

مقاله پژوهشی:

نقش توابع کرنل در افزایش دقت پیش‌بینی تبخیر روزانه در دو منطقه مرطوب و خشک ایران

میلاذ شرفی^{۱*}، سعید صمدیان فرد^۲

۱- انشجوی کارشناسی ارشد آبیاری و زهکشی، دانشکده کشاورزی، دانشگاه ارومیه، ارومیه

۲- استادیار بخش گروه مهندسی آب، دانشکده کشاورزی، دانشگاه تبریز، تبریز

* نویسنده مسئول: miladsharafi1@gmail.com

تاریخ دریافت: ۱۴۰۰/۰۵/۳۱؛ تاریخ پذیرش: ۱۴۰۰/۰۷/۱۷

چکیده

تبخیر به عنوان یک عامل کلیدی در مطالعات هیدرولوژیکی، آب و هوایی، مدیریت آب کشاورزی، برنامه‌ریزی آبیاری و غیره در نظر گرفته می‌شود. تبخیر به دلیل فعل و انفعالات عوامل مختلف آب و هوایی، یک پدیده پیچیده و غیرخطی است. بنابراین، برای تخمین تبخیر باید از مدل‌های پیشرفته مانند معادلات تجربی و هوش مصنوعی استفاده کرد. در سال‌های اخیر، معادلات تجربی به طور گسترده برای تخمین تبخیر استفاده شده است. در این تحقیق عملکرد مدل‌های رگرسیون فرایند گاوسی (GPR) و رگرسیون ماشین بردار پشتیبان (SVR) در تخمین تبخیر روزانه دو ایستگاه آمل و بم، در بازه زمانی ۲۰۲۰-۲۰۱۶ ارزیابی شده است. داده‌های روزانه هواشناسی میانگین دما، رطوبت نسبی، ساعات آفتابی و سرعت باد، به عنوان ورودی مدل‌های SVR و GPR برای تخمین تبخیر روزانه استفاده شد. در مطالعه حاضر چهار سناریو ترکیبی از پارامترهای هواشناسی به منظور تخمین تبخیر بکار گرفته شدند. نتایج حاصل از مدل‌های مذکور نشان داد که هر دو مدل SVR و GPR عملکرد قابل قبولی در تخمین تبخیر دارند (ضریب همبستگی حدود ۰/۹۴). همچنین با توجه به ارزیابی‌های انجام‌شده، مشخص شد که مدل GPR عملکرد بهتری نسبت به مدل SVR داشته است (جذر میانگین مربعات خطا به ترتیب ۱/۵۶ و ۱/۶۲). در تحقیق حاضر از کرنل PUK به دلیل داشتن دقت بالا، بیشترین ضریب همبستگی و کمترین خطا (۰/۹۴ و ۰/۸۴) استفاده گردید.

واژه‌های کلیدی: تخمین تبخیر، رگرسیون فرایند گاوسی، رگرسیون ماشین بردار پشتیبان، آمل، بم.

مقدمه

استفاده از روش های شبیه سازی با بکارگیری الگوریتم های یادگیری ماشینی و داده کاوی ضروری است. به طور مثال (Kishi et al., 2016)، در مطالعه خود استفاده از داده کاوی و یادگیری ماشین را به علت کمبود داده های آماری در منطقه مورد مطالعه ضروری دانستند. همچنین، (Mosavi et al., 2017) و (Issazadeh et al., 2017) نیز به علت پیچیدگی تعیین پارامترهای مربوطه استفاده از دو روش مذکور را امری مهم تلقی کردند.

با در نظر گرفتن میزان بارش های شهرستان های آمل و بم در چند سال گذشته، توجه به نقش بارش و تأثیر آن بر محیط زیست (Hashemin nasab et al., 2013) و همچنین کاربرد موفق مدل های رگرسیون فرایند گاوسی (GPR) و ماشین بردار پشتیبان (SVR) در بررسی پدیده های غیرخطی باعث شد که در این تحقیق به پیش بینی مقادیر تبخیر در هر یک از ایستگاه های آمل و بم با استفاده از مدل های مذکور پرداخته شود. علت انتخاب این دو ایستگاه متفاوت بودن اقلیم آب و هوایی در دو منطقه مرطوب شمالی و خشک جنوبی می باشد. بنابراین در تحقیق حاضر با در نظر گرفتن چهار سناریو ترکیبی از داده های هواشناسی روزانه شامل دما، رطوبت نسبی، ساعات آفتابی و سرعت باد، مقادیر روزانه تبخیر مدل سازی شده و دقت هر یک از روش های GPR و SVR مورد ارزیابی قرار گرفت.

مواد و روش ها

منطقه مورد مطالعه

در این پژوهش از داده های هواشناسی ایستگاه های سینوپتیک آمل و بم، استفاده شده است. شهر آمل در جلگه مازندران و در دو سوی رود هراز با ارتفاع ۷۶ متر از سطح دریا در ۵۲ درجه و ۲۱ دقیقه طول شرقی و ۲۶ درجه و ۲۵ دقیقه عرض شمالی قرار دارد (شکل ۱). آب و هوای آمل معتدل و مرطوب بوده و میزان متوسط بارش سالیانه شهرستان آمل ۷۵۰ میلی متر برآورد شده است. (Najafi et al., 2009)

آب و تأمین آن از جمله موضوعاتی است که از چالش های عمده فراوری انسان در قرن حاضر است. با توجه به تغییرات آب و هوایی جهانی، محققان مطالعات زیادی در مورد تبخیر در سراسر جهان و ارزیابی آن برای شناسایی در چرخه هیدرولوژی انجام داده اند (Miralles et al., 2016؛ Ghazvinian et al., 2020). تبخیر یک جز مؤثر و قابل توجه در چرخه هیدرولوژیکی است و تخمین میزان تبخیر در درجه اول برای مدیریت منابع آب، ارزیابی برنامه آبیاری و همچنین مدل سازی کشاورزی ضروری است (Shiri, 2011؛ Gundalia et al., 2013؛ Fan et al., 2016؛ Feng et al., 2018). تبخیر نقش مهمی در مدیریت منابع آب منطقه، تغییرات آب و هوا و کشاورزی دارد (Wang et al., 2017). تبخیر تا حد زیادی به داده های هواشناسی مانند تابش خورشید، دمای هوا، سرعت باد، رطوبت نسبی و فشار اتمسفر بستگی دارد (Vicente et al., 2018a؛ Fan et al., 2018b؛ Fan et al., 2018).

پیش بینی دقیق تبخیر یک چالش اساسی در مدیریت منابع آب حوضه های آبخیز است و مدل سازی آن در مناطقی که توزیع مکانی یا توزیع زمانی مجدد کافی اندازه گیری نشده است، از اهمیت زیادی برخوردار است (Dalkiliç et al., 2014). تشت تبخیر A یکی از ابزارهای استاندارد و مستقیم اندازه گیری تبخیر است که در سرتاسر دنیا به علت سهولت در کاربرد به عنوان شاخصی در تعیین تبخیر از دریاچه ها و مخازن مورد استفاده قرار می گیرد. در بیشتر ایستگاه ها ثبت دقیق تبخیر به دلیل محدودیت های ابزاری و مشکلات نگهداری عملی نیست (Ghorbani et al., 2017). تبخیر فرآیندی غیرخطی، پیچیده و ناپایدار است، به همین دلیل امکان ایجاد یک رابطه ریاضی از جمله با استفاده از همه عوامل مرتبط غیرممکن است (Ghaemi et al., 2019؛ Kim et al., 2012). به دلیل پیچیدگی تعیین تمام پارامترهای مربوطه و کمبود اطلاعات آماری، مدل سازی سیستم های هیدرولوژیکی به طور کامل غیرممکن است. بنابراین،

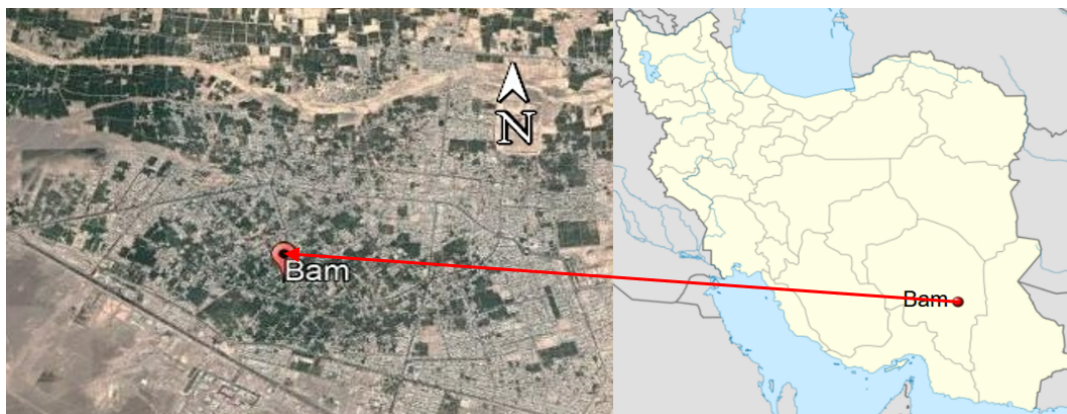


شکل ۱- موقعیت جغرافیایی شهرستان آمل در ایران

Figure 1 - Geographical location of Amol city in Iran

شرق استان کرمان واقع شده است (شکل ۲). آب و هوای بوم گرم و خشک است اما به علت مجاورت با کویر، آب و هوای متغیری دارد، به طوری که گاهی اوقات در تابستان‌ها گرم‌ترین و در زمستان‌ها سردترین نقطه کشور گزارش شده است.

شهرستان بوم از نظر جغرافیایی در ۵۸ درجه و ۲۱ دقیقه طول شرقی و ۲۹ درجه و ۶ دقیقه عرض شمالی و ارتفاع از سطح دریا ۱۰۵۰ متر و مساحت شهرستان بوم ۱۷/۷۵۵ کیلومتر مربع می‌باشد. شهر بوم در ۱۹۵ کیلومتری جنوب شرقی کرمان و در منطقه



شکل ۲- موقعیت جغرافیایی شهرستان بوم در ایران

Figure 2- Geographical location of Bam city in Iran

جدول ۱- پارامترهای آماری روزانه داده‌های مورد استفاده
Table 1- Daily statistical parameters of the data used

ایستگاه	پارامتر	واحد	میانگین	حداقل	حداکثر	انحراف از معیار	واریانس
آمل	دما (T)	°C	17/528	2/2	31/4	7/356	54/111
	رطوبت نسبی (RH)	%	81/382	49	97/158	6/609	43/679
	ساعات آفتابی (SSH)	hour	5.249	0	12/8	4/081	16/654
	سرعت باد (U2)	m/s	1/711	0	8/739	1/018	1/036

77/845	8/823	40/1	3/2	24/832	C'	دما (T)	بم
209/439	14/472	99/5	5/375	22/842	%	رطوبت نسبی (RH)	
6/615	2/572	12/4	0	8/654	hour	ساعات آفتابی (SSH)	
0/467	0/661	5/375	0/5	2/633	m/s	سرعت باد (U2)	

جدول 1- ۱ نشان دهنده میانگین، حداقل، حداکثر، انحراف از معیار و واریانس داده‌های مورد استفاده در این تحقیق می‌باشد. در این پژوهش از داده‌های روزانه هواشناسی ۵ سال آماری مربوط به سال‌های ۲۰۲۰-۲۰۱۶ استفاده شده است. متغیرهای هواشناسی روزانه شامل تبخیر از تشت (E)، میانگین دمای روزانه (T)، رطوبت نسبی (RH)، ساعات آفتابی (SSH) و سرعت باد (U₂) می‌باشد. در این پژوهش ۷۰٪ داده‌ها برای آموزش و ۳۰٪ برای آزمون در نظر گرفته شدند.

رگرسیون فرآیند گاوسی^۱ (GPR)

مدل‌های رگرسیون فرآیند گاوسی بر مبنای این فرض هستند که مشاهدات باید حامل اطلاعاتی درباره همدیگر باشد. این کار ترمیم طبیعی توزیع گاوسی است که میانگین و کواریانس آن به ترتیب بردار و ماتریس هستند (Yang et al., 2018). GPR به عنوان مجموعه‌ای از متغیرهای تصادفی تعریف شده است، که در آن تعدادی متغیر توزیع گاوسی چند متغیری دارند. X و Y به ترتیب دامنه ورودی و خروجی هستند. دامنه‌های n جفت (x_i, y_i) مستقل هستند و به طور مساوی استخراج و توزیع می‌شوند. فرض بر این است که فرآیند گاوسی بر روی X با میانگین عملکرد $\mu: Y \rightarrow Re$ و عملکرد کواریانس $k: X * X \rightarrow Re$ تعریف می‌شود. نشان دهنده هم‌ارز بودن عبارات می‌باشد. فرض اولیه GPR این است که y توسط $y = f(x) + \epsilon$ تعیین می‌شود، که در آن، ϵ نویز گاوسی با واریانس σ^2 است. در رگرسیون فرآیند گاوسی، برای هر متغیر ورودی X یک متغیر تصادفی f(x) وجود دارد که این مقدار تابع تصادفی f در آن مکان است.

$$Y = (Y_1, \dots, Y_n) \sim N(o, K = \sigma^2 I) \quad (1)$$

که در آن (I) ماتریس همانی است و $K_{ij} = k(x_i, x_j)$ ، از آنجا که $Y / X \sim N(o, K + \sigma^2 I)$ ، نرمال است، توزیع شرطی علامت‌های آزمایش با شرایط داده‌های آموزش و آزمایش $p(Y^*/Y, X, X^*)$ برابر است؛ در این شرایط $(Y^*/Y, X, X^*) \sim N(\mu, \sigma)$ ، بنابراین:

$$\mu = K(X, X_0)(K(X, X) + \sigma^2 I)^{-1} Y \quad (2)$$

$$\mu = K(X_0, X_0) - \sigma^2 I - (X, X_0)(K(X, X) + \sigma^2 I)^{-1} K(X, X_0) \quad (3)$$

$K(x, x')$ ماتریس $n \times n^*$ کواریانس ارزیابی شده در تمام جفت مجموعه داده‌های آزمون و آموزش است که برای سایر مقادیر $K(X, X)$ ، $K(X, X^*)$ و $K(X^*, X^*)$ که در آن X و Y بردار آموزش و برچسب داده‌های آموزشی y_i است، در حالی که X^* داده آزمون است. تابع کواریانس مشخص شده برای ایجاد یک ماتریس کواریانس مثبت نیمه نهایی k، که در آن $K_{ij} = k(K_{ij} = k(x_i, x_j))$ است، با تعیین هسته k و درجه نویز σ^2 ، معادله (۳) برای کسر کافی هستند. این تابع ساختار هندسی نمونه‌های آموزشی را در خود جای داده است. به عبارت دیگر، برای تولید پیش‌بینی‌های دقیق، توابع میانگین و کواریانس باید از داده‌های استفاده شده، که به آن‌ها ابر پارامتر گفته می‌شود، برآورد شود. (Pasolli et al., 2010)

در شکل ۳، مربعات نشان دهنده متغیرهای مشاهداتی و دوایر نمایانگر مجهولات می‌باشند که با توزیع y مشروط به مقادیر f بوده که با f یک گاوسین ایزوتروپیک به شکل زیر ارائه می‌شود:

¹ Gaussian process regression

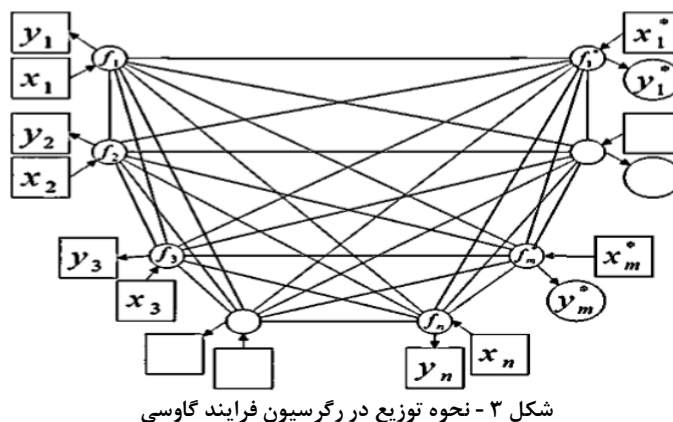


Figure 3- Distribution method in Gaussian process regression

تابع خطا در روش رگرسیون بردار پشتیبان به صورت گسترده در مسائل رگرسیونی کاربرد دارد، به صورت رابطه کتعیف می‌شود (Ahmadi et al., 2014):

$$\frac{1}{2}W^T W + C \sum_{i=1}^N \xi_i + C \sum_{i=1}^N \xi_i^* \quad (5)$$

معیارهای ارزیابی مدل

در این پژوهش برای ارزیابی عملکرد سناریوهای مختلف تعریف شده برای روش‌های GPR و SVR از پارامترهای آماری ضریب همبستگی (R)^۲، جذر میانگین مربعات خطا (RMSE)^۳ و میانگین خطای مطلق (MAE)^۴ استفاده شده است.

$$R = \frac{\left(\sum_{i=1}^n O_i P_i - \frac{1}{n} \sum_{i=1}^n O_i \sum_{i=1}^n P_i \right)}{\left(\sum_{i=1}^n O_i^2 - \frac{1}{n} \left(\sum_{i=1}^n O_i \right)^2 \right) \left(\sum_{i=1}^n P_i^2 - \frac{1}{n} \left(\sum_{i=1}^n P_i \right)^2 \right)} \quad (6)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (P_i - O_i)^2} \quad (7)$$

$$MAE = \frac{\sum_{i=1}^n |P_i - O_i|}{n} \quad (8)$$

در روابط ۶ تا ۸، n تعداد داده‌ها، O_i و P_i به ترتیب مقادیر مشاهداتی و پیش‌بینی شده تبخیر روزانه می‌باشند. به منظور بررسی امکان استفاده از ترکیب‌های متفاوتی از داده‌های هواشناسی در پیش‌بینی دقیق مقادیر تبخیر،

ماشین بردار پشتیبان (SVR)^۱

اولین کاربرد این روش در مسائل آب را برای شبیه سازی بارش رواناب ارائه کردند (Dibike et al., 2001). 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 کلی منجر می‌شود (Eskandari et al., 2011). در واقع می‌توان گفت که ماشین بردار پشتیبان سیستم یادگیری کارآمدی است که به منظور رسیدن به یک جواب بهینه از اصل استقرایی کمینه‌سازی خطای ساختاری استفاده می‌کند (Bilandi et al., 2014). مدل SVM^۴ یک الگوریتم محاسبات نرم تحت نظارت است که هم‌زمان خطاهای برآورد و پیچیدگی مدل را به حداقل می‌رساند (Benimam et al., 2020). برای مطالعه بیشتر به (Ahmadi et al., 2014) و (Hamel 2009) مراجعه گردد. در این روش در حالت غیرخطی مقادیر c برابر ۱ و اپسیلون برابر ۰/۰۰۱ می‌باشد. در یک مدل رگرسیونی ماشین بردار پشتیبان لازم است وابستگی تابعی متغیر وابسته y به مجموعه‌ای از متغیرهای مستقل x تخمین زده شود. فرض بر این است که مانند دیگر مسائل رگرسیونی، رابطه بین متغیرهای وابسته و مستقل توسط یک تابع معین f به علاوه یک مقدار اضافی نویز مشخص شود (Hamel, 2009):

$$y = f(x) + Noise \quad (4)$$

² Correlation coefficient

³ Root Mean Square Error

⁴ Mean absolute error

¹ support vector machine

1/6502 کاهش یافته است. همچنین در دیگر سناریوها نیز این نتیجه صادق است. لذا می‌توان نتیجه گرفت که مدل‌های مذکور برای مناطق خشک در پیش‌بینی مقادیر تبخیر نسبت به مناطق مرطوب قابلیت بالایی داشته باشند. همچنین سناریوهای شماره ۴ و ۲ به‌عنوان سناریوهای برتر انتخاب شدند. سناریو شماره ۱ تنها با پارامتر ورودی متوسط دما ضعیف‌ترین عملکرد را در هر دو مدل ارائه داد و می‌توان نتیجه گرفت که پارامتر دما به‌تنهایی برای برآورد میزان تبخیر شرط لازم ولی کافی نمی‌باشد. علاوه بر این، نتایج به دست آمده نشان داد که تابع کرنل پیرسون (PUK)^۱ در مقایسه با دیگر توابع کرنل، نقش عمده‌ای در افزایش دقت مدل‌های مورد مطالعه داشته است. همچنین، نمودارهای تغییرات زمانی مقادیر تبخیر پیش‌بینی شده با استفاده از روش‌های یادگیری ماشینی در مقایسه با نتایج حاصل از مشاهدات در مرحله صحت‌سنجی برای مدل‌های برتر در شکل ۴ قابل مشاهده می‌باشد.

چهار سناریو مختلف از متغیرهای هواشناسی دخیل در فرآیند تبخیر به‌عنوان ورودی مدل‌های SVR و GPR تعریف گردید (جدول ۲).

سپس با نرم‌افزار WEKA، محاسبات واسنجی بر روی ۷۰٪ از کل داده‌ها و محاسبات صحت‌سنجی هم بر روی ۳۰٪ از داده‌های باقیمانده انجام شد.

با تعریف چهار سناریو از پارامترهای هواشناسی در دو ایستگاه آمل و بم، محاسبات پیش‌بینی تبخیر با روش‌های SVR و GPR انجام گرفت. نتایج به‌دست آمده از هر یک از مدل‌ها با مقادیر مشاهداتی تبخیر مقایسه گردید و دقت روش‌های مذکور در مرحله صحت‌سنجی بر اساس معیارهای آماری ضریب همبستگی، جذر میانگین مربعات خطا و میانگین خطای مطلق محاسبه شده و در جدول ۳-۳ ارائه گردید. با توجه به نتایج ارائه شده در جدول ۳-۳ در ایستگاه آمل، سناریو چهارم مدل GPR با ضریب همبستگی ۰/۸۰۷۶، جذر میانگین مربعات خطای ۱/۱۸۷ و میانگین خطای مطلق ۰/۸۴۰۶، بهترین مدل در رگرسیون فرایند گاوسی شناخته شد. سناریو دوم مدل SVR نیز با دارا بودن ضریب همبستگی ۰/۸۰۴۷، جذر میانگین مربعات خطای ۱/۲۰۲۵ و میانگین خطای مطلق ۰/۸۳۰۱ رتبه دوم را در ایستگاه آمل به خود اختصاص داد. در ایستگاه بم، سناریو چهارم مدل GPR با ضریب همبستگی ۰/۹۴۲۷، جذر میانگین مربعات خطای ۱/۵۶۶۷ و میانگین خطای مطلق ۱/۱۸۹۹، بهترین مدل در رگرسیون فرایند گاوسی شناخته شد. سناریو چهارم مدل SVR نیز با دارا بودن ضریب همبستگی ۰/۹۳۵۹، جذر میانگین مربعات خطای ۱/۶۲۲۲ و میانگین خطای مطلق ۱/۲۱۱۲ رتبه دوم را در ایستگاه بم به خود اختصاص داد. نتایج این تحقیق نشان داد که عملکرد هر دو مدل SVR و GPR برای هر دو ایستگاه قابل ملاحظه و نسبتاً مشابه بوده (میانگین خطای مطلق به ترتیب ۰/۸۴۰۶، ۰/۸۷۷۱ و ضریب همبستگی ۰/۹۴۲۷، ۰/۹۳۵۹)، ولی با توجه به بررسی نتایج معیارهای ارزیابی مختلف، مدل GPR نسبت به مدل SVR عملکرد کمی بهتر داشته است.

در ایستگاه آمل مدل‌های بکار گرفته شده نسبت به ایستگاه بم عملکرد مناسب‌تری ارائه کردند. برای مثال در سناریو اول مدل SVR جذر میانگین مربعات خطا از 1/2625 به 1/6544 و در مدل GPR از 1/261 به

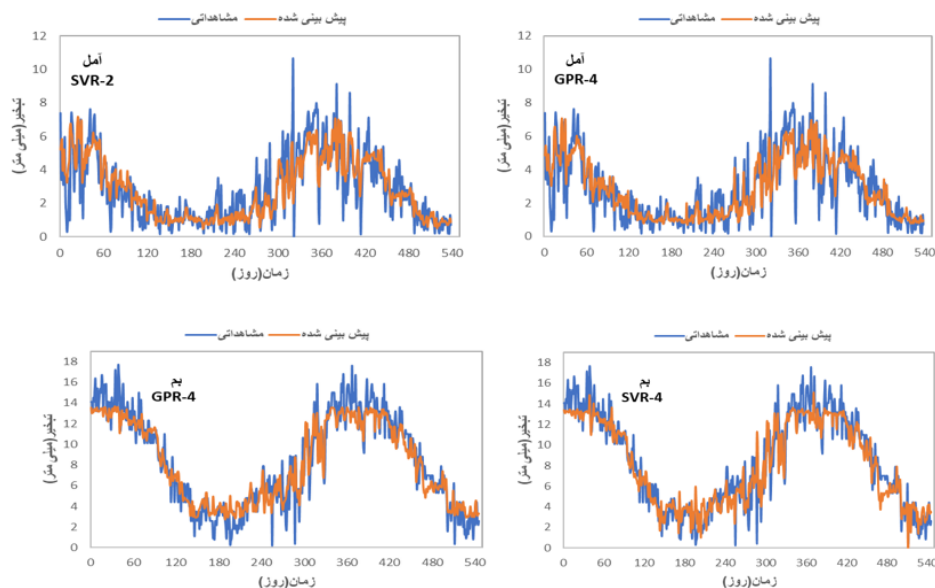
¹The Pearson VII function-based universal kernel

جدول 2- ترکیب‌های مختلف ورودی، برای تخمین مقدار تبخیر روزانه در مدل‌های SVM و GPR
 Table 2- Different input combinations to estimate daily evaporation in GPR and SVM models

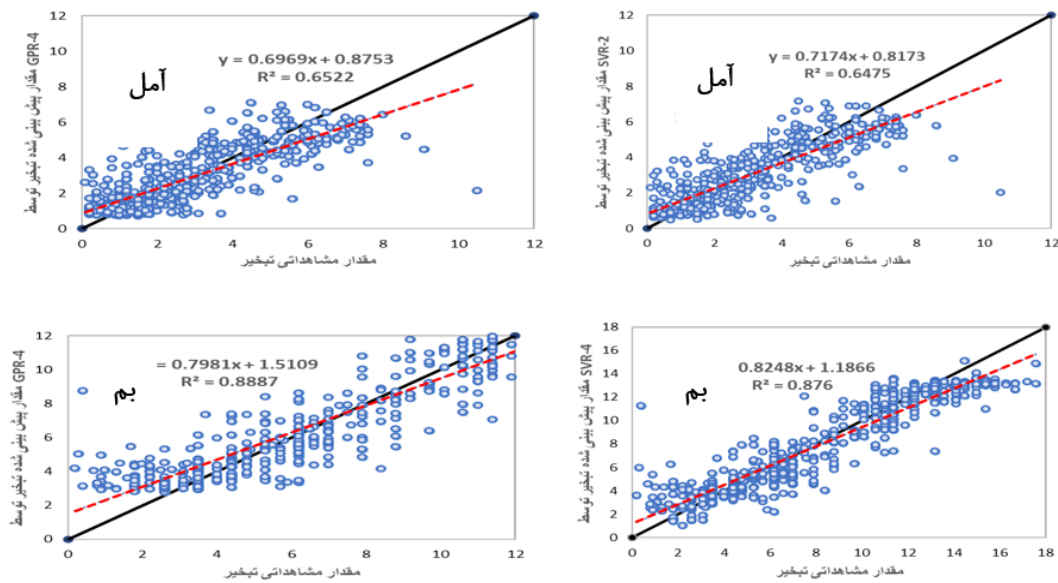
سناریو	پارامترهای ورودی	خروجی
1	T	E
2	RH	E
3	SSH	E
4	U2	E

جدول 3- نتایج معیارهای ارزیابی عملکرد دو مدل GPR و SVM در دو ایستگاه آمل و بجم
 Table 3- Statistical results of the computations for the defined scenarios for GPR and SVM models in Amol and Bam stations

ایستگاه	شماره سناریو	GPR			SVR		
		MAE	RMSE	R	MAE	RMSE	R
آمل	۱	۰/۸۹۵۹	۱/۲۶۱	۰/۷۸۱۳	۰/۸۹۰۸	۱/۲۶۲۵	۰/۷۸۱۷
	۲	۰/۸۳۴۷	۱/۱۹۶۴	۰/۸۰۴۵	۰/۸۳۰۱	۱/۲۰۲۵	۰/۸۰۴۷
	۳	۰/۸۴۱	۱/۲۰۴۲	۰/۸۰۱۳	۰/۸۵۹۲	۱/۲۳۱	۰/۷۹۵۹
	۴	۰/۸۴۰۶	۱/۱۸۷	۰/۸۰۷۶	۰/۸۷۷۱	۱/۲۲۱۱	۰/۷۹۹۲
بجم	۱	۱/۲۷۸۳	۱/۶۵۰۲	۰/۹۳۳۴	۱/۲۸۲۲	۱/۶۵۴۴	۰/۹۳۱۴
	۲	۱/۲۳۵۹	۱/۶۱۳۸	۰/۹۳۷۲	۱/۲۵۰۷	۱/۶۲۵۹	۰/۹۳۴۶
	۳	۱/۲۰۶۳	۱/۵۹۰۳	۰/۹۳۹۹	۱/۲۱۹	۱/۶۳۹۴	۰/۹۳۳۹
	۴	۱/۱۸۹۹	۱/۵۶۶۷	۰/۹۴۲۷	۱/۲۱۱۲	۱/۶۲۲۲	۰/۹۳۵۹



شکل ۴- نمودارهای تغییرات زمانی تبخیر روزانه با استفاده از مدل‌های GPR و SVR
 Figure 4- Diagrams of daily evaporation time changes using GPR and SVR models



شکل ۵- نمودارهای پراکنش تبخیر با استفاده از روش‌های GPR و SVR برای مدل‌های برتر
Figure 5- Evaporation distribution diagrams using GPR and SVR methods for superior models

پارامترهای آماری مورد مقایسه قرار گرفت و مشخص شد که در هر دو روش، مدل‌هایی که از همه پارامترهای هواشناسی (دمای روزانه، رطوبت نسبی، ساعات آفتابی و سرعت باد) به‌عنوان داده ورودی استفاده کردند، نتایج دقیق‌تری نسبت به سایر مدل‌ها ارائه دادند. همچنین نتیجه گرفته شد که در مقایسه مدل‌های برتر دو روش مورد مطالعه، در حالت کلی مدل GPR عملکرد بهتری نسبت به مدل SVR داشته است و از میان سناریوها مدل GPR4 بهترین عملکرد و پس از آن مدل SVR4 عملکرد قابل‌قبولی از خود نشان داده‌اند.

نتیجه‌گیری

تبخیر یکی از فرآیندهای مهم و مؤثری است که سبب از بین رفتن بیش از نیمی از نزولات جوی در مناطق خشک و نیمه خشک می‌شود. لذا در این پژوهش با استفاده از روش‌های GPR و SVR، در دو ایستگاه آمل و بم مقادیر تبخیر روزانه در طی دوره آماری ۲۰۲۰-۲۰۱۶ برآورد گردید. ورودی روش‌های مذکور با داده‌های هواشناسی میانگین دمای روزانه (T)، رطوبت نسبی (RH)، ساعات آفتابی (SSH) و سرعت باد (U) در چهار سناریو متفاوت بود. نتایج به‌دست آمده با استفاده از

مراجع

- 1- Ghazvinian H, Karami H, Farzin S, Mousavi SF. Effect of MDF-Cover for Water Reservoir Evaporation Reduction, Experimental, and Soft Computing Approaches. J Soft Comput Civ Eng 2020;4:98-110.
- 2- Miralles DG, Jiménez C, Jung M, Michel D, Ershadi A, McCabe MF, et al. The WACMOS-ET project – Part 2:
- 3- Evaluation of global terrestrial evaporation data sets. Hydrol Earth Syst Sci 2016;20:823-42.
- 4- Shiri J. Application of Artificial Intelligence to Estimate Daily Pan Evaporation Using Available and Estimated Climatic Data in the Khozestan Province (South Western Iran). Journal of Irrigation & Drainage Engineering, 2011, 137(137):412-425.
- 5- Gundalia M J, Dholakia M B. Estimation of Pan Evaporation Using Mean Air Temperature and, Radiation for Monsoon Season in Junagadh Region. International Journal of Engineering Research & Applications, 2013, 3(6): 64-70.
- 6- Fan J, Wu L, Zhang F, et al. Climate change effects on reference crop evapotranspiration across different climatic zones of China during 1956-2015. Journal of Hydrology, 2016, 542, 923-937.
- 7- Feng Y, Jia Y, Zhang Q, et al. National-scale assessment of pan evaporation models across different climatic zones of



China. *Journal of Hydrology*, 2018 ,564: 314-328

8- Wang L, Niu Z, Kisi O, Li C, Yu D. Pan evaporation modeling using four different heuristic approaches. *Comput Electron Agric* 2017;140:203–13.

9- VicenteSerrano S M, Bidegain M, TomasBurguera M, et al. A comparison of temporal variability of observed and model-based pan evaporation over Uruguay (1973–2014). *International Journal of Climatology*, 2018, 38(1): 337-350.

10- Fan J, Chen B, Wu L, et al. Evaluation and development of temperature-based empirical models for estimating daily global solar radiation in humid regions .*Energy*, 2018a, 144: 903-914.

11- Fan J, Wang X, Wu L, et al. New combined models for estimating daily global solar radiation based on sunshine duration in humid regions: A case study in South China. *Energy Conversion and Management*, 2018b, 156: 618-625.

12- Dalkiliç, Y., Okkan, U., & Baykan, N. (2014). Comparison of different Ann approaches in daily pan evaporation prediction. *Journal of Water Resource and Protection*, 6, 319– 326.

13- Ghorbani, M. A., Deo, R. C., Yaseen, Z. M., Kashani, M. H., & Mohammadi, B. (2018). Pan evaporation prediction using a hybrid multilayer perceptron-firefly algorithm (MLP-FFA) model: case study in North Iran. *Theoretical and applied climatology*, 133(3), 1119-1131

14- Ghaemi,A.;RezaieBalf,M.;Adamowski,J.;Kisi,O.;Quilty,J.On the applicability of maximum overlap discrete wavelet transform integrated with MARS and M5 model tree for monthly pan evaporation prediction. *Agric. For. Meteorol.* 2019, 278.

15- Kim, S.; Shiri, J.; Kisi, O. Pan Evaporation Modeling Using Neural Computing Approach for Different Climatic Zones. *Water Resour. Manag.* 2012, 26, 3231–3249.

16- Kisi, O., Genc, O., Dinc, S., & Zounemat-Kermani, M. (2016). Daily pan evaporation modeling using chi-squared automatic interaction detector, neural networks, classification and regression tree. *Computers and Electronics in Agriculture*, 122, 112-117.

17- Mosavi,A.,Bathla,Y.,&VarkonyiKocz y,A.(2017).Predictio the future using web knowledge:Of the art survey.*Advances in Intelligent Systems and Computing*, Springer Nature660, 341–349.

18- Issazadeh M, Shirzad M, Rezaei M, (2017). Evaluating the performance of artificial neural network and support vector machine in estimating daily evaporation values (Case study: Tabriz and Maragheh meteorological stations). *Natural Geography Research*, Volume 49, Samara 1, pp. 151-168

19- F. Hashemin nasab, M. Mousavi baygi, B. Bakhtiari, K. Davari, (2013). Predicting Rainfall Changes by Reducing the Scale of LARS-WG and HadCM3 Models in Kerman during the Next 20 Years (2011-2030), *Iranian Journal of Irrigation and Water Engineering*, Second Year, No. 12, pp. 43-58.

20- Najafi A, Matiei H, Najafi K. Feasibility study of eco-tourism development in rural areas using SWOT analytical model, (2009) Case study: Amol County *Journal of ScienceResearch of Iranian Geographical Society*, New Era, 6th, pp. 18 & 19.

21- Yang,D., Zhang,X., Pan R., Wang,Y., and Chen, Z. 2018 A novel Gaussian process regression model for state-of-health estimation of lithium-ion battery using charging curve. *Journal of Power Sources*. 384: 387–395.

22- Pasolli, L., Melgani, F., & Blanzieri, E. (2010). Gaussian process regression for estimating chlorophyll concentration in subsurface waters from remote sensing data. *IEEE Geoscience and Remote Sensing Letters*, 7(3), 464-468.

23- 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.

24- Eskandari, A . Nouri, R . Mearaji, H. And Kiaghadi, A. (2011). Development of a suitable model based on artificial neural network and support vector machine for biochemical oxygen demand forecasting, *fasting 5, ecology 38 (61)*, pp. 71-82.

25- Bilandi, M., Khashei siouki, E. and Sadeghi Tabas, S. “Daily Runoff Prediction with Minimal Squares Model of Support Vector Machine (Is-Svm15) *Journal of Soil and Water Conservation Research*, 2014, No. 6, pp 293-3,4.

26- Benimam, H., Si-Moussa, C., Laidi, M., Hanini, S., (2020)b. Modeling the activity coefficient at infinite dilution of water in ionic liquids using artificial neural networks and support vector machines. *Neural Comput. Appl.* 32 (12), 8635–8653.

27- Hamel,L. 2009. *Knowledge Discovery with Support Vector Machines*. Hoboken. N.J. John Wiley.

28- Ahmadi F, Radmanesh F., Mir Abbasi, “Comparison of Genetic Planning Methods and Support Vector Machines in Predicting Daily River Flow (Case Study: Barandozchay River)”, *Journal of Water and Soil (Crop Science and Technology)*, No. 6, 2014, pp. 1162-1171



Original Article:

The role of kernel functions in increasing the accuracy of daily evaporation prediction in wet and dry regions of Iran

Milad Sharafi^{1*}, Saeed Samadianfard²

1-Master student of Irrigation and Drainage, Faculty of Agriculture, Urmia University, Urmia

2-Assistant Professor, Department of Water Engineering, Faculty of Agriculture, University of Tabriz, Tabriz

*Corresponding Author E-mail: miladsharafi1@gmail.com

Received: 22-08-2021; Accepted: 09-10-2021

Abstract

Evaporation is considered as a key factor in hydrological, climatic, agricultural water management, irrigation planning, etc. studies. Evaporation is a complex and nonlinear phenomenon due to the interactions of various climatic factors. Therefore, advanced models such as experimental equations and artificial intelligence should be used to estimate evaporation. In recent years, experimental equations have been widely used to estimate evaporation. In this study, the performance of Gaussian process regression (GPR) and support vector machine regression (SVR) models in estimating the daily evaporation of Amol and Bam stations in the period 2020-2016 has been evaluated. Daily meteorological data on mean temperature, relative humidity, sunshine hours and wind speed were used as input of GPR and SVR models to estimate daily evaporation. In the present study, four combined scenarios of meteorological parameters were used to estimate evaporation. The results of the mentioned models showed that both GPR and SVR models have acceptable performance in estimating evaporation (correlation coefficient about 0.94). Also, according to the evaluations, it was found that the GPR model had a better performance than the SVR model (root mean square error of 1.56 and 1.62, respectively). In the present study, the PUK kernel was used due to its high accuracy, highest correlation coefficient and lowest error (0.94 and 0.84).

Keywords: Evaporation estimation, Gaussian process regression, Support vector machine regression, Amol, Bam.