

پیش‌بینی تبخیر از تشت با استفاده از مدل‌های محاسبات نرم در اقلیم فراخشک (مطالعه موردی: دشت سیستان، سیستان و بلوچستان)

مجتبی محمدی^{۱*}، معصومه فروزان‌فرد^۲، حمید غلامی^۳

تاریخ دریافت: ۱۴۰۱/۳/۲۰

تاریخ پذیرش: ۱۴۰۱/۱۰/۲۱

چکیده

پیش‌بینی بزرگی و الگوی تبخیر به‌ویژه برای پهنه‌هایی با اقلیم خشک، نیمه‌خشک و فراخشک همانند دشت سیستان ضروری است. از این‌رو در این پژوهش برای پیش‌بینی تبخیر تشت دشت سیستان از هشت مدل محاسبات نرم استفاده شد. بدین منظور، از پارامترهای متفاوت اقلیمی در دو ایستگاه سینوپتیک زابل و زهک و ایستگاه تبخیرسنجی چاه‌نیمه‌ها در پیکره هشت سناریوی ترکیبی برای ورودی مدل‌ها استفاده شد. به‌منظور ارزیابی مدل‌های مورد استفاده در این پژوهش از معیارهای ارزیابی ضریب تعیین، ریشه میانگین مربعات خطا، میانگین خطای مطلق و همچنین نمودار تیلور استفاده شد. یافته‌ها نشان داد که مدل با ورودی فقط دمای کمینه ماهانه عملکردی همانند مدل با هشت ورودی پارامتر هواشناسی دارد. همچنین یافته‌ها نشان داد که در بین هشت مدل، مدل جنگل تصادفی در پنج سناریو (سناریو ۲، ۵، ۶، ۷ و ۸ با مقادیر R^2 به ترتیب برابر ۰/۹۷، ۰/۹۸، ۰/۹۸ و ۰/۹۸) از هشت سناریو به‌عنوان بهترین مدل عمل کرده است. از این‌رو مدل جنگل تصادفی و مدل درختی به‌عنوان بهترین مدل محاسبات نرم به‌منظور شبیه‌سازی و برآورد تبخیر تشت دشت سیستان معرفی شدند.

کلیدواژه‌ها: پیش‌بینی، تبخیر، محاسبات نرم، اقلیم فراخشک، پهنه سیستان.

۱. استادیار گروه مدیریت و کنترل بیابان، دانشکده منابع طبیعی، مجتمع آموزش عالی سراوان، سراوان، نویسنده مسئول، Mohamadi613@gmail.com

۲. کارشناس ارشد مرتع‌داری، مجتمع آموزش عالی سراوان

۳. دانشیار گروه منابع طبیعی، دانشکده کشاورزی و منابع طبیعی، دانشگاه هرمزگان، بندرعباس

* این مقاله برگرفته از پژوهش مستقل بوده است.

مقدمه

(MARS) و رگرسیون درختی (M5Tree) (کیسی^۸، ۲۰۱۹)، مدل برنامه‌ریزی بیان ژن (GEP) (ماتر^۹، ۲۰۱۸)، هوش مصنوعی (کامل^{۱۰} و همکاران، ۲۰۲۱)، مدل الگوریتم نزدیک‌ترین همسایه (KNN) (فنگ و تیانگ^{۱۱}، ۲۰۲۱)، مدل‌های رگرسیون بردار پشتیبان و مدل درختی M5 (صمدیان‌فرد و پناهی، ۲۰۱۹؛ میرهاشمی و همکاران، ۲۰۲۰)، روش‌های مبتنی بر شبکه‌های عصبی مصنوعی (سپهری و همکاران، ۲۰۲۱؛ سیاسر و هنر^{۱۲}، ۲۰۲۰) و مدل‌های ماشین بردار پشتیبان و حداقل مربعات ماشین بردار پشتیبان (فراست و همکاران، ۲۰۲۱) در تخمین تبخیر و تبخیر و تعرق مورد استفاده قرار گرفته‌اند. بررسی پژوهش‌های اخیر در زمینه پیش‌بینی تبخیر/تعرق و اشکار ساخت که پژوهشگران تمرکز خود را بر روی دستیابی به مدل‌های پیش‌بینی بهتر، قوی‌تر و قابل اعتماد گذاشته‌اند. از این رو مطالعات صورت‌گرفته نشان داد که روش جدید محاسبات نرم برای توسعه و پیاده‌سازی مدل‌های پیش‌بینی ETo، خطاهای پیش‌بینی را کاهش داده‌اند (کایا^{۱۳} و همکاران، ۲۰۲۱؛ سیاحی و همکاران، ۲۰۲۱؛ شیری و همکاران، ۲۰۲۹). بررسی مطالعات در ایران نیز روشن ساخت که مطالعات اندکی در این خصوص انجام شده است. از آنجاکه پهنه وسیعی از سرزمین ایران در مناطق خشک و نیمه‌خشک قرار دارند، پژوهش در زمینه روش‌های جدید برآورد تبخیر و تعرق از اهمیت افزون‌تری برخوردار است. در این میان، پهنه سیستان که در شمال استان سیستان و بلوچستان در شرق ایران قرار گرفته، دارای اقلیم خشک و فراخشک است و کمیت تبخیر، مهم‌ترین و ناشناخته‌ترین مؤلفه بیلان آبی در این پهنه از کشور است (دانشکار آراسته و همکاران، ۲۰۰۵). از سویی دیگر، مخزن چاه‌نیمه‌ها با گنجایش تقریبی یک و نیم میلیارد مترمکعب، تنها منبع آب شرب شهرستان‌های زابل، زهک و زاهدان و سایر نواحی مسکونی سیستان است و سالانه حجم بسیاری از آن تبخیر

تبخیر به‌عنوان فرایند کلیدی در چرخه هیدرولوژیک، بر برنامه‌ریزی و مدیریت منابع آب تأثیر می‌گذارد (پرستلی و تیلور^۱، ۱۹۷۲). از این رو پیش‌بینی میزان و الگوی تبخیر بسیار مهم است (قاسم^۲ و همکاران، ۲۰۱۹). میزان تبخیر آب را می‌توان از طریق متغیرهای مختلف با استفاده از یک روش مستقیم یا غیرمستقیم برآورد کرد (موران^۳ و همکاران، ۱۹۹۶). در بین روش‌های مستقیم، اندازه‌گیری تبخیر از تشتک کلاس A و برآورد تبخیر و تعرق با استفاده از لایسیمتر رایج‌ترین روش‌ها هستند (استن‌هیل^۴، ۲۰۰۲). این روش برآورد دقیقی از تبخیر/تبخیر و تعرق در طول زمان ارائه می‌دهد (قربانی و همکاران، ۲۰۱۷). در روش‌های برآورد غیرمستقیم تبخیر و تعرق از متغیرهای هواشناسی همچون سرعت باد، ساعات آفتابی، بارش، حداکثر و حداقل درجه حرارت و رطوبت نسبی استفاده می‌شود (سنیخوانی و همکاران، ۲۰۱۸). با وجود این، این شکل از تخمین تبخیر و تعرق گیاه مرجع (ETo^۵) به دلیل ماهیت پویای پارامترهای پیش‌بینی آن، که برآمده از ویژگی‌های غیرخطی، غیرثابت و تصادفی آن‌هاست، پیچیده است (ماجهی^۶ و همکاران، ۲۰۱۹). از این رو توسعه مدل‌های پیش‌بینی هوشمند برای تبخیر و تعرق که قابل اعتماد و قوی باشند، به کانون اصلی مهندسان و دانشمندان منابع آب تبدیل شده است. در حال حاضر، توسعه و پیاده‌سازی مدل‌های مبتنی بر محاسبات نرم^۷ به‌منظور پیش‌بینی و حل مسائل پیچیده در رشته‌های مختلف پدیدار شده‌اند. این روش به دلیل کم‌هزینه بودن و عملکرد عالی آن‌ها (پناهی و همکاران، ۲۰۲۰)، روزبه‌روز در حال توسعه و گسترش در پژوهش‌ها از جمله پژوهش در برآورد تبخیر و تعرق است. به‌طور کلی، مطالعات متعددی از مدل‌های مختلف مانند حداقل مربعات ماشین بردار پشتیبان (LSSVM)، مدل اسپلاین چندمتغیره رگرسیون انطباقی

8. Kisi
9. Matter
10. Kamel
11. Feng and Tian
12. Siasar and Honar
13. Kaya

1. Priestley and Taylor
2. Qasem
3. Moran
4. Stanhill
5. Reference Evapotranspiration
6. Majhi
7. Soft computing

مواد و روش‌ها

منطقه مورد مطالعه

دشت سیستان در شمال استان سیستان و بلوچستان و در شرق ایران با محدوده جغرافیایی ۳۰ درجه و ۱۸ دقیقه تا ۳۱ درجه و ۲۰ دقیقه عرض شمالی و ۶۱ درجه و ۱۰ دقیقه تا ۶۱ درجه و ۵۰ دقیقه طول شمالی واقع شده است. این دشت با ارتفاع ۴۷۰ تا ۵۰۰ متر از سطح دریا، متوسط بارش سالانه ۵۹/۶ میلی‌متر و دمای ۹/۵- تا ۴۹ درجه سانتی‌گراد بر اساس شاخص خشکی دومارتن در طبقه‌بندی فراخشک قرار می‌گیرد.

داده‌های مورد استفاده

در این پژوهش از داده‌های ایستگاه سینوپتیک زابل و زهک به‌عنوان ورودی مدل و از داده‌های اندازه‌گیری تبخیر در ایستگاه تبخیرسنجی چاه‌نیمه به‌عنوان مقادیر مشاهداتی تبخیر از سطح چاه‌نیمه‌ها استفاده شد. جدول (۱) مشخصات ایستگاه‌های مورد استفاده را نشان می‌دهد.

جدول (۱): مشخصات ایستگاه‌های هواشناسی مورد استفاده برای مطالعه

Table (1): The characteristics of the meteorological stations used to study

نام ایستگاه	نوع ایستگاه	طول جغرافیایی	عرض جغرافیایی	میانگین بارش سالانه (میلی‌متر)	کمینه تبخیر سالانه (میلی‌متر)	بیشینه تبخیر سالانه (میلی‌متر)	ارتفاع از سطح دریا (متر)	سال تأسیس
زابل	سینوپتیک	۶۱° ۲۹'	۳۱° ۱۳'	۴۹/۲۵	۲۱/۲	۸۵۱	۴۸۹/۲	۱۳۶۲
زهک	سینوپتیک	۶۱° ۴۱'	۳۰° ۵۴'	۴۷/۹	۸۱/۸	۷۴۴/۲	۴۹۵	۱۳۷۳
چاه نیمه	تبخیرسنجی	۶۱° ۴۰'	۳۰° ۵۱'	۴۶/۲	۸۳	۹۰۰/۱	۴۹۶	۱۳۷۳

شبه‌سازی نماید یا خیر. سناریو شماره ۲ ترکیب سناریو شماره ۱ به‌علاوه متغیر با دومین ضریب همبستگی است. به همین ترتیب این روند برای فراهم‌آوری سایر سناریوها ادامه پیدا کرد. کارایی این روش توسط شرافتی و همکاران (۲۰۱۹) و خسروی و همکاران (۲۰۱۹) به اثبات رسیده است. از سوی دیگر مهم‌ترین چالش یک مدل محاسبات نرم، گزینش بهترین الگو از متغیرهای ورودی و همچنین در دسترس بودن مجموعه‌ای از داده‌های ورودی است.

می‌شود. بنابراین برآورد دقیق میزان تبخیر از سطح آب در این پهنه بسیار حائز اهمیت است. پژوهش‌های انجام‌شده درباره برآورد تبخیر در این پهنه بر روی مدل‌های تجربی و سنتی استوار بوده‌اند. برای نمونه، دانشکار آراسته و همکاران (۲۰۰۵) از روش بیلان حجمی، پیری (۲۰۱۲) از روش‌های تجربی، محمدابراهیم و همکاران (۲۰۱۷) از مدل SEBS، قادری و ناظر (۲۰۲۰) از مدل‌های سری زمانی، و بزی و همکاران (۲۰۲۱) از مدل WEAP استفاده کرده‌اند. بنابراین نبود یک پژوهش کامل به‌منظور برآورد تبخیر با استفاده از روش‌های جدید در این پهنه از کشورمان، یک ضعف اساسی است. از این‌رو در پژوهش حاضر به بررسی پتانسیل چند گروه از مدل‌های محاسبات نرم تازه توسعه‌یافته برای برآورد تبخیر تشت از سطح آب در منطقه سیستان پرداخته شده است.

افزون بر ساختار مدل، بهترین ترکیب ورودی پارامترها نیز تأثیر زیادی بر توانایی پیش‌بینی مدل و یافته‌های به‌دست‌آمده دارند. از این‌رو در نخستین گام ضریب همبستگی پیرسون بین تک‌تک پارامترهای هواشناسی با مقادیر تبخیر محاسبه شد. سپس بر اساس این ضریب همبستگی، تعداد هشت سناریو با ترکیب پارامترهای هواشناسی فراهم شد (جدول ۲). برای نمونه، سناریو شماره ۱ متغیر ورودی با بیشترین ضریب همبستگی پیرسون است. این سناریو نشان خواهد داد که آیا متغیری با همبستگی بالاتر با تبخیر، به‌تنهایی می‌تواند تبخیر را با دقت قابل قبولی

جدول (۲): ترکیب پارامترهای ورودی به مدل‌ها

Table (2): The combination of the input variables to the Models

سناریو (ترکیب ورودی)	پارامترهای ورودی
۱	$E = f(T_{min})$
۲	$E = f(T_{min}, T_{mean})$
۳	$E = f(T_{min}, T_{mean}, T_{max})$
۴	$E = f(T_{min}, T_{mean}, T_{max}, W_{mean})$
۵	$E = f(T_{min}, T_{mean}, T_{max}, W_{mean}, RH_{max})$
۶	$E = f(T_{min}, T_{mean}, T_{max}, W_{mean}, RH_{max}, RH_{mean})$
۷	$E = f(T_{min}, T_{mean}, T_{max}, W_{mean}, RH_{max}, RH_{mean}, RH_{min})$
۸	$E = f(T_{min}, T_{mean}, T_{max}, W_{mean}, RH_{max}, RH_{mean}, RH_{min}, R)$

T_{min} = کمینه دمای ماهانه، T_{mean} = متوسط دمای ماهانه، T_{max} = بیشینه دمای ماهانه، W_{mean} = متوسط سرعت ماهانه باد، RH_{max} = بیشینه رطوبت نسبی، RH_{min} = کمینه رطوبت نسبی، RH_{mean} = متوسط رطوبت نسبی، R = مجموع بارش ماهانه

مدل‌های مورد استفاده

است. به منظور مطالعه بیشتر می‌توان به مطالعه محمدی و همکاران (۲۰۲۱) مراجعه کرد. شایان ذکر است که تمامی مراحل اجرای مدل‌ها در محیط برنامه نویسی R صورت گرفته است.

در پژوهش حاضر از هشت مدل محاسبات نرم به منظور شبیه‌سازی و پیش‌بینی تبخیر تشت از آب سطحی در پهنه سیستان استفاده شد. به منظور جلوگیری از فزونی مطالب گزیده‌ای از توصیفات مدل‌ها در جدول (۳) آورده شده

جدول (۳): گزیده‌ای از مدل‌های محاسبات نرم مورد استفاده

Table (3): A summary of the soft computing models

منبع	پکیج مورد استفاده	مخفف	نام لاتین	نام فارسی
Liaw and Wiener, 2002	party	RF	Random forest	الگوریتم جنگل تصادفی
Helleputte and Gramme, 2017	LiblineAR	SVM	Support Vector Machine	ماشین بردار پشتیبان
Zema et al., 2020	brnn	ANN	Artificial Neural Networks	شبکه عصبی مصنوعی
Nobre and Neves, 2019	xgboost	XGBoost	Extremem Gradient Bossting	ارتقای گرادیان شدید
Zahiri and Nezaratian, 2020	RWeka	TM	Tree Model	مدل درختی
Zahiri and Nezaratian, 2020	earth	MARS	Multivariate Adaptive Regression Splines	مدل اسپلاین چندمتغیره رگرسیون انطباقی
Karatzoglou et al., 2019	Kernlab	GP	Gaussian Process	مدل فرایند گوسی
Gramacy, 2019	monomvn	BART	Baysian Additive Regression Tree	مدل بیزی جمعی رگرسیون درختی

معیارهای ارزیابی

همبستگی، انحراف معیار سری‌های زمانی و $RMSD^1$ استوار است. روابط و اطلاعات بیشتر در مطالعه تیلور (۲۰۰۱) درج شده است.

به منظور ارزیابی مدل‌های مورد استفاده در این تحقیق از معیارهای ارزیابی ضریب تعیین، ریشه میانگین مربعات خطا و میانگین خطای مطلق به شرح زیر استفاده شد:

نتایج

همان‌گونه که بیان گردید، در این پژوهش از داده‌های تبخیر تشت ایستگاه تبخیرسنجی واقع شده در محل چاه نیمه‌های

در این مطالعه، علاوه بر روش‌های ارزیابی نام‌برده از نمودار تیلور (تیلور، ۲۰۰۱) نیز به منظور ارزیابی کارایی مدل‌ها استفاده شد. این نمودار بر مبنای ارتباط هندسی بین ضریب

1. Root Mean Square Deviation

همبستگی پارامترهای مختلف با پارامتر تبخیر تشت برای ایستگاه‌ها نشان می‌دهد. افزون بر آن به‌منظور تقویت این روش و تعیین جایگاه و اهمیت نسبی پارامترهای ورودی مدل از تابع varImp بسته R caret (Kuhn, 2012) استفاده شد؛ که یافته‌های این روش (شکل ۲) نیز هم‌راستا با یافته‌های جدول (۶) هستند؛ از این رو از این دو روش برای تعیین بهترین ترکیب ورودی مدل‌ها (جدول ۲) استفاده شد.

یافته‌های شبیه‌سازی مدل‌ها

در این پژوهش همان‌طور که بیان شد، از هشت مدل محاسبات نرم برای شبیه‌سازی و پیش‌بینی مقادیر تبخیر تشت در دشت سیستان با تأکید بر چاه‌نیمه‌های سیستان استفاده گردید. برای این هدف از ۷۰٪ داده‌های برای مرحله آموزش و ۳۰٪ داده‌ها برای مرحله صحت‌سنجی بهره‌برده شد. به‌منظور جلوگیری از فزونی نوشتار، فقط نتایج بهترین مدل در هر سناریو در ایستگاه‌ها ارائه می‌شود (جدول ۷ یافته‌های ایستگاه سینوپتیک زابل، و جدول ۸ یافته‌های ایستگاه سینوپتیک زهک). همان‌طور که جدول ۷ و ۸ نشان می‌دهد، خروجی‌های دو ایستگاه زابل و زهک بسیار به هم نزدیک و مشابه یکدیگرند. که یکی از دلایل آن، نزدیکی دو ایستگاه به یکدیگر و شرایط اقلیمی و توپوگرافی یکسان در دشت سیستان است. از این رو با هدف جلوگیری از فزونی نوشتار فقط به ارائه و تحلیل نمودارها و یافته‌های ایستگاه زابل پرداخته می‌شود.

سیستان استفاده شد. جدول (۴) مقادیر ماهانه تبخیر تشت ایستگاه تبخیرسنجی چاه‌نیمه در بازه زمانی ۱۳۹۰-۱۴۰۰ را نشان می‌دهد. از آنجا که داده‌های هواشناسی ایستگاه‌های سینوپتیک زابل و زهک نسبت به داده‌های ایستگاه تبخیرسنجی چاه‌نیمه کامل‌تر و با دوره آماری طولانی‌تری است و از طرفی همبستگی بالا بین مقادیر تبخیر تشت ایستگاه چاه‌نیمه با مقادیر تبخیر تشت دو ایستگاه زابل و زهک (به ترتیب ۰/۹۹ و ۰/۹۸)، باعث شد از مقادیر تبخیر تشت این دو ایستگاه سینوپتیک در ادامه مراحل کار استفاده گردد. جدول (۵) ماتریس همبستگی بین مقادیر تبخیر تشت و شکل (۱) نمودار پراکنش مقادیر تبخیر تشت این سه ایستگاه را نشان می‌دهد.

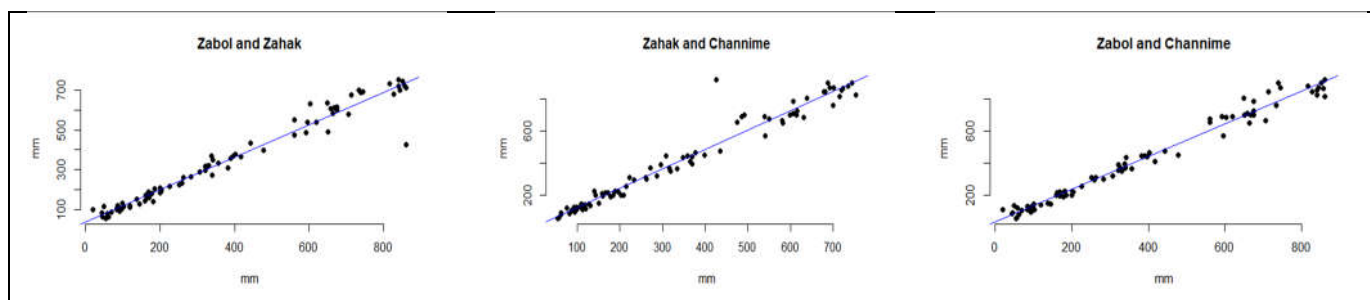
تعیین بهترین ترکیب ورودی

همان‌طور که بیان شد، در ابتدا با استفاده از همبستگی متغیرهای اقلیمی با مقادیر تبخیر تشت، بهترین ترکیب (سناریوی) ورودی برای مدل‌ها تعیین گردید (جدول ۲)؛ که در ادامه هر هشت مدل با استفاده از این سناریوهای ورودی اجرا شد. یافته‌ها نشان داد که بیشترین همبستگی در پیوند با پارامتر کمینه دمای ماهانه با مقدار ضریب همبستگی ۰/۹۶ برای هر دو ایستگاه زابل و زهک و متوسط دمای ماهانه با مقدار ۰/۹۵ برای ایستگاه زابل و مقدار ۰/۹۴ برای ایستگاه زهک است. کمترین مقدار نیز برای پارامتر مجموع بارش ماهانه با ضریب همبستگی ۰/۴۲- برای ایستگاه زابل و مقدار ۰/۴۶- برای ایستگاه زهک است. جدول (۶) مقادیر

جدول (۴): مقادیر ماهانه تبخیر تشت (میلی‌متر) ایستگاه چاه‌نیمه در بازه زمانی ۱۳۹۰-۱۴۰۰

Table (4): Monthly values of Actual Evaporation (mm) of ChahNimeh Station in period of 2011-2021

سالانه	شهریور	مرداد	تیر	خرداد	اردیبهشت	فروردین	اسفند	بهمن	دی	آذر	آبان	مهر
۴۹۹۰/۷	۶۸۴/۶	۸۵۶/۸	۸۶۵/۵	۷۰۷/۱	۴۵۲/۲	۲۹۹/۵	۲۰۰	۱۲۶/۵	۹۱/۵	۱۱۷/۳	۲۰۴/۳	۳۸۲/۲



شکل (۱): نمودار پراکنش مقادیر تبخیر تشت ایستگاه‌های مورد مطالعه

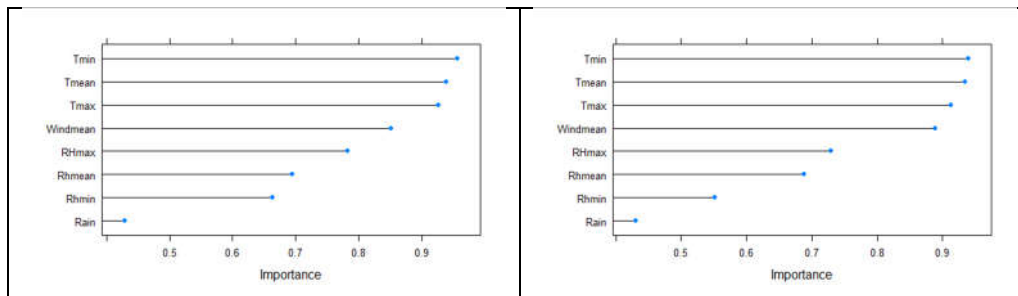
Figure (1): Scatter Plot of Monthly total of Actual Evaporation for the selected stations

جدول (۶): مقادیر ضریب همبستگی پارامترهای هواشناسی ورودی با پارامتر تبخیر تشت

Table (6): Correlation coefficients between input variables and Actual Evaporation

پارامتر ایستگاه	T_{min}	T_{mean}	T_{max}	W_{mean}	RH_{max}	RH_{mean}	RH_{min}	R
زابل	۰/۹۶	۰/۹۵	۰/۹۳	۰/۹۲	-۰/۸۲	-۰/۷۹	-۰/۶۹	-۰/۴۲
زهک	۰/۹۶	۰/۹۴	۰/۹۳	۰/۹۱	-۰/۸۵	-۰/۷۸	-۰/۷۴	-۰/۴۶

T_{min} = کمینه دمای ماهانه، T_{mean} = متوسط دمای ماهانه، T_{max} = بیشینه دمای ماهانه، W_{mean} = متوسط سرعت ماهانه باد، RH_{max} = بیشینه رطوبت نسبی، RH_{mean} = متوسط رطوبت نسبی، RH_{min} = کمینه رطوبت نسبی، R = مجموع بارش ماهانه



شکل (۲): اهمیت پارامترهای ورودی به مدل‌ها در ایستگاه‌های مورد مطالعه

Figure (2): The relative importance of input variable to the models in the selected stations

جدول (۷): نتایج بهترین مدل برای شبیه‌سازی مقادیر تبخیر تشت در سناریوهای ترکیبی (ایستگاه زابل)

Table (7): Result of the Best Model for Simulation of Evaporation Based on input variable combinations (Zabol Station)

صحت سنجی			آموزش			مخفف	بهترین مدل	شماره سناریو
MAE	R^2	RMSE	MAE	R^2	RMSE			
44.21	0.97	57.30	56.40	0.93	75.10	MT	Model Tree	1
47.55	0.98	61.57	53.80	0.94	71.50	RF	Random Forest	2
48.11	0.97	61.27	61.00	0.92	80.80	MT	Model Tree	3
44.90	0.96	59.90	38.70	0.97	49.40	ANN	Bayesian Regularized Neural Networks	4
46.00	0.97	59.00	42.70	0.96	56.60	RF	Random Forest	5
46.30	0.97	59.24	41.80	0.96	56.00	RF	Random Forest	6
45.70	0.97	58.23	42.00	0.96	56.47	RF	Random Forest	7
45.30	0.97	58.30	42.20	0.96	57.00	RF	Random Forest	8

جدول (۸): نتایج بهترین مدل برای شبیه‌سازی مقادیر تبخیر تشت در سناریوهای ترکیبی (ایستگاه زهک)

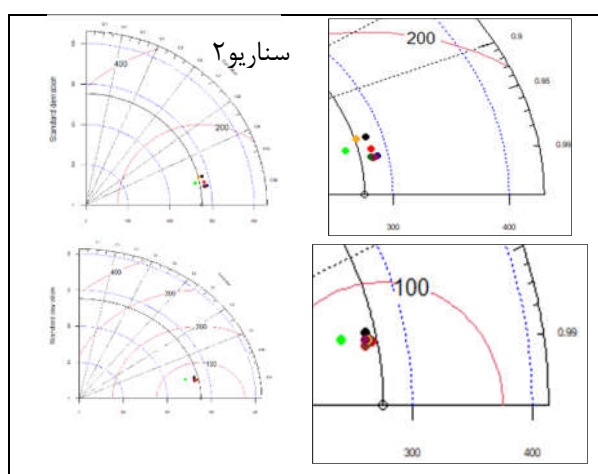
Table (8): Result of the Best Model for Simulation of Evaporation Based on input variable combinations (Zahak Station)

صحت سنجی			آموزش			مخفف	بهترین مدل	شماره سناریو
MAE	R^2	RMSE	MAE	R^2	RMSE			
52.26	0.97	68.42	45.20	0.94	60.76	MT	Model Tree	1
51.70	0.97	68.40	40.65	0.95	54.90	RF	Random Forest	2
55.54	0.97	72.90	44.40	0.94	61.30	MT	Model Tree	3
50.18	0.97	64.30	32.40	0.97	43.40	ANN	Bayesian Regularized Neural Networks	4
51.90	0.98	64.90	36.00	0.96	48.40	RF	Random Forest	5
51.87	0.98	65.10	35.20	0.97	48.10	RF	Random Forest	6
52.20	0.97	65.40	35.20	0.97	47.39	RF	Random Forest	7
51.90	0.98	65.00	34.51	0.97	46.70	RF	Random Forest	8

سناریو شماره ۷ با هفت ورودی (جدول ۲) با مقدار ۵۸/۲۳ و سناریو شماره ۸ با هشت ورودی با مقدار ۵۸/۳ در رتبه‌های بعدی قرار دارند. بر پایه معیار MAE نیز بهترین عملکرد مربوط به سناریو شماره ۱ با مقدار ۴۴/۲۱ است. از این رو بر پایه معیارهای ارزیابی $RMSE$ ، MAE و R^2 بهترین ترکیب ورودی برای پیش‌بینی و شبیه‌سازی مقادیر تبخیر تشت در دشت سیستان، سناریو ۱ یعنی فقط دمای کمینه ماهانه با مدل درختی می‌باشد.

همچنین یافته‌های جدول ۷ و ۸ نشان می‌دهد که در بین مدل‌ها، مدل جنگل تصادفی (RF) در پنج سناریو (سناریو ۲، ۵، ۶، ۷ و ۸) از هشت سناریو به‌عنوان بهترین مدل نمود داشته است. مدل درختی نیز در دو سناریو (سناریو ۱ و ۳) نمود داشته است. از این رو مدل جنگل تصادفی (RF) و مدل درختی (MT) به‌عنوان بهترین مدل‌های محاسبات نرم به‌منظور شبیه‌سازی و برآورد تبخیر تشت دشت سیستان معرفی می‌شوند.

شکل (۴) نیز نمودار مقادیر شبیه‌سازی شده توسط مدل‌های منتخب و تبخیر مشاهداتی در سناریوهای مختلف نشان می‌دهد. همان‌طور که این شکل‌ها آشکار می‌سازند، در نقاط کمینه و بیشینه، مقادیر شبیه‌سازی شده و مشاهداتی اختلاف بیشتری نسبت به هم دارند؛ به طوری که در نقاط بیشینه و کمینه، مقادیر شبیه‌سازی شده بیشتر از مقادیر مشاهداتی هستند. اما در مقادیر دیگر بین مقادیر شبیه‌سازی شده و مشاهداتی اختلاف چندانی وجود ندارد.



Obs SVM BRR RF GP M5 MARS XGB ANN

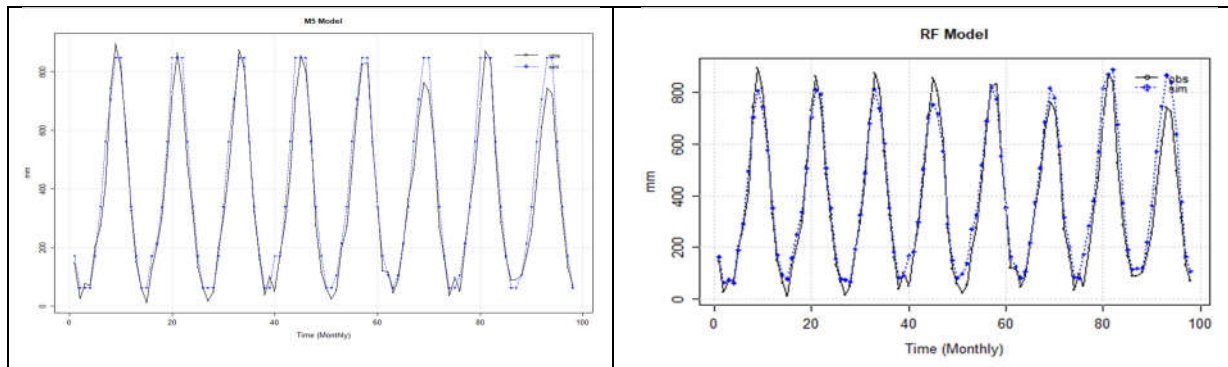
شکل (۳): دیاگرام تیلور برای تمامی مدل‌های پیش‌بینی

Figure (3): The Taylor diagram for all predictive models

همان‌طور که جدول (۷) نشان می‌دهد، بهترین مدل برای سناریو شماره ۱، مدل درختی با مقدار R^2 برابر ۰/۹۷ و $RMSE$ برابر ۵۷/۳، برای سناریو شماره ۲، مدل جنگل تصادفی با مقدار R^2 برابر ۰/۹۸ و $RMSE$ برابر ۶۱/۷، برای سناریو شماره ۳، مدل درختی با مقدار R^2 برابر ۰/۹۷ و $RMSE$ برابر ۶۱/۲۷، برای سناریو شماره ۴، مدل شبکه عصبی مصنوعی با مقدار R^2 برابر ۰/۹۶ و $RMSE$ برابر ۵۹/۹، و برای سناریوهای ۵ تا ۸ مدل جنگل تصادفی با R^2 برابر ۰/۹۷ و $RMSE$ به ترتیب برابر ۵۹، ۵۹/۲۴، ۵۸/۲۳ و ۵۸/۳ در مرحله صحت‌سنجی است. افزون بر معیارهای اشاره شده، از نمودار تیلور نیز به‌منظور ارزیابی کارایی مدل‌ها و انتخاب مناسب‌ترین مدل نیز استفاده گردید. که یافته‌های نمودار تیلور نیز با معیارهای ارزیابی در جدول (۷) و (۸) همخوانی داشته و این نمودار نیز یافته‌های جدول (۷) و (۸) را تأیید کردند. شکل (۳) نمودار تیلور دو سناریو شماره ۲ و شماره ۸ در ایستگاه زابل را نشان می‌دهد.

همان‌طور که یافته‌ها نشان می‌دهد با افزایش ورودی متغیرها به مدل، تفاوت چندانی در دقت مدل‌ها مشاهده نشد. برای نمونه سناریوی یک با فقط یک ورودی (دمای کمینه ماهانه) در سنجش با سناریو شماره ۸ با هشت ورودی (جدول ۲) مقدار R^2 در هر دو سناریو ۱ و ۸ برابر ۰/۹۸ در مرحله صحت‌سنجی است. افزون بر آن با افزایش تعداد ورودی‌ها مقدار $RMSE$ افزایش چندانی نداشته است. برای نمونه مقدار $RMSE$ در سناریو ۱ از مقدار ۵۷/۳ به مقدار ۵۸/۳ در سناریو ۸ فزونی داشته است.

بر پایه معیار R^2 بهترین عملکرد مربوط به سناریو شماره ۲ با مقدار ۰/۹۸ در مرحله صحت‌سنجی است. ترکیب ورودی این سناریو شامل دمای کمینه ماهانه و دمای متوسط ماهانه ایستگاه سینوپتیک زابل است. همچنین سناریوهای شماره ۱، ۳، ۵، ۶ و ۷ با مقدار ۰/۹۷ در رتبه بعدی و سناریو ۴ با مقدار ۰/۹۶ در رتبه پایانی قرار دارد. بر پایه معیار $RMSE$ بهترین عملکرد مربوط به سناریو شماره ۱ با مقدار ۵۷/۳ در مرحله صحت‌سنجی است. این سناریو تنها ترکیب ورودی دمای کمینه ماهانه را دارد. بر پایه این معیار ارزیابی،



شکل (۴): نمودار مقادیر شبیه‌سازی‌شده و تبخیر تست مدل‌های منتخب در سناریوهای یک و هشت
Figure (4): Observation and Simulation of Evaporation using Selected Models in Scenarios 1 and 8

بحث و نتیجه‌گیری

تبخیر یکی از مهم‌ترین پارامترهای اقلیمی و هیدرولوژیکی در هر پهنه‌ای محسوب می‌شود. دانستن مقدار تبخیر و به‌کارگیری مدل مناسب در برآورد آن در مدیریت منابع آبی و کشاورزی ارزش فراوانی دارد؛ به‌ویژه در پهنه‌های خشک و فراخشک همچون دشت سیستان که تنها منبع آب شرب، کشاورزی و صنعت در این پهنه، منبع آب سطحی چاه‌نیمه‌هاست. این در حالی است که پژوهش‌های کنونی در پیوند با مدل‌های برآورد و پیش‌بینی پدیده‌های غیرخطی آشکار ساخته‌اند که مدل‌های محاسبات نرم توانایی و انعطاف‌پذیری بایسته‌ای در برآورد و پیش‌بینی این پدیده‌ها داشته‌اند. از این‌رو در این پژوهش به بررسی کارایی هشت مدل نو از مدل‌های محاسبات نرم برای برآورد تبخیر تست در پهنه سیستان پرداخته شد. برای این هدف از هشت ترکیب ورودی پارامترهای هواشناسی در پیکره هشت سناریو استفاده گردید.

یافته‌ها نشان داد که با افزایش ورودی متغیرها به مدل، تفاوت چندانی در دقت مدل‌ها مشاهده نشد. برای نمونه اندازه R^2 سناریوی یک با فقط ورودی دمای کمینه ماهانه برابر با سناریو شماره ۸ با هشت ورودی برابر $0/98$ در مرحله صحت‌سنجی است. افزون بر آن با افزایش تعداد ورودی‌ها مقدار RMSE افزایش چندانی نداشته است. برای نمونه مقدار RMSE در سناریو ۱ از مقدار $57/3$ به مقدار $58/3$ در سناریو ۸ فزونی داشته است. همچنین اندازه MAE از $44/21$ در سناریو ۱ به $45/3$ در سناریو ۸ افزایش یافته است. از این‌رو یافته‌ها نشان داد که مدل با فقط ورودی دمای کمینه ماهانه

عملکردی همانند مدل با هشت ورودی پارامتر هواشناسی دارد. این یافته‌ها با نتایج دانشفراز (۲۰۱۶) که به بررسی شبکه عصبی مصنوعی برای برآورد تبخیر از تشتک با استفاده از دو مدل ورودی شامل مدل ۱ با هفت ورودی و مدل ۲ با یک ورودی یعنی کمینه درجه‌حرارت در ایستگاه تبریز پرداختند همخوانی دارد، به‌طوری که این محققان نیز در پژوهش خویش بیان داشتند که اگرچه مدل ۲ با ورودی تنها یک عامل درجه‌حرارت کمینه را دارد، با توجه به ضریب تعیین بالای $0/9$ و خطای یکسان با مدل ۱ با هفت متغیر ورودی، عملکردی همانند مدل ۱ در برآورد تبخیر داشته است. کیسی و همکاران (۲۰۱۹) بیان داشتند که پیش‌بینی با استفاده از مدل‌های مختلف یادگیری ماشین در مجموعه داده‌های گوناگون منجر به یافته‌های متفاوتی می‌شود. یک چرایی این تفاوت وجود ساختارهای متفاوت الگوریتم‌ها است. به‌عبارتی دیگر، توانایی محاسباتی و پیچیدگی مدل‌های مختلف محاسبات نرم، پی‌آیندهای متفاوتی دارد. یافته‌های این پژوهش نیز نشان داد که در بین هشت مدل، مدل جنگل تصادفی (RF) در پنج سناریو (سناریو ۲، ۵، ۶، ۷ و ۸) از هشت سناریو به‌عنوان بهترین مدل عمل کرده است. مدل درختی نیز در دو سناریو (سناریو ۱ و ۳) نمود داشته است. از این‌رو مدل جنگل تصادفی (RF) و مدل درختی (MT) به‌عنوان بهترین مدل محاسبات نرم به‌منظور شبیه‌سازی و برآورد تبخیر تست دشت سیستان معرفی شدند. این یافته‌ها با یافته‌های خلیلی نفتچالی و همکاران (۲۰۱۷) که مدل درختی M5 را به‌عنوان بهترین مدل برآورد تبخیر بیرجند معرفی نمودند همخوانی دارد. همچنین

در بین مدل‌های مورد ارزیابی معرفی نمودند. همچنین ستاری و اسمعیل‌زاده (۲۰۱۷) نیز در پژوهشی مدل درختی M5 را برای برآورد Kp در مناطق خشک و نیمه‌خشک معرفی کرد. در پایان می‌توان بیان داشت مهم‌ترین یافته این پژوهش معرفی یک مدل محاسبات نرم برای برآورد دقیق تبخیر تشت دشت سیستان با استفاده از یک پارامتر هواشناسی یعنی دمای کمینه ماهانه است. روشی ساده بدون نیاز به جداول و معادلات پیچیده و همچنین بدون نیاز به پارامترهای بی‌شمار هواشناسی، پارامتر دمای کمینه که در بیشتر ارگان‌های سازمانی قابل دسترس است. تنها محدودیت این پژوهش این است که در پهنه برگزیده‌ای از ایران انجام شده است و یافته‌های آن برای مناطق با اقلیم‌های ناهمسان انجام‌پذیر نیست. پیشنهاد ما انجام پژوهش‌های مشابه در مناطق با شرایط اقلیمی متفاوت است.

با یافته‌های خسروانی و همکاران (۲۰۱۹) نیز همخوانی دارد. محققانی که به بررسی کارایی مدل‌های داده‌کاوی در عراق پرداختند، نشان دادند که مدل‌های درخت تصمیم همانند RF و M5 در سنجش با مدل‌های ترکیبی ANFIS نتایج خوبی را ارائه می‌دهند. این پژوهشگران دلیل نتایج خوب این مدل‌ها را دو دلیل اصلی بیان کردند: نخست مدل‌های مبتنی بر درخت هیچ فرضی در مورد توزیع داده‌ها ندارند و دوم آنکه این مدل‌ها نسبت به داده‌های پرت حساس نیستند، زیرا آن‌ها را در مناطق کوچکی از فضای ویژگی جدا می‌کنند. همچنین با یافته‌های شرافتی و همکاران (۲۰۱۹) که به بررسی کارایی مدل‌های داده‌کاوی برای برآورد تابش خورشید در بورکینافاسو آفریقا پرداختند همخوانی دارد. یافته‌های این پژوهشگران نیز مدل RF را برای شبیه‌سازی تابش خورشیدی

منابع

- Bazzi, H., Ebrahimi, H. and Aminnejad, B., 2021. A comprehensive statistical analysis of evaporation rates under climate change in Southern Iran using WEAP (Case study: Chahnimeh Reservoirs of Sistan Plain). *Ain Shams Engineering Journal*, 12(2), 1339-1352.
- Daneshfaraz, R. 2016. Sensitivity Analysis of the Effective Parameters upon Daily Evaporation Using Garson Equation and Artificial Neural Network (Case Study: Tabriz city). *Geography and Planning*, 19(54), 127-142. (Persian)
- Daneshkar Arasteh, P., Tajrishi, M., Mirlatif, M. and Saghafian, B., 2005. Statistical model of free water surface evaporation using the volume balance method in Chahnimeh reservoir, Sistan-Iran. *Pajouhesh va Sazandegi*. 18 (68): 2-14. (Persian)
- Farasat, M., Seyedian, M. and daab, K., 2021. Evaporation Modeling of Free Surface Water Using SVM and LSSVM Models. *Irrigation and Water Engineering*, 11(3), 272-288. doi: 10.22125/iwe.2021.128205. (Persian)
- Feng, K. and Tian, J., 2021. Forecasting reference evapotranspiration using data mining and limited climatic data, *European Journal of Remote sensing*, 54: sup2, 363-371
- Ghaderi, A. and yosefian nazer, H., 2020. Prediction of Evaporation Changes in Half Well Reservoirs of Sistan by Using Randomized Simulator Models. *Geography and Development Iranian Journal*, 18(58), 225-238. doi: 10.22111/gdij.2020.5362. (Persian)
- Ghorbani, M.A., Deo, R.C., Yaseen, Z.M., H. Kashani, M. and Mohammadi, B., 2017. Pan evaporation prediction using a hybrid multilayer perceptron-firefly algorithm (MLPFFA) model: case study in North Iran. *Theoret. Appl. Climatol.* doi: z10.1007/s00704-017-2244-0.
- Gramacy, R.B., 2019. Package 'monomvn'. Package 'monomvm' version 1.9-13. Retrieved from <https://cran.r-project.org/web/packages/momomvm/index.html>
- Helleputte, T. and Gramme, P., 2017. LiblineaR: Linear predictive models based on the LIBLINEAR C/C++ Library. *R package version*, 2-10
- Kamel, A. H., Afan, H. A., Sherif, M., Ahmed, A. N., and El-Shafie, A. 2021. RBFNN versus GRNN modeling approach for sub-surface evaporation rate prediction in arid region. *Sustainable Computing: Informatics and Systems*, 30, 100514.
- Kaya, Y. Z., Zelenakova, M., Üneş, F., Demirci, M., Hlavata, H., and Mesaros, P. 2021. Estimation of daily evapotranspiration in Košice City (Slovakia) using several soft computing techniques. *Theoretical and Applied Climatology*, 144(1), 287-298.
- Karatzoglou, A., Smola, A., Hornik, K. and Karatzoglou, M. A. 2019. Package 'kernlab'. CRAN R Project.
- Khalili Naft Chali, A., Khashei Siuki, A. and Shahidi, A. 2017. Compare KNN and M5 decision tree models in anticipation of evaporation and comparison with empirical equations (Case Study of Birjand). *Iranian Journal of Irrigation & Drainage*, 11(3), 356-366. (Persian)
- Khosravi, K., Daggupati, P., Alami, M. T., Awadh, S. M., Ghareb, M. I., Panahi, M. and Yaseen, Z. M., 2019. Meteorological data mining

- and hybrid data-intelligence models for reference evaporation simulation: A case study in Iraq. *Computers and Electronics in Agriculture*, 167, 105041.
15. Kisi, O., Heddad, S. and Yaseen, Z.M., 2019. The implementation of univariable schemebased air temperature for solar radiation prediction: New development of dynamic evolving neural-fuzzy inference system model. *Appl. Energy* 241, 184–195.
 16. Liaw, A. and Wiener, M., 2002. Classification and regression by randomForest. *R news*, 2(3), 18-22.
 17. Majhi, B., Naidu, D., Mishra, A.P. and Satapathy, S.C., 2019. Improved prediction of daily pan evaporation using Deep-LSTM model. *Neural Comput. Appl.* <https://doi.org/10.1007/s00521-019-04127-7>
 18. Mattar, Mohamed A., 2018. "Using gene expression programming in monthly reference evapotranspiration modeling: A case study in Egypt", *Agricultural Water Management*, Elsevier, vol. 198(C), p. 28-38.
 19. Mirhashemi, S., Panahi, M. and Zareei, L., 2020. Evaluation of M5P Algorithm for Estimation of Potential Evapotranspiration, Minimum and Maximum Temperature (Case study: Sari Weather Station). *Journal of Meteorology and Atmospheric Science*, 2(4), 287-295. (Persian).
 20. Mohammad Ebrahim, M., Mohammadrezapour, O. and Akbarzadeh seghaleh, H., 2017. Evaluating SEBES Model to Estimate Actual Evapotranspiration using MODIS Sensor Data in Regional Scale (Case Study: Sistan Plain). *Iranian journal of Ecohydrology*, 4(4), 1141-1150. doi: 10.22059/ije.2017.63243. (Persian)
 21. Mohammadi, M., Vagharfard, H., Mahdavi Najafabadi, R., Daneshkar Arasteh, P. and Nazemosadat, M., 2021. Rainfall-runoff Modelling of Coastal Watersheds near Hormuz Strait Using Data Mining. *Iranian Journal of Soil and Water Research*, 52(2), 313-327. doi: 10.22059/ijswr.2021.309641.668732
 22. Moran, M.S., Rahman, A.F., Washburne, J.C., Goodrich, D.C., Wertz, M.A. and Kustas, W.P., 1996. Combining the Penman-Monteith equation with measurements of surface temperature and reflectance to estimate evaporation rates of semiarid grassland. *Agric. For. Meteorol.* <https://doi.org/10.1007/s12549-010-0046-9>
 23. Nobre, J. and Neves, R. F., 2019. Combining principal component analysis, discrete wavelet transform and XGBoost to trade in the financial markets. *Expert Systems with Applications*, 125, 181-194.
 24. Panahi, M., Gayen, A., Pourghasemi, H. R., Rezaie, F., and Lee, S. 2020. Spatial prediction of landslide susceptibility using hybrid support vector regression (SVR) and the adaptive neuro-fuzzy inference system (ANFIS) with various metaheuristic algorithms. *Science of the Total Environment*, 741, 139937.
 25. Piri, H., 2012. Assesment of Computational Methods of Estimation of Potential Evapotranspiration Using Lysimeter Data (Case Study: Sistan Plain). *Irrigation and Water Engineering*, 3(1), 50-61. (Persian)
 26. Priestley, C.H.B. and Taylor, R.J., 1972. On the assessment of the surface heat flux and evaporation using large-scale parameters. *Mon. Weather Rev.* 100, 81–92.
 27. Qasem, S.N., Samadianfard, S., Kheshtgar, S., Jarhan, S., Kisi, O., Shamshirband, S. and Chau, K.-W., 2019. Modeling monthly pan evaporation using wavelet support vector regression and wavelet artificial neural networks in arid and humid climates. *Eng*
 28. Samadianfard, S. and Panahi, S., 2019. Estimating Daily Reference Evapotranspiration using Data Mining Methods of Support Vector Regression and M5 Model Tree. *Journal of Watershed Management Research*. 9(18), 157-167. (Persian).
 29. Sanikhani, H., Kisi, O., Maroufpoor, E. and Yaseen, Z.M., 2018. Temperature-based modeling of reference evapotranspiration using several artificial intelligence models: application of different modeling scenarios. *Theor. Appl. Climatol.* 1–14. <https://doi.org/10.1007/s00704-018-2390-z>.
 30. Sattari, M. T. and Esmailzadeh, B. (2017). Performance Assessment of M5 Tree Model and Genetic Programming in Tabriz Station Reference Evapotranspiration Modeling. *Water Resources Engineering*, 9(31), 11-20. (Persian).
 31. Sayyahi, F., Farzin, S., and Karami, H. 2021. Forecasting daily and monthly reference evapotranspiration in the Aidoghmoush basin using multilayer perceptron coupled with water wave optimization. *Complexity*, 2021.
 32. Sepehri, S., Abbasi, F., Zarei, G. and Nakhjavanimoghaddam, M., 2021. Investigation of Artificial Neural Network Based Models and Sensitivity Analysis for Reference Evapotranspiration Estimating. *Iranian Journal of Irrigation & Drainage*, 14(6), 2089-2099. (Persian)
 33. Sharafati, A., Khosravi, K., Khosravinia, P., Ahmed, K., Salman, S. A., Yaseen, Z. M. and Shahid, S., 2019. The potential of novel data mining models for global solar radiation prediction. *International Journal of Environmental Science and Technology*, 16(11), 7147-7164.
 34. Shiri, J., Zounemat-Kermani, M., Kisi, O., and Mohsenzadeh Karimi, S. 2020. Comprehensive assessment of 12 soft computing approaches for modelling reference evapotranspiration in humid locations. *Meteorological Applications*, 27(1), e1841.
 35. Siasar, H. and Honar, T., 2020. Comparison of Performance of GLM, RF and DL Models in Estimation of Reference Evapotranspiration in

- Zabol Synoptic Station. *jwmr*. 11 (22): 210-219. (Persian)
36. Stanhill, G., 2002. Is the class a evaporation pan still the most practical and accurate meteorological method for determining irrigation water requirements? *Agric. For. Meteorol.* [https://doi.org/10.1016/S0168-1923\(02\)00132-6](https://doi.org/10.1016/S0168-1923(02)00132-6).
- 37.
38. Taylor, K. E., 2001. Summarizing multiple aspects of model performance in a single diagram. *Journal of Geophysical Research: Atmospheres*, 106(D7), 7183-7192.
39. Zahiri, J. and Nezaratian, H., 2020. Estimation of transverse mixing coefficient in streams using M5, MARS, GA, and PSO approaches. *Environmental Science and Pollution Research*, 1-14
40. Zema, D. A., Lucas-Borja, M. E., Fotia, L., Rosaci, D., Sarnè, G. M. and Zimbone, S. M., 2020. Predicting the hydrological response of a forest after wildfire and soil treatments using an Artificial Neural Network. *Computers and Electronics in Agriculture*, 170, 105280.

Predicting Pan Evaporation in a Hyper-Arid Climate Using Soft Computing Models: A Case Study of Sistan Plain, Sistan-Baluchistan, Iran

Mojtaba Mohammadi^{1*}, Masoomeh Forozanfar², Hamid Gholami³

Received: 2022/6/10

Accepted: 2023/1/11

Expanded abstract

Introduction: Considering the fact that evaporation affects the planning and operations of water resources as a key process in the hydrologic cycle, predicting its magnitude and patterns, particularly in arid, semi-arid, and hyper-arid environments such as Sistan plain (in northern Sistan-Baluchistan Province, Iran) is of great importance. On the other hand, as accurate estimation of Pan evaporation is regarded as one of the main aspects of water management in such regions, it is crucially important to accurately simulate the pan evaporation based on the available regional meteorological parameters. Therefore, this study sought to investigate the capabilities of soft computing techniques for estimating monthly evaporation in Sistan plain. The results of the study could be helpful for the management of water resources in the Sistan area, allowing the policymakers to develop future projects of water resource management/development plans for the region based on Evaporation estimations.

Materials and Methods: various meteorological parameters, including maximum, minimum, and average temperature rates, relative humidity, wind speed, and precipitation rate were used to predict monthly Evaporation using the consistent and uninterrupted historical time series data (1994–2021) collected from three meteorological stations (Zabol, Zahak, and Chahnmeh).

The main purpose of this study was to assess the performance of a soft computing model in simulate pan evaporation. To this end, nine soft computing models, including Model Tree (MT), Random Forest (RF), Support Vector Machines (SVM), Bayesian Ridge Regression (BRR), Gaussian Process (GP), Extreme Gradient Boosting (XGB), Artificial Neural Network (ANN), and Multivariate Adaptive Regression Splines (MARS) were used to predict evaporation at the meteorological stations selected for this research.

1. Assistant Professor (Corresponding Author), Department of Desert Management & Control, Faculty of Natural Resource, Higher education complex of Saravan, Iran.

2. Department of Desert Management & Control, Faculty of Natural Resource, Higher education complex of Saravan, Iran.

3. Associate Professor, Department of Watershed Management Engineering, Faculty of Natural Resources, Hormozgan

University, Bandar Abbas, Iran.

DOI: 10.22052/deej.2021.11.36.43

On the other hand, the model's performance was assessed using statistical measures, including coefficient of determination (R^2), root mean square error (RMSE), mean absolute error (MAE), and Taylor diagram. Moreover, to construct predictive models, the dataset was divided into training (70%) and validation (30%) data. Then, eight combinations of input parameters were selected for Zabol and Zahak sites based on the Pearson correlation coefficient between the individual input parameters and evaporation. Finally, the best input combination and the optimal values for different models were determined using the R programming language.

Results: The different input combinations were determined for the two sites independently based on the inclusion of the weather parameters with the highest coefficient Pearson with evaporation. Then, each model was run using various fixed sets of parameters. For the Zabol station, the minimum temperature rate indicated the greatest correlation coefficient with (0.96), followed by average temperature (0.95) and smallest by rainfall (-0.42). It was also found that the outputs of Zabol and Zahak stations were very close and similar to each other, which could partly be attributed to the proximity of the two stations and their same topography and climatic conditions in the Sistan plain.

On the other hand, the results of assessing the models' performance indicated that in the validation stage, MT (whose $R^2 = 0.97$ and $RMSE=57.3$) delivered the best performance in Zabol station under Scenario 1, RF (with its R^2 and $RMSE$ being 0.98 and 61.7, respectively) performed the best under Scenario 2, MT with its R^2 and $RMSE$ being 0.97 and 61.27, respectively, showed the best performance under Scenario 3, the ANN and MT (whose R^2 and $RMSE$ reported as being 0.96 and 59.9, respectively) put in the best performance under Scenario 4, and RF with $R^2 = 0.97$ and $RMSE=59, 59.24, 58.23,$ and 58.3 delivered the best performance under scenarios 5 to 8, respectively.

Moreover, the results suggested that adding the number of input variables to the models made no difference in their accuracy level. It was also found that out of eight scenarios investigated in this study, the RF model delivered the performance under five scenarios (2, 5, 6, 7, and 8), and that the MT performed well under two scenarios (1 and 3). Therefore, RF and MT could be introduced as the best soft computing models for simulating and estimating the pan evaporation of the Sistan plain.

Conclusion: The results showed that with increasing the input of variables to the model, there was not much difference in the accuracy of the models. For example, R^2 of scenario 1 with only the minimum monthly temperature input is equal to scenario 8 with eight inputs equal to 0.98 in the validation stage. Therefore, the findings showed that the model with only the minimum monthly temperature input has the same performance as the model with eight inputs. The main contribution of this study was introducing a soft computing model to accurately estimate the pan evaporation of the Sistan plain using a meteorological parameter (minimum monthly temperature).

Keywords: Prediction, Evaporation, Soft Computing, Hyper-Arid Region, Sistan Plain.