفاده برای مدل‌سازی تبخیر تعرق پتانسیل ماهانه در طول دوره آماری برای هر یک از ایستگاه‌ها را نشان می‌دهد. موقعیت ایستگاه‌های مورد مطالعه در شکل ۱، آورده شده است.

مناسبی برای برآورد تبخیر روزانه برخوردار هستند (Tezel and Buyukyildiz, 2015).

### مواد و روش‌ها

در این مطالعه، به منظور تعیین تبخیر از سطح آزاد آب، از پارامترهای هواشناسی سال‌های ۱۳۷۶ تا ۱۳۹۳ در سه ایستگاه گرگان، کلالة و بندر ترکمن واقع در استان گلستان



شکل (۱): موقعیت ایستگاه‌های مورد مطالعه

روزانه برداشت و برای سنجش دقت سایر روش‌ها استفاده شد.

مقادیر روزانه داده‌ها شامل دمای حداقل، میانگین و حداکثر، رطوبت نسبی کمینه، میانگین و بیشینه، ساعات آفتابی و سرعت باد است (جدول ۱). مقادیر تبخیر از تشت به صورت

گرگان	۱۸	۱۳	۲۳	۵۲	۸۷	۷۰	۲	۶
ن								
کل	۱۸	۱۲	۲۴	۴۸	۸۷	۶۸	۶	۱/۶
اله								
بندر ترکمن	۱۸	۱۴	۲۲	۶۰	۸۵	۷۲	۳/۴	۶/۳
من								

جدول (۱): میانگین پارامترهای هواشناسی ایستگاه‌های مورد مطالعه طی سال‌های ۱۳۷۳-۱۳۹۳

ایستگاه	میانگین	حداقل	حداکثر	رطوبت نسبی	رطوبت نسبی	سرعت باد	ساعات آفتابی
گرگان	۱۸	۱۲	۲۴	۴۸	۸۷	۶۸	۱/۶
ن							
کل	۱۸	۱۲	۲۴	۴۸	۸۷	۶۸	۱/۶
اله							
بندر ترکمن	۱۸	۱۴	۲۲	۶۰	۸۵	۷۲	۳/۴
من							



### ماشین بردار پشتیبان (SVM)

در سال‌های اخیر، ابزار پیشرفته‌ای که در رابطه باهوش مصنوعی کاربردهای زیادی در روش یادگیری ماشین دارد، ماشین بردار پشتیبان می‌باشد. استفاده از این روش موفقیت‌آمیز بوده است زیرا از آن در طبقه‌بندی اطلاعات و اخیراً در مسائل رگرسیونی استفاده شده است. اولین کاربرد این روش در مسائل آب را Dibike et al. (2001) 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، ۲۰۱۱).

ماشین بردار پشتیبان از جمله روش‌های نسبتاً جدیدی است که در سال‌های اخیر کارایی خوبی نسبت به روش‌های قدیمی‌تر برای دسته‌بندی و رگرسیون از جمله شبکه‌های عصبی پرسپترون چندلایه نشان داده است. از برتری‌های مدل ماشین بردار پشتیبان نسبت به سایر روش‌ها آن است که با داده‌های آموزشی کمتر و متغیرهای کمتری کار می‌کند اما نسبت به تغییر متغیرها بسته به داده‌های آموزشی حساس است. ماشین بردار پشتیبان (SVM) یکی از روش‌های یادگیری با ناظر است که از آن برای طبقه‌بندی و رگرسیون استفاده می‌شود. SVM توسط Vapnik (1998) معرفی شده است و بر پایه تئوری یادگیری آماری بنا نهاده شد. SVM روشی برای طبقه‌بندی دوتایی در فضای ویژگی‌های دلخواه است و از این روش مناسب برای مسائل پیش‌بینی بشمار می‌رود. در بیشتر مسائل جهان واقعی، تابع تقریب خطی استفاده‌های عملی محدودی دارد. Vapnik (1998) با استفاده از مفهوم حاصل ضرب داخلی نشان داد که ابتدا می‌توان بردار ورودی  $x$  را با یک تبدیل غیرخطی به یک فضای با بعد زیاد انتقال داد و در آن فضای حاصل ضرب داخلی را انجام داد و ثابت کرد که اگر یک هسته متقارن، شرایط قضیه مرسر را داشته باشد، اعمال این هسته در فضای ورودی با بعد کم می‌تواند حاصل ضرب را به شدت کاهش دهد. اگر بردار ورودی  $x_i$  به وسیله تابع غیرخطی  $\Phi(x)$  داخل فضای ویژه نگاشت شود تابع تصمیم به شکل زیر خواهد بود.

پارامترهای ورودی به مدل‌های مورد مطالعه الگوهای مختلف ورودی به مدل‌های SVM و LSSVM در جدول ۲ ارائه شده است  
جدول (۲): ترکیب‌های مختلف ورودی، برای تخمین مقدار تبخیر روزانه از تشت

نام الگو	پارامترهای ورودی	پارامترهای خروجی
۱	در هر الگو	در هر الگو
۲	EVA	$n-u_2$
۳	EVA	Tmean
۴	EVA	u-n
۵	EVA	Tmin-u
۶	EVA	Tmax-n
۷	EVA	Tmax-u
۸	EVA	Tmin-Tmax
۹	EVA	Tmean-n
۱۰	EVA	Tmin-n-u
۱۱	EVA	Tmax-n-u
۱۲	EVA	Tmin-Tmax-n
۱۳	EVA	Tmin-Tmax-u
۱۴	EVA	Tmin-n-u
۱۵	EVA	n-u
۱۶	EVA	Tmean-RHmean-n-u
		Tmin-Tmax-RHmin-n-u

EVA

تبخیر واقعی،  $n$ : تعداد ساعات آفتابی،  $T$  درجه حرارت (درجه سانتیگراد)،  $u$ : سرعت باد،  $RH_{min}$  رطوبت نسبی مینیمم،  $RH_{mean}$  رطوبت نسبی میانگین،  $T_{max}$  درجه حرارت بیشینه  
با توجه به آمار موجود، از ۷۰ درصد از داده‌ها برای آموزش و از ۳۰ درصد دیگر برای صحت سنجی مدل‌های ماشین بردار پشتیبان و حداقل مربعات ماشین بردار پشتیبان استفاده گردید.

$$0 \leq a_i \leq c, i = 1, 2, \dots, l$$

$$0 \leq a_i^* \leq c, i = 1, 2, \dots, l$$

اطلاعات کمی در مورد انتخاب تابع غیرخطی مناسب (x) در دسترس می‌باشد. بعلاوه، محاسبه  $\Phi(X_i)$  و  $\Phi(x_j)$  در فضای ویژه بسیار پیچیده می‌باشد. محاسبات در فضای ورودی با استفاده از تابع هسته‌ای  $\Phi$   $K(X_i, Y_i) = \Phi(X_i) \cdot \Phi(Y_i)$  انجام می‌شود تا تولید داخلی در فضای ورودی را ایجاد کند. توابعی معروف‌ترین توابع کرنل به صورت زیر می‌باشند (Yu et al, 2006).

$$\text{تابع کرنل گاوسی} (\Delta) = \exp\left(-\frac{\|u-v\|^2}{2\sigma^2}\right)$$

در نهایت، تابع تصمیم رگرسیون بردار پشتیبان غیرخطی به صورت معادله زیر خواهد بود (Yu et al., 2006).

$$f(x_i) = \sum_{i=1}^l (-a_i + a_i^*) k(x_i, x_j) + b \quad (6)$$

چنانچه w بردار ضرایب و b ثابت مشخصه‌های تابع رگرسیونی و  $\Phi$  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.

**ماشین بردار پشتیبان حداقل مربعات (LSSVM)**  
 Suykens and Vandewalle (1999) ماشین بردار پشتیبان با بهینه‌سازی حداقل مربعات (LSSVM) را معرفی کردند که فرمول‌بندی آن محدودیت مشابهی با SVM را بکار می‌برد. ویژگی اصلی این مدل این است که از نظر محاسباتی بهتر از SVM عمل می‌کند. در این حالت، آموزش نیاز به حل یک مجموعه توابع خطی بجای مسئله برنامه‌نویسی دوگانه که حل کلاسیک SVM می‌باشد، دارد (Khemchandani, 2009). این روش به طور مؤثری پیچیدگی الگوریتم را کاهش می‌دهد. درحالی‌که در روش SVM از بردارهای پشتیبان برای آموزش و حل مسئله رگرسیونی استفاده می‌شود، در روش LSSVM همه داده‌های آموزشی برای حل مسئله بهینه‌سازی و تولید نتایج استفاده می‌شود. در روش LSSVM داده‌های آموزشی به فضای هسته‌ای نگاشت می‌شوند و برای برقراری توازن بین

اولین کاربرد این روش در مسائل آب را Dibike و همکاران (2001) 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، ۲۰۱۱). 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 (خطای مجاز  $\epsilon$ ) مشخص شود. در مدل رگرسیون SVM تابعی مرتبط با متغیر وابسته y که خود تابعی از چند متغیر مستقل x برآورد می‌شود. مشابه سایر مسائل رگرسیونی فرض می‌شود رابطه میان متغیر مستقل و وابسته با تابع جبری مانند  $f(x)$  به علاوه مقداری اغتشاش (خطای مجاز  $\epsilon$ ) مشخص شود.

$$f(X) = W^T \times \varphi(x) + b \quad (1)$$

$$y = f(x) + \text{noise} \quad (2)$$

چنانچه W بردار ضرایب و b ثابت مشخصه‌های تابع رگرسیونی و  $\varphi$  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. فضای ویژه نگاشت شود، تابع تصمیم به شکل زیر خواهد بود.

$$f(W, X) = W \times \varphi(x) + b \quad (3)$$

مسئله رگرسیون غیرخطی می‌تواند مانند مسئله بهینه‌سازی بیان شود:

$$\text{Minimize } \frac{1}{2} \sum_{ij}^1 (a_i - a_i^*) (a_j - a_j^*) \times (\varphi(x_i) \cdot \varphi(x_j)) \quad (4)$$

$$+ \epsilon \sum_{i=1}^1 (a_i - a_i^*) - \sum_{i=1}^1 y_i (a_i - a_i^*)$$

$$\text{با محدودیت‌های } \sum_{i=1}^1 (a_i - a_i^*) = 0$$



می‌باشد و به صورت زیر توصیف می‌شود ( Yue et al., 2006).

(۱۰)

$$\text{minimize}_{\xi, w, b} \frac{c}{2} \|\xi\|^2 + \frac{1}{2} w^2$$

خواهد بود. ضمن اینکه برای ارزیابی بهتر، نتایج مدل‌ها نسبت به خط ۱:۱ نیز موردسنجش قرار گرفت.

$$R^2 = 1 - \frac{\sum_{i=1}^n (E_o - E_p)^2}{\sum_{i=1}^n (E_o - \bar{E}_o)^2} \quad (12)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (E_p - E_o)^2}{n}} \quad (8)$$

MBE =

$$\frac{1}{n} \sum_{i=1}^n \left( \frac{|E_o - E_p|}{\max(E_o) - \min(E_o)} \right) \times 100 \quad (9)$$

تعداد داده‌ها،  $E_o$  مقادیر تبخیر مشاهده شده،  $E_p$  مقادیر تبخیر برآورد شده است. هرچقدر مقادیر جذر میانگین مربعات خطا و میانگین کارایی خطابه صفر و ضریب همبستگی خطی به یک نزدیک‌تر باشد، مدل شبیه‌سازی را با دقت بیشتری انجام داده است.

## نتایج و بحث

شاخص‌های آماری مقدار تبخیر از سطح آزاد آب در مطالعه حاضر برای مدل‌سازی تبخیر مرجع ماهانه با داده‌های ۱۷ سال از فروردین ۱۳۷۶ تا شهریور ۱۳۸۹ برای آموزش و از شهریور ۱۳۸۹ تا ۱۳۹۳ برای صحت‌سنجی انتخاب شد. در مرحله بعد مقادیر بهینه مشخصه‌های مدل SVM و LSSVM تعیین شد.

شاخص‌های آماری محاسبه شده با مدل SVM برای

ایستگاه‌های مورد مطالعه با الگوهای ورودی مختلف

جدول ۳، معیارهای ارزیابی الگوهای مختلف را در ایستگاه گرگان، بندر ترکمن و کلاله را نشان می‌دهد. با توجه به جدول ۳، الگوی شماره ۱۶ در ایستگاه گرگان، بندر ترکمن و کلاله به ترتیب دارای بیشترین  $R^2$  ۰/۹۷، ۰/۹۷ و ۰/۹۸ و کمترین RMSE ۶/۷، ۶/۱۹ و ۵/۱ و MBE ۰/۲۵، ۰/۱۱ و

خطاهای آموزشی و تابع هموار از پارامترهای تنظیم استفاده می‌کند که برای همه نمونه‌ها یکسان است و می‌تواند به عنوان یک پیش‌فرض در نظر گرفته شود. روش LSSVM قادر به حل هر دو مسئله طبقه‌بندی و رگرسیون می‌باشد (Valyon and Horvath, 2005). رگرسیون ماشین بردار پشتیبان با حداقل مربعات شامل حل مسئله بهینه‌سازی زیر با محدودیت‌های

$$(W \cdot \varphi(x_i) + b) - y_i + \xi_i = 0$$

$$i = 1, 2, \dots, M$$

(۱۱)

که در آن  $w$  و  $b$  به ترتیب مقادیر وزن‌ها و اربیبی تابع رگرسیون هستند که از طریق حداقل سازی تابع هدف در رابطه زیر تعیین می‌گردند:

$$(Y) \min_{w, b} J(w, e) = \frac{1}{2} W^T W + \frac{\gamma}{2} \sum_{i=1}^n (e_i^2)$$

با محدودیت:  $y_i = W^t \sigma(x_i) + b + x_i$

که در آن‌ها  $\gamma$ : پارامتر تنظیم‌کننده بخش خطا است و  $e_i$ : خطای داده‌های آموزشی می‌باشد.

## معیارهای ارزیابی

شاخص‌های آماری ضریب تبیین  $R^2$ ، RMSE، MBE به منظور بررسی و ارزیابی دقت مدل‌ها استفاده شد. ضریب تبیین هم نشان‌دهنده میزان ارتباط مقادیر برآورد شده با مقادیر محاسبه شده است که هر چه میزان آن بیشتر باشد این ارتباط نزدیک‌تر است و اختلاف مقادیر برآورد شده و مقادیر محاسبه شده کمتر خواهد بود. شاخص RMSE- نشان‌دهنده عدم اطمینان مطلق مدل می‌باشد، این شاخص هر چه به سمت عدد صفر میل کند نشان از تخمین بهتر مدل خواهد بود (Singh and San, 2014). پارامتر ارزیابی MBE نشان می‌دهد که اولاً آیا مدل مقدار متغیر موردنظر را کم یا زیاد برآورد می‌کند و ثانیاً مقدار کمی آن چقدر است. زمانی که MBE برابر صفر است، مشخص می‌شود که مدل فضای مطالعه شده را خوب برآورد کرده و هیچ‌گونه انحرافی وجود ندارد. ضریب همبستگی هم نشان‌دهنده میزان ارتباط مقادیر برآورد شده با مقادیر محاسبه شده است که هر چه مقدار آن بیشتر باشد، این ارتباط نزدیک‌تر است و اختلاف مقدار برآورد شده با مقدار محاسبه شده کمتر

است. Kisi و Guven (2011) در تخمین میزان تبخیر از سطح آزاد آب به این نتیجه رسیدند که مدل SVM دارای دقت بالایی در پیش‌بینی تبخیر است. Tezal and Buyukyildiz (2015) نتایج تحقیق خود نشان داد که مدل ماشین بردار پشتیبان از دقت مناسبی برای برآورد تبخیر روزانه برخوردار هستند.

و ۰/۱۵ می‌باشد. الگوی شماره ۱۶ شامل پارامترهای حداقل و حداکثر دما، حداقل و حداکثر رطوبت نسبی، سرعت باد و ساعات آفتابی بوده است؛ بنابراین الگویی که دارای بیشترین پارامترهای ورودی بوده دارای دقت بیشتری در پیش‌بینی تبخیر از سطح آزاد آب بوده است. همچنین با توجه به جدول ۳، مدل SVM در ایستگاه کلاله دارای بهترین پیش‌بینی نسبت به دو ایستگاه دیگر بوده

جدول (۳): شاخص‌های آماری مقدار تبخیر از سطح آزاد آب با استفاده از مدل SVM

نام ایستگاه	شاخص	۱	۲	۳	۴	۵	۶	۷	۸	۹	۱۰	۱۱	۱۲	۱۳	۱۴	۱۵	۱۶
گرگان	RMSE	۶,۷	۷,۶	۷,۱	۶,۸	۷,۱۵	۷,۷	۷,۵	۶,۷	۶,۷	۶,۷	۶,۷	۱۳	۶,۹۴	۶,۸۴	۶,۷	۶,۷
	R <sup>2</sup>	۰,۹۶	۰,۹۶	۰,۹	۰,۶	۰,۹۷	۰,۹۷	۰,۸۵	۰,۹۰	۰,۹۰	۰,۹۰	۰,۹۰	۰,۹	۰,۹۰	۰,۹۰	۰,۹۰	۰,۹۷
ترکمن	MBE	۰,۲۵	۰,۳۶	۰,۴	۰,۷۸	۰,۷۷	۰,۳۵	۰,۳۸	۰,۲۷	۰,۳۴	۰,۴۴	۰,۳۳	۰,۳۲	۰,۵	۰,۵	۰,۲۷	۰,۲۵
	RMSE	۶,۷۱	۶,۲۱	۹,۹۷	۷,۶۵	۷,۶۵	۷,۰۲	۷,۲۸	۶,۷۷	۷,۲۶	۶,۹۶	۶,۸۵	۷,۰۶	۷,۱۱	۷,۱۰	۶,۷۹	۶,۱۹
کلاله	R <sup>2</sup>	۰,۹۵	۰,۹۶	۰,۹۶	۰,۹۶	۰,۹۶	۰,۹۶	۰,۹۶	۰,۹۵	۰,۹۶	۰,۹۶	۰,۹۶	۰,۹۶	۰,۹۶	۰,۹۶	۰,۹۶	۰,۹۷
	MBE	۰,۱۴	۰,۱۴	۰,۱۴	۰,۱۰	۰,۱۰	۰,۱۸	۰,۳۷	۰,۳۴	۰,۳	۰,۵۱	۰,۶۱	۰,۵۸	۰,۵۶	۰,۶۵	۰,۱۹	۰,۱۱
	RMSE	۶,۷۵	۶,۷۵	۸,۸۶	۶,۷۵	۶,۷۵	۶,۷۵	۶,۷۵	۶,۷۵	۶,۷۵	۶,۷۵	۶,۷۵	۶,۷۵	۶,۷۵	۶,۷۵	۶,۷۵	۵,۱
	R <sup>2</sup>	۰,۹۸	۰,۹۸	۰,۹۸	۰,۹۸	۰,۹۸	۰,۹۸	۰,۹۸	۰,۹۸	۰,۹۸	۰,۹۸	۰,۹۸	۰,۹۸	۰,۹۸	۰,۹۸	۰,۹۸	۰,۹۸
	MBE	۰,۱۶	۰,۱۷	۰,۱۸	۰,۱۶	۰,۱۳	۰,۱۶	۰,۱۷	۰,۱۶	۰,۱۶	۰,۱۶	۰,۱۶	۰,۱۶	۰,۱۰	۰,۱۶	۰,۱۶	۰,۱۵

ایستگاه دیگر بوده است. با توجه به جدول‌های ۳ و ۴، در همه ایستگاه‌های مورد مطالعه مدل LSSVM دارای R<sup>2</sup> بیشتر و RMSE و MBE کمتری نسبت به مدل SVM بوده است. با توجه به جدول ۳ و ۴، در ایستگاه‌های گرگان و بندر ترکمن، مدل LSSVM خطا را به ترتیب ۱۳ و ۱۹ درصد کاهش داده است. Wang و همکاران (2016) به مدل‌سازی تبخیر از تشت با استفاده از شش روش پرسپترون چندلایه، ژنتیک فازی، دو روش رگرسیون و حداقل مربعات ماشین بردار پشتیبان در اقلیم‌های مختلف چین پرداختند. نتایج حاصل از این آزمایش نشان داد که مدل حداقل مربعات ماشین بردار پشتیبان مورد مطالعه نسبت به روش‌های رگرسیون تبخیر تعلق را با دقت بیشتری تعیین کردند

### شاخص‌های آماری محاسبه شده با مدل LSSVM برای ایستگاه‌های مورد مطالعه با الگوهای ورودی مختلف.

با توجه به جدول ۳، الگوی شماره ۱۶ در ایستگاه گرگان، بندر ترکمن و کلاله به ترتیب دارای بیشترین R<sup>2</sup> ۰/۹۸، ۰/۹۸ و ۰/۹۸ و کمترین RMSE ۵/۸، ۵/۸، ۶/۱۶ و MBE ۰/۱۵، ۰/۱۳ و ۰/۱۵ می‌باشد. الگوی شماره ۱۶ شامل پارامترهای حداقل و حداکثر دما، حداقل و حداکثر رطوبت نسبی، سرعت باد و ساعات آفتابی بوده است؛ بنابراین الگویی که دارای بیشترین پارامترهای ورودی بوده دارای دقت بیشتری در پیش‌بینی تبخیر از سطح آزاد آب بوده است. همچنین با توجه به جدول ۴، مدل LSSVM در ایستگاه بندر ترکمن دارای بهترین پیش‌بینی نسبت به دو



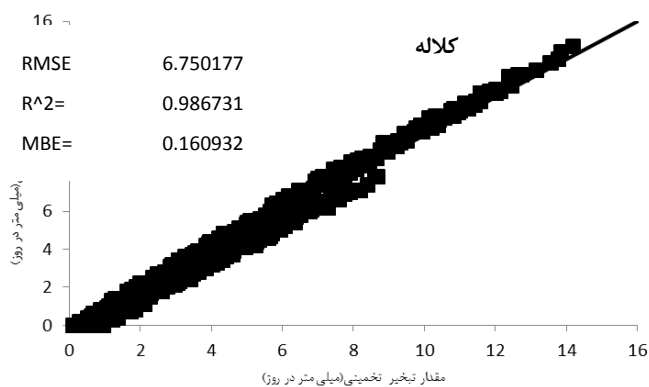
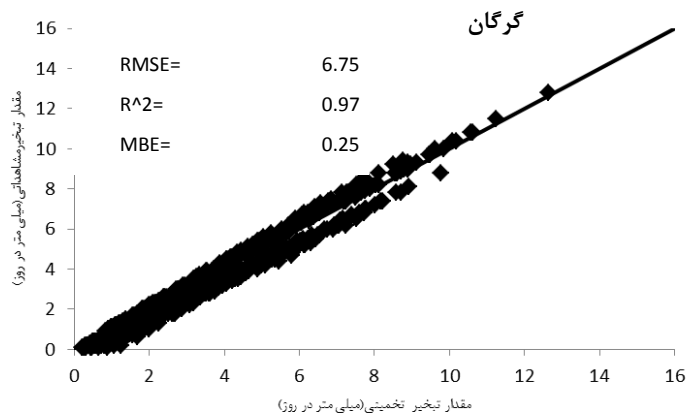
جدول (۴): شاخص‌های آماری مقدار تبخیر از سطح آزاد آب با استفاده از مدل LSSVM

نام ایستگاه گرگان	شاخص	۱	۲	۳	۴	۵	۶	۷	۸	۹	۱۰	۱۱	۱۲	۱۳	۱۴	۱۵	۱۶
ایستگاه گرگان	RMS E	۶,۴	۷,۷	۱۰,۷	۷,۳	۴,۰	۳,۷	۷,۵	۷,۹	۶,۷	۷,۶	۷,۵	۱۳,۴	۷,۸	۷,۳	۷,۶	۵,۸
	R <sup>2</sup>	۰,۹	۰,۹۸	۰,۹۸	۰,۹۸	۰,۹	۰,۹	۰,۹	۰,۹۸	۰,۹	۰,۹۸	۰,۹	۰,۹۷	۰,۹۸	۰,۹۰	۰,۹	۰,۹
	MBE	۰,۳	۰,۱۷	۰,۱۷	۰,۱۵	۰,۱	۰,۱	۰,۱	۰,۱۷	۰,۲	۰,۱۷	۰,۱۷	۰,۳۲	۰,۱۷	۰,۱۷	۰,۱	۰,۱
ترکمن	RMS E	۶,۷	۵,۷	۷,۸	۷,۷۵	۷,۴	۷,۵	۷,۶	۷,۷	۷,۸	۷,۶	۷,۷	۷,۷	۷,۸	۷,۸	۷,۹	۵,۰
	R <sup>2</sup>	۰,۹	۰,۹۸	۰,۹۷	۰,۹۷	۰,۹	۰,۹	۰,۹	۲۱,۰۶	۰,۹	۰,۹	۰,۹	۰,۹۷	۰,۹۷	۰,۹۷	۰,۹	۰,۹
	MBE	۰,۲	۰,۲۵	۰,۲۱	۰,۱۹	۰,۲	۰,۲	۰,۲	۰,۲۰	۰,۲	۰,۱	۰,۲۰	۰,۲۰	۰,۲۰	۰,۲۰	۰,۲	۰,۱
کلاله	RMS E	۱۰,۰	۷,۲۷	۱۳,۹	۱۰,۰	۶,۷	۹,۹	۶,۷	۱۰,۵۱	۹,۷	۱۰,۷۲	۱۰,۸۲	۱۰,۴۴	۱۰,۹	۱۰,۹	۱۰,۰	۶,۱
	R <sup>2</sup>	۰,۹	۰,۹۹	۰,۹۸	۰,۹۸	۰,۹	۰,۹	۰,۹	۰,۹۸	۰,۹	۰,۹۸	۰,۹۸	۰,۹۸	۰,۹۸	۰,۹۸	۰,۹	۰,۹
	MBE	۰,۱	۰,۰۲	-۰,۲	۰,۱۹	۰,۱	۰,۱	۰,۱	۰,۱۷	۰,۱	۰,۱	۰,۱۶	۰,۱۶	۰,۱۶	۰,۱۸	۰,۱	۰,۱

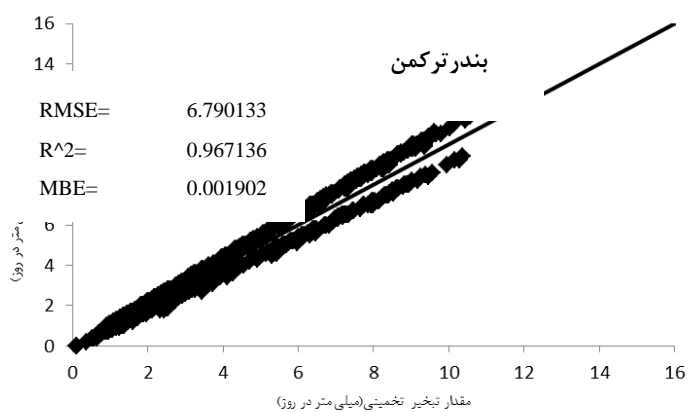
با توجه به شکل ۲، نمودار پراکندگی مقادیر مشاهداتی و تخمینی مدل SVM را در الگوی بهینه (الگوی شماره ۱۶) نشان می‌دهد. با توجه به نمودارهای رسم شده در هر سه ایستگاه، در ایستگاه کلاله بالاترین مقدار R2 و کمترین مقدار RMSE و MBE برای الگوی بهینه نسبت به دو ایستگاه مورد مطالعه بود

### نتایج پراکندگی مقادیر مشاهداتی و تخمینی در بهترین حالت مدل

با توجه به اینکه در سه ایستگاه مورد مطالعه، الگوی ۱۶ دارای بیشترین R2 و کمترین RMSE و MBE بود به‌عنوان الگوی بهینه انتخاب شد و نمودار پراکنش داده‌ها برای این الگو رسم شد.



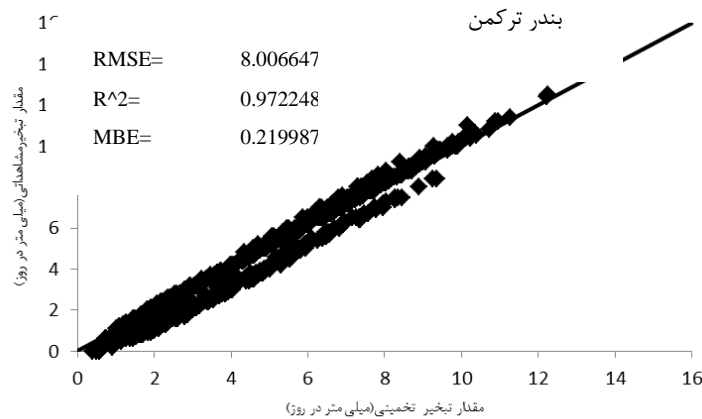
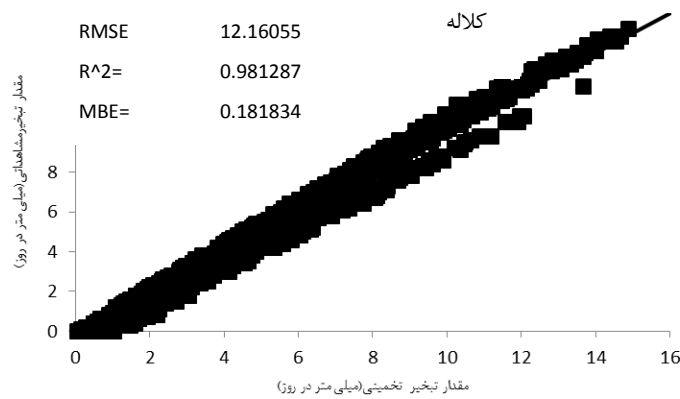
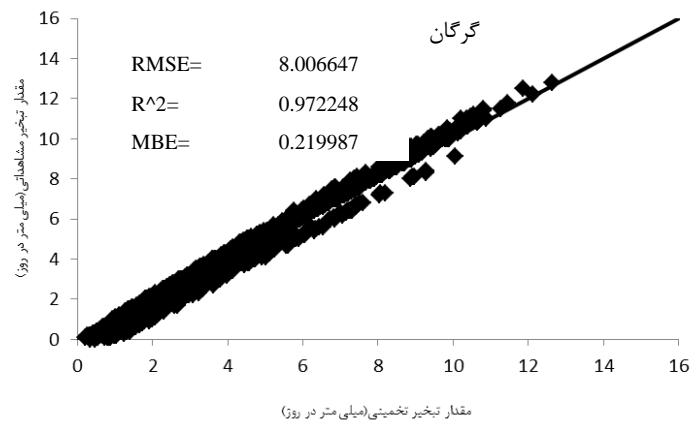
شکل (۲): نمودارهای پراکندگی مقادیر مشاهده‌ای و تخمینی در بهترین حالت مدل SVM



ادامه شکل (۲): نمودارهای پراکندگی مقادیر مشاهده‌ای و تخمینی در بهترین حالت مدل SVM

### پراکندگی مقادیر مشاهداتی و تخمینی در بهترین حالت مدل LSSVM

شکل ۳، نمودارهای الگوی بهینه را در سه ایستگاه گرگان، کلاله و بندر ترکمن برای مدل LSSVM نشان می‌دهد



شکل (۳): نمودارهای پراکندگی مقادیر مشاهداتی و تخمینی در بهترین حالت مدل LSSVM

هر دو مدل ماشین بردار پشتیبان SVM و حداقل مربعات ماشین بردار پشتیبان (LSSVM) در تعیین تبخیر روزانه از تشت دقیق و کارآمد ارزیابی شده است. همان طور که مشاهده می شود کلیه ایستگاهها دارای ضریب تبیین بالای ۰/۹۶ هستند که نشان دهنده عملکرد مناسب هر دو مدل SVM و LSSVM است. با توجه به شکل های ۲ و ۳، مدل LSSVM نتایج نزدیک تری در برآورد تبخیر از تشتک دارد. مقایسه مقادیر مشاهداتی و تبخیر حاصل از دو مدل نشان دهنده عملکرد بهتر مدل LSSVM نسبت به مدل SVM می باشد.

فایده اصلی مدل LSSVM این است که از نظر محاسباتی بهتر از SVM عمل می کند. این روش بطور موثری پیچیدگی های الگوریتم را کاهش می دهد. در حالی که در روش SVM از بردارهای پشتیبان برای آموزش و حل مسئله رگرسیونی استفاده می شود. در روش LSSVM همه داده های آموزشی برای حل مسئله بهینه سازی و تولید نتایج استفاده می شود.

۲. بین الگوهای ورودی به مدل SVM و LSSVM.

الگوی شماره ۱۶ دارای بهینه ترین برازش بوده است.

۳. نتایج ترکیب های الگوهای ۱ تا ۱۶ در تعیین تبخیر از سطح آزاد آب هر سه ایستگاه نشان داد که حذف یکی از پارامترها تأثیر چندانی در تعیین تبخیر ندارد؛ بنابراین، در صورت فقدان یکی از پارامترهای ورودی این تحقیق، تعیین تبخیر را با مشکل جدی روبه رو نمی کند.

### سپاسگزاری

این پژوهش به شماره ۶/۱۵۱۸ با حمایت مالی دانشگاه گنبد کاووس اجرا گردیده است. در اینجا لازم است از مدیریت محترم پژوهشی دانشگاه که شرایط مناسب برای انجام این تحقیق را فراهم نموده اند تشکر و قدردانی شود

نتایج ترکیب های الگوهای ۱ تا ۱۶ در تعیین تبخیر از سطح آزاد آب هر سه ایستگاه نشان داد که حذف یکی از پارامترها تأثیر چندانی در تعیین تبخیر ندارد؛ بنابراین، در صورت فقدان یکی از پارامترهای ورودی این تحقیق، تعیین تبخیر را با مشکل جدی روبه رو نمی کند. با توجه به نتایج این مطالعه، عملکرد مدلها بسیار نزدیک به یکدیگر بود؛ این نتیجه در مطالعه Tezel و Buyukyildiz (۲۰۱۵) نیز تأیید شده است. همچنین با توجه به نزدیک بودن نتایج، نمی توان یک ترکیب را به عنوان ترکیب مطلقاً برتر انتخاب کرد و این مسئله در ایستگاه های مختلف می تواند متغیر باشد؛ بنابراین، در هر ایستگاه بایستی به صورت جداگانه بهترین ترکیب و مدل برای تعیین تبخیر تعیین شود. مقایسه نتایج جدول های ۲ و ۳ نشان می دهد که میزان تبخیر را می توان با استفاده از یک ورودی نیز بسیار خوب تخمین زد و اضافه کردن سایر ورودیها در بهبود نتایج نقش کمتری دارد. یکی از مهم ترین دلایل این موضوع وجود همبستگی بین ورودیها با تبخیر از تشت ایستگاه های گرگان، کلاله و بندر ترکمن است. در حالت کلی، عملکرد

### نتیجه گیری

به دلیل اهمیت برآورد پدیده تبخیر و تعرق مرجع در سیستم برنامه ریزی و مدیریت آبیاری، در این پژوهش کارایی مدل های شبیه ساز SVM و LSSVM به عنوان سیستمی خبره برای پیش بینی این پدیده استفاده شد. کارایی این مدل ها در برآورد تبخیر از سطح آزاد آب در ایستگاه های گرگان، بندر ترکمن و کلاله مورد بررسی قرار گرفت و نتایج زیر به دست آمد:

۱. مدل LSSVM با ترکیب پارامترهای دمای

کمینه، دمای حداکثر، رطوبت نسبی حداقل، رطوبت نسبی حداکثر، سرعت باد و تعداد ساعات آفتابی در برآورد تبخیر ایستگاه های گرگان، کلاله و بندر ترکمن در مقایسه با مدل SVM مناسب تر است.

## منابع

- احمدی، ف.، س. آیشم، ک. خلیلی و ج. بهمنش. ۱۳۹۶. ارزیابی عملکرد شبکه عصبی مصنوعی (ANN) و ماشین بردار پشتیبان (SVM) در تخمین مقادیر روزانه تبخیر (مطالعه موردی: ایستگاه‌های هواشناسی تبریز و مراغه). نشریه آب‌و خاک (علوم و صنایع کشاورزی)، پژوهش‌های جغرافیای طبیعی. ۱(۴۹): ۱۵۱-۱۶۸.
- اسکندری، ع.، ر. نوری، ح. معراجی و ا. کیا قادری. ۱۳۹۱. توسعه مدلی مناسب بر مبنای شبکه عصبی مصنوعی و ماشین بردار پشتیبان برای پیش‌بینی بهنگام اکسیژن خواهی بیوشیمیایی ۵ روزه، محیط شناسی. ۳۸(۶۱): ۷۱-۸۲.
- رضایی، ا.، ع. خاشعی سیوکی و شهیدی، ع. ۱۳۹۳. طراحی شبکه پایش سطح آب زیرزمینی با استفاده از مدل حداقل مربعات ماشین بردار پشتیبان (LS-SVM)، تحقیقات آب‌و خاک ایران (علوم کشاورزی ایران). ۴۵(۴): ۳۸۹-۳۹۶.
- رضایی، ا.، ع. شهیدی و ع. خاشعی سیوکی. ۱۳۹۲. ارزیابی کارایی مدل حداقل مربعات ماشین بردار پشتیبان در پیش‌بینی سطح ایستابی (مطالعه موردی: دشت رامهرمز)، نشریه آبیاری و زهکشی ایران. ۴(۷): ۵۱۰-۵۲۰.
- زارع ابیانه، ح.، ح. نوری، ع. م. لیاقت، ح. نوری و ا. کریمی. ۱۳۹۰. مقایسه‌ی روش پنمن مانیتث فائو و تحت تبخیر کلاس A با داده‌های لایسیمتری در برآورد تبخیر و تعرق گیاه برنج در منطقه‌ی آمل، پژوهش‌های جغرافیای طبیعی. ۴۳(۷۶): ۷۱-۸۳.
- شادمانی، م و ص. معروفی. ۱۳۹۰. مقایسه چند روش برآورد تبخیر روزانه از تشتک مطالعه موردی منطقه کرمان. مجله علوم و فنون کشاورزی و منابع طبیعی، علوم آب‌و خاک. ۱۵(۵۵): ۶۹-۸۳.
- شایان نژاد، م. ۱۳۸۵. مقایسه روش‌های شبکه‌های عصبی مصنوعی و پنمن مانیتث در محاسبه تبخیر و تعرق پتانسیل، همایش ملی مدیریت شبکه‌های آبیاری و زهکشی، دانشگاه شهید چمران اهواز. ۵۳۷-۵۳۱.
- Allen, R. G., L. S. Preira., D. Raes and M. Smith, 1998. Crop evapotranspiration guidelines for computing crop water requirement. FAO Irrigation and Drainage Paper, NO.56.
- Asefa, T., M. W. Kemblowski. M. Mckee and A. Khalil. 2006. Multi-time scale stream flow prediction: The support vector machine approach. Hydrology. 318:7-16.
- Asefa, T., M. W. Kemblowski. M. Mckee and A. Khalil. 2004. Support vector-based groundwater head observation networks design. Water Resource Research. 40. W11509.
- Behzad, M., K. Asghari M. Eazi and M. Pallhang. 2009. Generalization performance of support vector machines and neural networks in runoff modeling. Expert Systems with Applications. 36: 7624-7629.
- Chang F. J., K. Y. Chang and L. C. Chang. 2008. Counter-propagation fuzzyneural network for city flood control system. J. Hydrol. 358: 24-34.
- Dibike, Y., S. Velickov, D. Solomatine and M. Abbott. 2001. Model induction with of support vector machines: Introduction and applications, J. Computing in Civil Engineering, 15(3): 208-216.
- Eslamian, S. S., J. Abedi-Koupai., M. J. Amiri and S. A. Gohari. 2009. Estimation of daily reference evapotranspiration using support vector machines and artificial neural networks in greenhouse. Environmental Sciences. 4: 439-447.
- Guo, x., x. Sun and x. Ma. 2011. Prediction of daily crop reference evapotranspiration value through a least- square support vector machine model. Hydrology Reserch. 42(4). 268- 274.
- Guyen, A and Kişi, O. 2011. Daily pan evaporation modeling using linear genetic programming technique, 29 (2):135-145.
- Kisi, O. and M. Cimen. 2010. Evapotranspiration modelling using support vector machines. Hydrological Sciences. 54 (5): 918-928.
- Kisi, O. and M. Zounemat-Kermani. 2014. Comparison of Two Different Adaptive Neuro-Fuzzy Inference Systems in Modelling Daily Reference Evapotranspiration. Water Resources Management, 28: 2655-2675.
- Khemchandani, R., S. Jayadeva and S. Chandra. 2009. Regularized least squares fuzzy support vector machine time series forecasting. Expert System with Application. 36: 132-138.
- Liong, SY. And C Sivapragasam. 2002. Flood stage forecasting with support vector machines. American Water Resource Association. 38:173-186.



- Liu, S., J. Bai., Z. Jia., L. Jia, H. Zhou, and L. Lu. 2010. Estimation of evapotranspiration in the Mu Us Sandland of China. *Hydrology and Earth System sciences*. 14: 573-584.
- Moghadamnia, A., M. Ghafari., J. Piri., and D. Han. 2008. Evaporation estimation using support vector machines technique. *Engineering and Technology*. 33:14-22.
- Pai, P., F. and W. C. Hong. 2007. A recurrent support vector regression model in rainfall forecasting. *Hydrological Process*. 21:819-827.
- Singh, H. and A. Sankarasubramanian. 2014. Systematic uncertainty reduction strategies for developing streamflow forecasts utilizing multiple climate models and hydrologic models. *Water Resources Research*. 50(2): 1288-1307.
- Suykens, J. A. K and J. Vandewalle .1999. *Least Squares Support Vector Machine Classifiers*. Kluwer Academic Publishers. Printed in the Netherlands.
- Tezel, G and M. Buyukyildiz. 2015. Monthly evaporation forecasting using artificial neural networks and support vector machines. *Theoretical and Applied Climatology*.
- Tezal, G and M. Buyukyildiz, M. 2015. Modeling of daily pan evaporation in subtropical climate ANN LS-SRV fuzzy logic and ANFIS. *Thero Appl Climatol*.
- Valyon, J. and Horvath, G. 2005. A robust LS-SVM regression. *World Academy of Science, Engineering and Technology*. 7:148-153.
- Vapnik, V.N. 1998. *Statistical Learning Theory*. Wiley. New York.
- Wang, L. Kisi, M. Zounemat-Kermani and H. Li. 2016. Pan evaporation modeling using six different heuristic computing methods in different climates of China.
- Yu, P. S., S. T. Chen and I. F. Chang. 2006. Support vector regression for real-time flood stage forecasting. *Hydrology*. 328: 704-716.



## Evaporation Modeling of Free Surface Water Using SVM and LSSVM Models

Masumeh Farasati<sup>۱</sup>, Morteza Seyedian<sup>۲</sup>, Kimia Daab<sup>۳</sup>

### Abstract

Evaporation is one of the most important and influential processes in the water cycle. Evaporation results in the loss of more than half of the precipitation in arid areas. Evaporation pan is used as an indicator for determining the evaporation of lakes and reservoirs due to the ease of interpreting its data around the world. On the other hand, the study on evaporation from the pan and the rate of evapotranspiration of the reference plant shows that there is a linear and direct relation between evaporation from the pan and evapotranspiration of the reference plant. Therefore, by correctly recording the amount of evaporation from the bath, the evapotranspiration of the reference plant can be estimated. The empirical relationships presented for estimating evaporation from free surfaces, considering meteorological parameters as inputs, are highly diverse. The accuracy of empirical relationships varies in different regions and needs calibration in each area. Also, it does not have high accuracy and access to all input parameters is difficult or time consuming. The aim of this study was to evaluate the efficiency of backup vector machine and least squares support vector machine for estimating evaporation from free water level in Golestan province. In this research, three synoptic stations (Kelaleh, Gorgan and Bandar-Turkman) were used for daily weather data for 17 years (1997-2015). The results showed that the input patterns with relative humidity input parameters, maximum relative humidity, wind speed and sunshine hours with the highest R<sup>2</sup> and the lowest RMSE and MBE.

**Keywords:** Evaporation, SVM model, LSSVM model, Pan

---

<sup>۱</sup>. Associate professor, Faculty of Agriculture, Gonbad University, Gonbad, Iran. Corresponding author, Email: Farasati2760@gmail.com

<sup>۲</sup>. Assistant professor, Faculty of Agriculture, Gonbad university, Gonbad, Iran. s.m.seyedian@gmail.com

<sup>۳</sup>. Undergraduate student, Faculty of Agriculture, Gonbad university, Gonbad, Iran. daab@gmail.com



## Evaporation Modeling of Free Surface Water Using SVM and LSSVM Models

Masumeh Farasati<sup>1</sup>, Morteza Seyedian<sup>2</sup>, Kimia Daab

### Introduction

Evapotranspiration is one of the main components of water balance in an area and is one of the effective factors for accurate irrigation planning. Therefore, accurate estimation of plant water needs has a significant effect on reducing the water crisis, especially in areas.

Accurate estimation of evapotranspiration plays an important role in studies such as global climate change, environmental evolution and control of water resources.

In this regard, intelligent models can be useful due to their ability to solve nonlinear and complex phenomena. These models include SVM and LSSVM.

### Methodology

In this study, in order to determine the evaporation from the free surface of water, meteorological parameters from 1997 to 2014 in three stations of Gorgan, Kalaleh and Bandar Turkmen located in Golestan province were used to implement SVM and LSSVM models. Parameters required determining evaporation included day, month, year, minimum and maximum and average temperature, minimum, maximum and average relative humidity and sunny hours.

According to available statistics, 70% of the data were used for training and the other 30% for validation of support vector machine models and minimum squares of support vector machine.

Statistical indices of coefficient of explanation  $R^2$ , RMSE, MBE were used to evaluate and evaluate the accuracy of the models. The coefficient of explanation also indicates the relationship between the estimated values and the calculated values. The higher the value, the closer the relationship is and the smaller the difference between the estimated values and the calculated values. RMSE index - indicates the absolute uncertainty of the model,

### Discussion and Conclusion

According to Tables, pattern number 16 in Gorgan, Turkmen port and Kalaleh stations have the highest  $R^2$  0.97, 0.97 and 0.98, respectively, and the lowest RMSE 6.7, 6.19 and 5.1 and MBE 0.25, 0.11 and 0.15. Pattern No. 16 included the parameters of minimum and maximum temperature, minimum and maximum relative humidity, wind speed and sunny hours; Therefore, the model with the most input parameters has been more accurate in predicting evaporation from the free surface of water. Also, the SVM model in Kalaleh station had the best forecast compared to the other two stations.

<sup>1</sup>. Associate professor, Faculty of Agriculture, Gonbad University, Gonbad, Iran. Corresponding author, Email: Farasati2760@gmail.com

<sup>2</sup>. Assistant professor, Faculty of Agriculture, Gonbad university, Gonbad, Iran. s.m.seyedian@gmail.com





In all stations studied, the LSSVM model had more  $R^2$  and less RMSE and MBE than the SVM model. According to Tables 3 and 4, in Gorgan and Bandar-e-Turkmen stations, the LSSVM model reduced the error by 13% and 19%, respectively.

Considering that in the three stations under study, model 16 had the highest  $R^2$  and the lowest RMSE and MBE was selected as the optimal model.

The results of combinations of patterns 1 to 16 in determining the evaporation from the free surface of water of all three stations showed that the omission of one of the parameters has little effect in determining the evaporation; Therefore, in the absence of one of the input parameters of this study, the determination of evaporation does not face a serious problem. According to the results of this study, the performance of the models was very close to each other.

The results showed that among the input models to SVM and LSSVM models, model 16 with the input parameters of minimum relative humidity, maximum relative humidity, wind speed and sunny hours had the highest  $R^2$  and the lowest RMSE and MBE. The LSSVM model at the Turkmen port station had the best forecast compared to the other two stations. Also, in all studied stations, LSSVM model had more  $R^2$  and less RMSE and MBE than SVM model.

The LSSVM model with the combination of minimum temperature, maximum temperature, minimum relative humidity, maximum relative humidity, wind speed and number of sunny hours in estimating the evaporation of Gorgan, Kalaleh and Turkmen port stations is more appropriate than the SVM model.

Between the input patterns to the SVM and LSSVM models, pattern 16 has the most optimal fit. The main advantage of the LSSVM model is that it performs computationally better than SVM. This method effectively reduces the complexity of the algorithm. While in SVM method, support vectors are used to teach and solve the regression problem. In the LSSVM method, all training data is used to solve the problem of optimization and production of results.

## The most important references

Pai, P., F. and W. C. Hong. 2007. A recurrent support vector regression model in rainfall forecasting. *Hydrological Process*. 21:819-827.

Singh, H. and A. Sankarasubramanian. 2014. Systematic uncertainty reduction strategies for developing streamflow forecasts utilizing multiple climate models and hydrologic models. *Water Resources Research*. 50(2): 1288-1307.

Suykens, J. A. K and J. Vandewalle .1999. *Least Squares Support Vector Machine Classifiers*. Kluwer Academic Publishers. Printed in the Netherlands.

Tezel, G and M. Buyukyildiz. 2015. Monthly evaporation forecasting using artificial neural networks and support vector machines. *Theoretical and Applied Climatology*.

Tezal, G and M. Buyukyildiz, M. 2015. Modeling of daily pan evaporation in subtropical climate ANN LS-SRV fuzzy logic and ANFIS. *Thero Appl Climatol*.