



مقایسه برآوردهای تبخیر تعرق مرجع روزانه با روش‌های داده‌کاوی و سامانه نیاز آبی گیاهان در استان البرز

آزاده صداقت^{۱*}، نیازعلی ابراهیمی پاک^۲، آرش تافته^۳، سیده نرگس حسینی^۱

تاریخ دریافت: ۱۴۰۱/۰۱/۲۱

تاریخ پذیرش: ۱۴۰۲/۰۲/۰۴

چکیده

تبخیر تعرق مرجع (ET₀) عامل مهمی در تعیین نیاز آبی و برنامه‌ریزی آبیاری گیاهان است و برای تخمین آن، معمولاً از روش نسبتاً دقیق پنمن مانیتیت فائو ۵۶، استفاده می‌شود. هدف از این پژوهش ارزیابی عملکرد روش‌های شبکه عصبی مصنوعی (ANNs)، جنگل تصادفی (RF) و ماشین بردار پشتیبان (SVM) برای تخمین مقدار ET₀ روزانه در استان البرز است. در این پژوهش از داده‌های ده ساله (۱۳۸۹ تا ۱۳۹۹) پنج ایستگاه سینوپتیک استان شامل (مشکین‌دشت، هشتگرد، اشتهارد، طالقان و کرج) استفاده شد و نتایج حاصل با داده‌های سامانه نیاز آبی گیاهان مورد مقایسه قرار گرفت. نتایج نشان داد که برآورد مقدار ET₀ روزانه به روش ANN بر اساس برآوردهای سامانه پیشنهادی، دقت بالاتری نسبت به سایر روش‌ها دارد. در ایستگاه مشکین‌دشت، مقادیر آماره‌های EF و NRMSE در روش ANN برای هر دو مرحله آموزش و آزمون به ترتیب برابر با ۰/۹۶ و ۰/۱۱ و در روش RF به ترتیب برای آموزش برابر با ۰/۹۶ و ۰/۱۱ و برای آزمون برابر با ۰/۹۵ و ۰/۱۲ و در ایستگاه کرج، مقادیر آماره‌های EF و NRMSE در روش ANNs به ترتیب برای آموزش برابر با ۰/۹۶ و ۰/۱۱ و برای آزمون برابر با ۰/۹۵ و ۰/۱۲ و در روش RF برای آموزش به ترتیب برابر با ۰/۹۶ و ۰/۱۲ و برای آزمون برابر با ۰/۹۵ و ۰/۱۳ به دست آمد. با توجه به بالا بودن دقت تخمین مقدار ET₀ روزانه به روش‌های ANNs و RF می‌توان این دو روش را جهت تخمین ET₀ روزانه برای استان البرز توصیه کرد.

واژه‌های کلیدی: تبخیر تعرق مرجع، سامانه نیاز آبی، داده‌کاوی، پنمن مانیتیت فائو ۵۶

مقدمه

ساعات آفتابی، فشار سطح، رطوبت نسبی، سطح ابرناکی و تحت تأثیر غیرمستقیم بارش، خصوصیات خاک و گیاه و کیفیت آب قرار دارد (Salarian et al., 2014). برای تعیین مقدار ET₀ می‌توان از لایسیمتر که دقت اندازه‌گیری بالایی دارند (Allen et al., 2011)، یا از مدل‌های ریاضی مبتنی بر داده‌های اقلیمی استفاده کرد (Ferreira et al., 2019). تغییرات اقلیمی که در سطح جهان در حال رخ دادن است تأثیر زیادی روی مقدار ET₀ دارند. مقدار تبخیر تعرق، در اقلیم‌های گرم در مناطقی مانند غرب آفریقا (Onyutha,)

تبخیر تعرق (ET)، شامل تبخیر از سطح آب و تعرق از سطح گیاه می‌باشد و یکی از اجزای چرخه آب در طبیعت را تشکیل می‌دهد. برآورد ET₀ نقش مهمی در تعیین نیاز آبی گیاهان، برنامه‌ریزی آبیاری و مدیریت منابع آب دارد. تعیین نیاز آبی گیاهان برای مناطقی که آبیاری در سطح وسیعی انجام می‌شود دارای اهمیت ویژه‌ای است (Ferreira et al., 2019). مقدار ET₀ تحت تأثیر مستقیم متغیرهای سرعت باد،

^۲ دانشجویار مؤسسه تحقیقات خاک و آب، سازمان تحقیقات، آموزش و ترویج کشاورزی، کرج، ایران
^۳ استادیار مؤسسه تحقیقات خاک و آب، سازمان تحقیقات، آموزش و ترویج کشاورزی، کرج، ایران

^۱ محقق مؤسسه تحقیقات خاک و آب، سازمان تحقیقات، آموزش و ترویج کشاورزی، کرج، ایران
(*نویسنده مسئول: Azadehsedaghat65@gmail.com)

نحوه ارجاع مقاله:

گزارش‌های ثبت شده نشان می‌دهند که روش ANNs^۲ اغلب عملکرد بهتری نسبت به تکنیک‌های مرسوم دارد (Huo et al., 2014; Shiri et al., 2012). همچنین پژوهشگرانی هم، گزارش کردند روش SVM^۳ عملکرد خوبی را برای تخمین ET₀ نشان داده است (Fan et al., 2018; Wen et al., 2015). Allawi et al., (2019) در پژوهشی توانایی شبکه عصبی مصنوعی بر پایه شعاعی (RBF-NN) و روش‌های رگرسیون بردار پشتیبان (SRV) برای پیش‌بینی مقدار ET₀ در منطقه گرمسیری مالزی بررسی کردند. آنها گزارش کردند در مقایسه با مقادیر تبخیر تعرق اندازه‌گیری شده، دقت روش RBF-NN بالاتر از SRV بود. Alves et al., (2017) در پژوهشی به منظور تخمین مقدار ET₀ روزانه در برزیل با استفاده از روش ANNs و با بهره‌گیری از داده‌های ۱۰ ساله رطوبت نسبی روزانه، شار گرما، سرعت باد در ارتفاع ۲ متری و تابش خالص خورشیدی، نشان دادند که تنها با استفاده از دمای هوا به‌عنوان متغیر ورودی به روش ANNs می‌توان با دقت خوبی مقدار ET₀ را تخمین زد. با توجه به مطالعه‌هایی که اخیراً انجام شد، مقدار ET₀ تحت تأثیر عوامل مختلف هواشناسی در بعد زمان و مکان تغییر می‌کند و از مهم‌ترین پارامتر برای مدیریت و برنامه‌ریزی آب مصرفی و طراحی سیستم‌های آبیاری می‌باشد (Salarian et al., 2014). در بیشتر این مطالعات، پیش‌بینی ET₀ با مدل‌های هوشمند از پارامترهای هواشناسی زیادی به‌عنوان ورودی مدل استفاده شده است (Aytek, 2008; Kumar et al., 2008; Lin et al., 2002). اگرچه انتخاب پارامترهای ورودی بیشتر در راستای افزایش دقت پیش‌بینی‌ها بوده، اما ذکر این نکته ضروری است که دسترسی به کلیه پارامترها در همه مناطق امکان پذیر نیست. بنابراین یافتن مدل‌هایی با روش‌های نوین با حداقل ورودی هواشناسی برای مدل‌سازی ET₀ که ماهیت غیرخطی و پیچیده دارد، امری ضروری به‌نظر می‌رسد. استان البرز بر روی رسوبات آبرفتی رشته کوه البرز بوجود آمده و از سه بخش کوهستانی و کوهپایه‌ای در شمال و منطقه‌ای هموار و دشتی در جنوب تشکیل شده است. بنابراین تعیین اثرات تغییرات آب و هوا بر ET₀ برای استان البرز با شرایط توپوگرافی خاص، نیاز به بررسی و آنالیز فراوان با

(Yin et al., 2010) و جنوب چین (2016)، و مهم‌ترین عامل تغییر مقدار ET₀ در استرالیا (Roderick and Farquhar, 2004)، و ایران (Dinpashoh et al., 2012) عمدتاً ناشی از سرعت باد است، درحالی‌که در هند (Chattopadhyay and Hulme, 1997) رطوبت نسبی و در چین (Liu et al., 2012) فشار بخار آب می‌باشد. بنابراین با دانستن اهمیت عوامل تأثیرگذار و تأثیر ترکیبی عوامل می‌توان به‌طور دقیق‌تر تغییرات مقدار ET₀ را تحلیل کرد (Su et al., 2015). Liang et al., (2010) گزارش کردند که عوامل اقلیمی نظیر میانگین و حداکثر دمای هوا، رطوبت نسبی و ساعت آفتابی^۱ از مهم‌ترین متغیرهای اقلیمی مؤثر بر مقدار ET₀ بوده و با داشتن این متغیرها، مقدار ET₀ برای دوره‌های آینده قابل محاسبه و پیش‌بینی خواهد بود. Seifi et al., (2011) در پیش‌بینی مقدار ET₀ در ایستگاه سینوپتیک کرمان با استفاده از مدل ترکیبی رگرسیون چندگانه و تحلیل مؤلفه‌های اصلی نشان دادند که متغیرهای شدت تابش، رطوبت نسبی، ساعات آفتابی، دمای حداقل و دمای حداکثر برای برآورد مقدار ET₀ از اهمیت بیشتری نسبت به سایر متغیرها برخوردار است. مقدار ET₀ ماهیت غیرخطی، پویا و پیچیده دارد، و استخراج فرمولی که بتواند تمام فرآیندهای فیزیکی درگیر با آن را نشان دهد دشوار است (Wen et al., 2015). بنابراین، مدل‌های یادگیری ماشین به دلیل ظرفیت بالایی که دارند برای رسیدگی به مسائل پیچیده و غیرخطی، جایگزین مناسبی برای تکنیک‌های سنتی هستند (Ferreira et al., 2019). مدل‌های یادگیری ماشین می‌توانند برای داده‌های منطقه‌ای که تلفیقی از داده‌های به دست آمده از چندین ایستگاه هواشناسی در یک منطقه می‌باشد، توسعه پیدا کنند. اگرچه مدل‌های یادگیری ماشین از ظرفیت تعمیم بالایی برخوردار هستند، اما ممکن است عملکرد این مدل‌ها به دلیل غیرخطی بودن و شباهت کم بین داده‌های مورد استفاده برای آموزش، کاهش یابد (Shiri et al., 2014). در دهه‌های اخیر، پژوهش‌هایی برای تخمین مقدار ET₀ با استفاده از مدل‌های یادگیری ماشین انجام شده است (Kisi et al., 2017; Alizamir, 2018).

³ Support Vector Mashine¹ Sunshine² Artificial Neural Networks

روش فائو-۵۶

مقدار ET_0 برآورد شده از سامانه نیاز آبی گیاهان بر اساس روش پنمن-مانتیت فائو ۵۶ بود، که به عنوان مقدار واقعی برای آموزش و آزمون روش های مختلف داده کاوی در نظر گرفته شد. معادله ۱ به عنوان معتبرترین روش جهت تخمین مقدار ET_0 می باشد (Allen et al., 1998; Shiri, 2017). در روش پنمن-مانتیت فائو ۵۶ تخمین مقدار ET_0 در مقیاس زمانی روزانه محاسبه شده است (Allen et al., 1998).

$$ET_0 = \frac{0.408D(R_n - G) + \gamma \left(\frac{900}{T + 273} \right) U_2 (e_s - e_a)}{D + \gamma(1 + 0.34U_2)} \quad (1)$$

ET_0 تبخیر تعرق مرجع ($mm \ d^{-1}$)، D شیب منحنی فشار بخار R_n ($kpa^\circ C^{-1}$)، G جریان حرارتی خاک ($MJ \ m^{-2} \ d^{-1}$)، U_2 سرعت باد در ارتفاع ۲ متر $e_s - e_a$ ($m \ s^{-1}$)، T دمای متوسط در ارتفاع ۲ متری ($^\circ C$)، γ ضریب رطوبتی ($kpa^\circ C^{-1}$)، عوامل مورد استفاده در این پژوهش شامل میانگین دمای هوای روزانه، میانگین رطوبت نسبی، متوسط سرعت وزش باد^۴ در ارتفاع دو متری و ساعت آفتابی می باشد.

توصیفی از روش های مدل سازی

مدل سازی ET_0 با سه روش شبکه عصبی مصنوعی (ANNs)، ماشین بردار پشتیبان (SVM) و جنگل تصادفی (RF) انجام شد. روش ANN استفاده شده در این پژوهش شامل دو نوع شبکه عصبی پرسپترون چند لایه (MLP)^۵ و تابع پایه شعاعی (RBF)^۶ می باشد. اساس شبکه عصبی مرکب (هر دو نوع شبکه) توسط Picton (2000) و Hagan et al. (1996) ارائه شده است. برای کاهش خطا در مدل سازی، شبکه عصبی در ۱۰۰ تکرار تنظیم شد. میانگین ۵۰ خروجی از ۱۰۰ تکرار که خطای کمتری دارند، به عنوان خروجی نهایی برای تخمین ET_0 انتخاب شد. اساس روش SVM استفاده شده در این پژوهش، توسط Vapnik (2000) ارائه شده است. روش SVM دارای چهار نوع کرنل که شامل کرنل خطی، چند جمله ای، شعاعی و حلقوی است، که نوع کرنل استفاده شده در این پژوهش هسته تابع پایه شعاعی (RBF) که شامل

روش های جدید با دقت و اعتبار بالا دارد. هدف از این پژوهش برآورد مقدار ET_0 با تحلیل اثر چهار عامل عمده هواشناسی شامل میانگین دمای هوا، میانگین رطوبت نسبی، ساعت آفتابی و سرعت باد، با کاربرد سه روش داده کاوی (ANNs، RF^۱ و SVM) می باشد. در نهایت، سهم هر یک از عوامل هواشناسی بر ET_0 به صورت کمی برآورد شد و با مقدار ET_0 حاصل از سامانه نیاز آبی گیاهان مقایسه گردید.

مواد و روش ها

موقعیت محدوده مورد مطالعه

مساحت استان البرز ۵۸۳۳ کیلومتر مربع می باشد. این استان اقلیم متنوعی دارد، به گونه ای که از اقلیم بیابانی در قسمت های جنوبی شروع و تا اقلیم های نیمه مرطوب و مرطوب در قسمت های شمالی استان ادامه پیدا می کند. ایستگاه های هواشناسی مورد مطالعه در این تحقیق شامل کرج، اشتهارد، مشکین دشت، هشتگرد و طالقان بودند و داده های استفاده شده در این مطالعه از ایستگاه های سینوپتیک ثبت شده در سامانه نیاز آبی گیاهان موسسه تحقیقات خاک و آب کشور^۲ استفاده شده است. مشخصات ایستگاه ها در جدول ۱ آورده شد. داده های مورد استفاده تنها برگرفته از اطلاعات ایستگاه های هواشناسی منطقه بوده که فاقد داده های لایسیمیتری می باشند و مربوط به دوره ده ساله (۱۳۸۹-۱۳۹۹) بود. از هر ایستگاه به تعداد ۴۰۱۸ داده در دسترس بود که مجموع داده های استفاده شده در پنج ایستگاه برابر با ۲۰۰۹۰ است.

جدول ۱- مشخصات ایستگاه های سینوپتیک استان البرز

Table 1- Details of synoptic stations of Alborz province

station	Elev.	Lat.	Lon.	Climate*
Karaj	1293	35.80	50.93	Semi-arid
Eshtehard	1191	35.72	50.37	Arid
MeshkinDasht	1200	35.76	50.96	Arid
Hashtgerd	1613	36.00	50.74	Semi-arid
Taleghan	1857	36.18	50.77	Semi-arid

**Elev: Elevation (m), Lat: Latitude ($^\circ N$), Lon: ($^\circ E$), Climate based on De Martonne aridity index

برآورد مقدار ET_0 سامانه نیاز آبی گیاهان

مقدار ET_0 در این پژوهش از سامانه نیاز آبی گیاهان موسسه تحقیقات خاک و آب استخراج شده است.

⁴ Wind Speed

⁵ Multilayer perceptron's

⁶ Radial basis function

¹ Random Forest

² <http://www.swri.ir/en-S/DouranPortal/1/page/Home>

³ FAO-56 Penman-Monteith

کم‌برآوردی را در تخمین مقدار ET_0 نشان می‌دهد (Gopinathan, 1988). از NRMSE به‌عنوان شاخص خطای پیش‌بینی مدل استفاده می‌شود. بنابراین اگر مقدار NRMSE کمتر از ۱۰٪ باشد نتیجه تخمین عالی، بین ۱۰ تا ۲۰٪، خوب؛ اگر بین ۲۰ تا ۳۰٪، قابل قبول؛ و برای مقادیر بیشتر از ۳۰، ضعیف می‌باشد (Sandhu and Irmak., 2020). شاخص d شاخص توافق ویلموت (۱۹۸۲) می‌باشد که از دامنه صفر (بدون توافق) تا یک (توافق کامل) می‌باشد که میزان درستی و صحت برآوردها را نشان می‌دهد (Willmott et al., 2012). همچنین مقدار EF از منفی بی‌نهایت برای بدترین برآورد تا یک برای تطابق کامل تغییر کرده و کفایت و کارایی مدل را نشان می‌دهد و بالای ۰/۵ عملکرد قابل قبول مدل را نشان می‌دهد (Karimipour and Banitalebi, 2020).

نتایج و بحث

توصیف آماری

میانگین آماری متغیرهای مذکور و همچنین مقدار ET_0 حاصل از سامانه نیاز آبی گیاهان در جدول ۲ آورده شده است. تمامی داده‌های ورودی و خروجی در پنج ایستگاه منتخب در این تحقیق، دارای دامنه تغییرات بالایی (اختلاف بین کمترین و بیشترین مقدار) می‌باشند (جدول ۲). هنگامی که تغییرات داده‌های ورودی و خروجی زیاد باشد، نشان دهنده این است که قابلیت تعمیم‌پذیری^۶ مدل افزایش و هنگامی که تغییرپذیری کم است، قابلیت تعمیم‌پذیری آن نیز کاهش می‌یابد. به عبارت دیگر افزایش قابلیت تعمیم‌پذیری مدل یعنی مدل می‌تواند روی داده‌های جدید نیز عملکرد خوبی داشته باشد. به‌منظور نشان دادن پراکندگی نسبی داده‌ها به میانگین، از ضریب تغییرات (CV) استفاده بهره‌گیری از توابع انتقالی در محدوده داده‌های به‌کار گرفته شده برای آموزش با دقت بیشتری خواهد بود (Ungaro et al., 2005). محدوده تغییرات و میانگین متغیرها به دلیل اینکه انتخاب داده‌های آموزش و آزمون به‌صورت تصادفی انجام شد، مشابه است.

عامل گاما (γ) می‌باشد. روش SVM دارای عوامل تنظیم کننده شامل C و ν هسته است. روش RF یک روش جدید و قدرتمند است که پیشرفت‌های فراوانی را در فنآوری داده‌کاوی داشت. اساس این روش توسط (Breiman, 2001) ارائه شده است. در این پژوهش برای مدل‌سازی ET_0 از نرم‌افزار STATISTICA 12 استفاده شد.

گام‌های ساخت توابع انتقالی

فرضیه‌های پیش‌پردازش و رگرسیون داده‌ها، از جمله تشخیص داده‌های پرت‌ها^۱، آزمون نرمال کردن داده‌ها، بررسی چندخطی بودن برای همه عوامل انجام شد (Berry, 1993). برای حذف داده‌های پرت از روش «دامنه میان چارکی (IQR)»^۲ (Seo, 2006) استفاده شد. فرض صحت رگرسیون خطی، بررسی هم راستایی بین متغیرهای مستقل است. همچنین در تمام مدل‌ها این مسئله از طریق آماره عامل تورم واریانس (VIF)^۳ بررسی شد. متغیرهایی که VIF آن‌ها کمتر از ۱۰ باشد مشکل هم‌راستایی نخواهند داشت (Hocking, 2013). که در این تحقیق VIF تمام ورودی‌ها با SPSS.26 بررسی شد. در این پژوهش از مجموع داده‌ها برای هر ایستگاه (۴۰۱۸)، دوسوم داده‌ها (۲۶۷۸) برای مرحله آموزش و یک‌سوم (۱۳۴۰) از داده‌ها برای مرحله آزمون انتخاب شده است.

معیارهای ارزیابی دقت و قابلیت اعتماد مدل‌ها

برای بررسی دقت و قابلیت اعتماد توابع ایجادشده از معیارهای آماره‌های ضریب تعیین (R^2)، مجذور میانگین مربعات خطا ($RMSE^4$)، میانگین انحراف خطای (MBE^5)، مجذور میانگین خطای نرمال شده (NRMSE) و ضریب کارایی مدل (EF) بهره‌گیری شد. از لحاظ آماری حداقل مقدار RMSE برابر صفر بوده و زمانی است که مقادیر پیش‌بینی شده با مقادیر اندازه‌گیری شده برابر باشند. هرچقدر مقدار این آماره کمتر باشد، طبیعتاً مدل برآورد کننده دارای دقت بیشتری خواهد بود (Adab et al., 2020). MBE عمدتاً برای نشان دادن اریب در مدل مورد استفاده قرار می‌گیرد و مقدار مثبت آن نشان‌دهنده بیش برآوردی و مقدار منفی آن

⁴ Root mean square error

⁵ Mean Bias Error

⁶ Generalization

¹ Outlier data

² InterQuartile Range

³ Variance Inflation Factor

ET₀ در همه ایستگاه‌ها با تابش خورشید، سرعت باد، میانگین دمای هوا همبستگی مثبت و معنی‌داری دارد و با میانگین رطوبت نسبی همبستگی منفی و معنی‌داری را نشان می‌دهد با افزایش تابش خورشید و بالا رفتن دمای هوا و همچنین افزایش سرعت باد، مقدار ET₀ افزایش یافته و با افزایش رطوبت هوا مقدار ET₀ کاهش می‌یابد (You et al., 2020).

همبستگی بین عوامل ورودی و خروجی مدل همبستگی به بررسی ارتباط دو یا چند متغیر می‌پردازد و ضریب آن را محاسبه می‌نماید. متداول‌ترین نمایه برای بیان همبستگی بین متغیرهای دارای اندازه‌های مقیاس فاصله‌ای و نسبی استفاده از ضریب همبستگی پیرسون^۱ است که در این پژوهش از آن استفاده شده است. با توجه به جدول ۳، مقدار

جدول ۲- عوامل آماری متغیرهای مورد مطالعه

Table 2 - Statistical factors of the studied variables

Station	factors	Minimum	Maximum	Mean	Standard Deviation	CV(%)	Skewness
Hashtgerd (n=4018)	T _{mean}	-9.6	33.3	14.5	9.3	64.3	-0.07
	RH _{mean}	9	100	48.09	0.32	42.49	0.42
	W	0.22	4.48	1.55	0.51	33.00	0.61
	S	0.00	14.45	9.09	3.48	38.24	-0.87
	ET ₀	0.39	9.12	3.68	2.16	58.84	0.25
Karaj (n=4018)	T _{mean}	-7.6	35.9	16.0	9.3	58.3	-0.09
	RH _{mean}	8.5	100	49.07	0.28	35.67	0.42
	W	0.22	5.45	1.54	0.56	36.17	0.81
	S	0.00	13.50	8.42	3.54	42.05	-1.00
	ET ₀	0.33	9.08	3.77	2.15	57.10	0.24
Taleghan (n=4018)	T _{mean}	-10.9	31.3	12.0	9.0	74.8	-0.07
	RH _{mean}	5.00	97	53.73	0.25	29.62	0.26
	W	0.00	4.03	1.46	0.50	34.57	0.30
	S	0.00	14.45	9.73	3.36	34.56	-1.29
	ET ₀	0.41	8.51	3.51	2.10	59.73	0.21
MeshkinDasht (n=4018)	T _{mean}	-7.6	35.9	15.9	9.3	58.3	-0.09
	RH _{mean}	0.00	100	49.24	0.27	35.34	0.40
	W	0.00	4.18	1.49	0.53	35.24	0.42
	S	0.00	13.50	8.38	3.53	42.14	-0.98
	ET ₀	0.36	8.98	3.72	2.12	56.91	0.23
Eshtehard (n=4018)	T _{mean}	-7.5	36.0	15.9	9.5	59.7	-0.10
	RH _{mean}	0.00	100	48.19	0.30	39.38	0.27
	W	0.00	4.58	1.58	0.62	39.50	0.52
	S	0.00	13.84	8.59	3.43	39.90	-1.28
	ET ₀	0.41	10.74	3.88	2.23	57.67	0.23
Alborz province (n=20090)	T _{mean}	-10.9	36.0	14.9	9.4	63.3	-0.07
	RH _{mean}	0.00	100	49.66	0.13	36.70	0.31
	W	0.00	5.45	1.53	0.55	35.95	0.57
	S	0.00	14.45	8.84	3.51	39.65	-1.05
	ET ₀	0.33	10.74	3.71	2.16	58.12	0.24

** T_{mean}: Mean temperature (°C), RH_{mean}: Mean relative humidity (%), W: Wind Speed (ms⁻¹), S: Sunshine (h), ET₀ (mm day⁻¹)

جدول ۳- رابطه همبستگی بین عامل‌های هواشناسی با مقدار ET₀Table3- Correlation between meteorological factors and ET₀

Stations		Mean Temperature (°C)	Mean relative Humidity (%)	Wind Speed (ms ⁻¹)	Sunshine (h)
Eshtehard	ET ₀	0.92**	-0.75**	0.22**	0.57**
Taleghan	ET ₀	0.93**	-0.59**	0.54**	0.53**
Karaj	ET ₀	0.93**	-0.70**	0.054**	0.70**
MeshkinDasht	ET ₀	0.93**	-0.70**	0.04**	0.70**
Hashtgerd	ET ₀	0.93**	-0.70**	0.24**	0.66**
Alborz province	ET ₀	0.92**	-0.69**	0.25**	0.61**

* و ** به ترتیب همبستگی معنی‌دار در سطوح ۵ و ۱ درصد را نشان می‌دهند.

¹ Pearson coefficient of correlation

ارزیابی عملکرد روش‌های مختلف در برآورد ET_0

در این پژوهش مقدار ET_0 روزانه در پنج ایستگاه سینوپتیک استان البرز با سه روش ANNs، RF و SVM برآورد شد (جدول ۷) و سپس عملکرد روش باهم مورد مقایسه قرار گرفت.

ویژگی‌های روش ANNs

با توجه به جدول ۴، ANNs بکار برده شده در این پژوهش دارای سه لایه داده ورودی، میانی و خروجی بودند. از توابع فعال‌سازی Sine، Tanh، Logistic، Gaussian، Exponential و Identity در لایه‌های میانی و خروجی استفاده شد (جدول ۴). از ۵۰ شبکه انتخاب شده در هر مدل ایجاد شده از هر دو نوع RBF و MLP هستند، اما در هر پنج ایستگاه، تعداد MLP استفاده شده در ایجاد مدل بیشتر از RBF است. نتایج نشان داد که شبکه عصبی MLP نسبت به RBF در برآورد مقدار ET_0 از عملکرد بهتری برخوردار است. در این پژوهش تعداد سلول‌ها (نورون‌های) پنهان شبکه عصبی از ۳ تا ۳۰ متغیر است. Exponential و Identity و به ترتیب تابع فعال‌سازی غالب در لایه پنهان و خروجی بوده و بهترین تابع برای توصیف روابط بین متغیرهای ورودی و ET_0 می‌باشند.

ویژگی‌های روش SVM

در این پژوهش از رگرسیون بردار پشتیبان نوع دوم استفاده شده است. مشخصات روش SVM که در این مطالعه استفاده شده، در جدول ۵ آمده است.

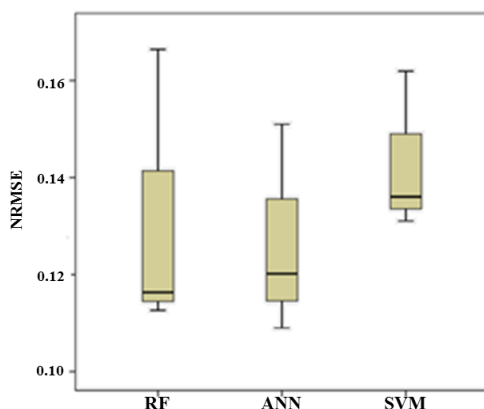
ویژگی‌های روش RF

یکی از مهم‌ترین ویژگی‌های روش RF، ساده‌سازی مدل از طریق تعیین اهمیت عامل^۱ (تخمین‌گرها) می‌باشد. نتایج اهمیت عامل‌ها برای توابع مختلف که با روش RF به دست آمده، در جدول ۶ ارائه شده است. این روش عامل‌ها را بر اساس سهم عامل در دقت تخمین و کاهش خطا، مقدار ۰ تا ۱ می‌دهد (Sedaghat et al., 2022). در جدول ۶ مشاهده می‌شود در همه ایستگاه‌ها به جز طالقان میانگین دمای هوا بیشترین سهم را در افزایش دقت تخمین ET_0 یا به عبارت

دیگر بیشترین سهم را در کاهش خطای تخمین ET_0 دارد. هم‌چنین سرعت باد کمترین سهم را در افزایش دقت تخمین مقدار ET_0 دارد. اما در طالقان نتایج عکس نسبت به سایر مناطق مشاهده شد. به طوری که میانگین دمای هوا کمترین و میانگین رطوبت نسبی بیشترین سهم را در افزایش دقت تخمین ET_0 دارد. احتمالاً به دلیل اینکه منطقه طالقان ارتفاع از سطح دریای بالاتری نسبت به سایر مناطق مورد بررسی دارد (جدول ۱)، با افزایش ارتفاع سبب کاهش دما و به تبع آن کاهش مقدار ET_0 و افزایش رطوبت نسبی می‌شود (Raziei et al., 2005). Liang et al., (2010) گزارش کردند که عوامل اقلیمی نظیر میانگین و حداکثر دمای هوا، رطوبت نسبی و ساعات آفتابی از مهم‌ترین متغیرهای اقلیمی مؤثر بر مقدار ET_0 بوده و با داشتن این متغیرها، مقدار ET_0 برای دوره‌های آینده قابل محاسبه و پیش‌بینی خواهد بود. بنابراین می‌توان گفت در مناطق با ارتفاع زیاد از سطح دریا، میانگین رطوبت نسبی و در ارتفاع‌های کمتر میانگین دمای هوا مهم‌ترین مؤلفه در برآورد مقدار ET_0 می‌باشد. با توجه به جدول ۷، نتایج مقایسه سه روش در برآورد مقدار ET_0 ، نشان داد که در هر دو مرحله آموزش و آزمون دقت روش ANNs نسبت به دو روش RF و SVM به طور نسبی بهتر بود. هم‌چنین مشاهده شد نتایج دو روش ANNs و RF بسیار به هم نزدیک می‌باشد. مقایسه میانگین نتایج حاصل از سه روش براساس آماره NRMSE در مرحله آزمون، در شکل ۱ ارائه شده است. با توجه به شکل ۱ مشاهده می‌شود در مرحله آزمون میانگین خطای روش ANNs از دو روش دیگر کمتر و میانگین خطای روش SVM نسبت به دو روش دیگر بیشتر می‌باشد. همان‌طور که در شکل ۱ مشاهده می‌شود تفاوت معنی‌داری بین میانگین خطای دو روش ANNs و RF وجود ندارد. مقایسه میانگین نتایج بین ایستگاه‌های مختلف بر اساس آماره NRMSE با سه روش در شکل ۲ نشان داده شده است. نتایج نشان می‌دهد (شکل ۲) که دقت سه روش به کار رفته در این پژوهش در کرج و مشکین نسبت به سایر مناطق استان بیشتر و میانگین خطای روش‌های بکار رفته در این دو ایستگاه نسبت به بقیه کمتر می‌باشد. اما میانگین خطای روش‌های بکار رفته در

¹ Important parameter

نامناسب برای سازگاری با شرایط آتی باعث خسارات جبران‌ناپذیر خواهد شد (Fazeli Khiavi et al., 2020).



شکل ۱- مقایسه NRMSE سه روش در برآورد ET_0 در مرحله آزمون

Figure 1- Comparison the NRMSE of three methods for estimating of ET_0 in the test steps

طالقان نسبت به سایر ایستگاه‌ها بیشتر و در نتیجه دقت روش‌های بکار رفته در این ایستگاه پایین‌تر است. با توجه به شکل ۳ مشاهده می‌شود که با افزایش ارتفاع از سطح دریا، میانگین خطا در تخمین مقدار ET_0 افزایش می‌یابد. ارتفاع از سطح دریا ایستگاه‌های اشتهارد برابر با ۱۱۹۱، مشکین‌دشت ۱۲۰۰ و کرج برابر با ۱۲۹۳ متر از سطح دریا است (جدول ۱). اشتهارد کمترین ارتفاع از سطح دریا را دارد. ولی میانگین خطای آن کمی بیشتر از دو ایستگاه کرج و مشکین‌دشت می‌باشد. شرایط منطقه‌ای اشتهارد به صورت دشت می‌باشد، در نتیجه در دشت‌ها با افزایش تابش خورشید میزان دما افزایش یافته و بر میزان ET_0 اثر خواهد گذاشت، که این موضوع با پژوهش‌های (Goudarzi et al., 2018) و Mehrazar et al., (2018) مطابقت دارد. افزایش دما، مقدار ET_0 و تغییرات زمانی بارش در دشت‌ها باعث کاهش منابع آب و افزایش نیاز آبی گیاهان شده و در صورت تصمیم‌گیری

جدول ۴- ویژگی‌های روش ANNs بکار رفته در این پژوهش

Table 4 - Characteristics of the ANNs method used in this research

Stations		Eshtehard	Karaj	Taleghan	Hashtgerd	MeshkinDasht	Alborz province
Network type	MLP	41	46	44	43	47	49
	RBF	9	4	6	7	3	1
Number of hidden cells	Range	3-30	3-27	3-30	3-30	3-28	3-28
Output layer function	Tanh	9	10	10	11	6	9
	Identity	19	15	13	20	9	15
	Sine	5	8	4	8	11	7
	Exponential	10	9	15	7	14	10
	Logistic	7	8	8	4	10	9
Hidden layer function	Gaussian	9	4	6	7	3	1
	Identity	4	3	6	8	3	2
	Sine	11	9	1	5	8	7
	Exponential	10	14	14	8	11	11
	Tanh	7	12	9	13	15	15
	Logistic	9	8	14	9	10	14

جدول ۵- مشخصات روش SVM به کار برده شده در این پژوهش

Table 5 - Characteristics of the SVM method used in this research

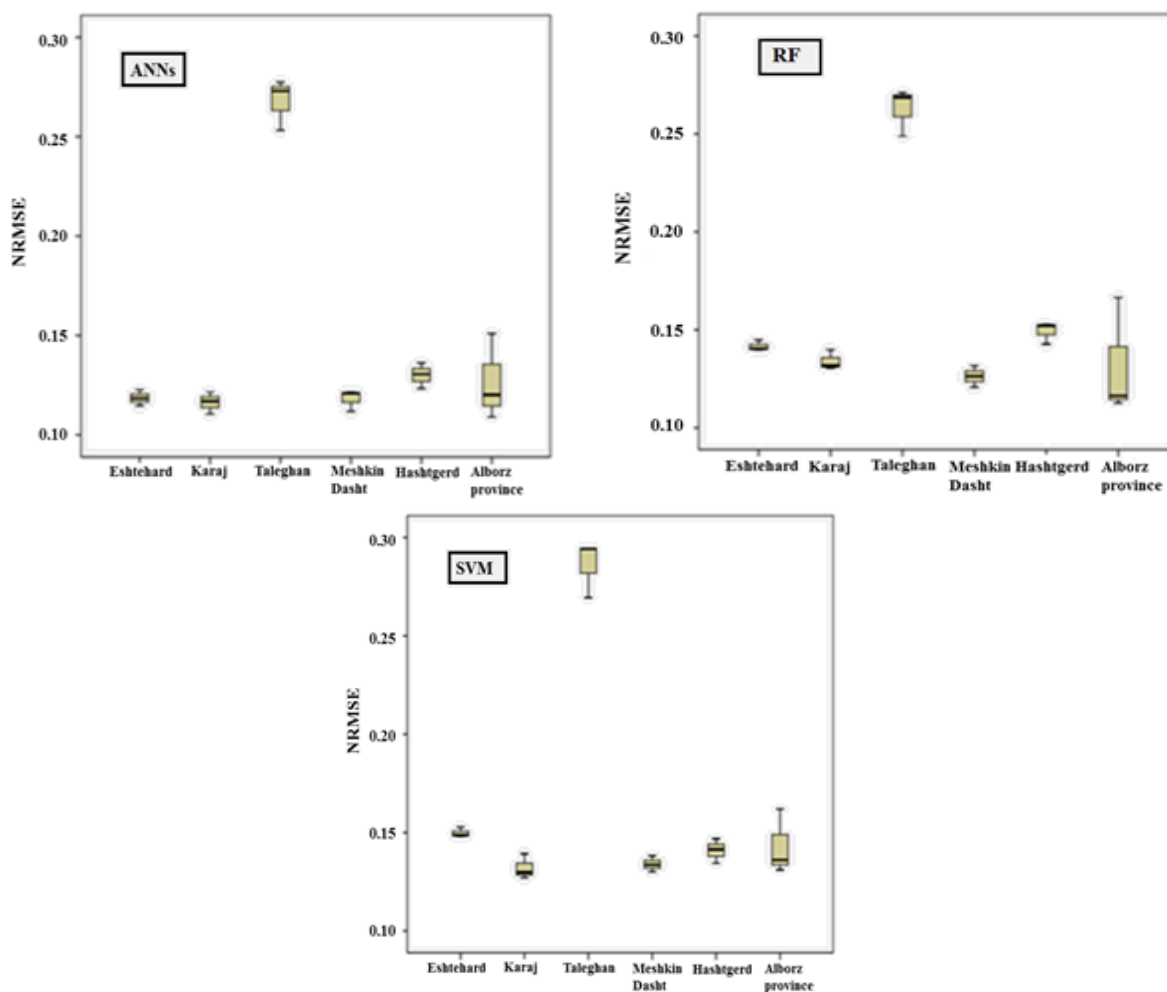
	γ^*	nu	C	Number of support vectors	Cross-validation error
Eshtehard	0.25	0.2	9	352	0.011
Karaj	0.25	0.1	2	804	0.012
Taleghan	0.25	0.21	2	876	0.061
Hashtgerd	0.27	0.3	10	393	0.016
MeshkinDasht	0.24	0.2	10	410	0.013
Alborz province	0.2	0.1	2	969	0.011

* γ ، nu و C پارامترهای تنظیم کننده هسته در روش SVM می‌باشند.

جدول ۶- اهمیت عامل به دست آمده با روش RF

Table 6- Importance of the factor obtained by RF method

factor	Mean Temperature (°C)	Sunshine (h)	Mean relative humidity (%)	Wind Speed (ms ⁻¹)
Eshtehard	1	0.51	0.59	0.20
Karaj	1	0.74	0.60	0.13
Taleghan	0.75	0.97	1	0.91
Hashtgerd	1	0.67	0.63	0.16
MeshkinDasht	1	0.60	0.55	0.12
Alborz province	1	0.60	0.58	0.18



شکل ۲- میانگین NRMSE بین ایستگاه‌های مختلف در سه روش بکار رفته در این پژوهش
Figure 2- Mean NRMSE of different stations in the three methods used in this study

می‌باشد ولی در اشتهارد که ارتفاع کمتری نسبت به مشکین‌دشت و کرج دارد (جدول ۱) بیشتر می‌باشد. احتمالاً به خاطر این شرایط، میانگین خطای تخمین مقدار ET_0 در اشتهارد نسبت به مشکین‌دشت و کرج بیشتر شد. اما با توجه به کمتر بودن ارتفاع ایستگاه مشکین‌دشت نسبت به کرج، تا

با توجه به جدول ۲، ضریب تغییرات دما و مقدار ET_0 در اشتهارد به ترتیب برابر ۵۹/۷۱ و ۵۷/۶۷، در مشکین‌دشت برابر ۵۸/۳۶ و ۵۶/۹۱ و در کرج برابر ۵۸/۳۳ و ۵۷/۱ می‌باشد. بنابراین تغییرات دما و مقدار ET_0 در دو ایستگاه مشکین‌دشت و کرج که تقریباً ارتفاع از سطح دریای یکسانی دارند، برابر

NRMSE به ترتیب برای آموزش برابر با ۰/۹۶، ۰/۴۲ و ۰/۱۱، و برای آزمون برابر با ۰/۹۵، ۰/۴۵ و ۰/۱۲ می‌باشد. در روش SVM مقادیر آماره‌های RMSE، R^2 و NRMSE به ترتیب برای آموزش برابر با ۰/۵، ۰/۹۴ و ۰/۱۴، و برای آزمون برابر با ۰/۵۱، ۰/۹۴ و ۰/۱۴ می‌باشد (جدول ۷).

حدودی میانگین خطای تخمین در مشکین‌دشت نسبت به کرج کمتر می‌باشد. در مشکین‌دشت، در روش ANNs مقادیر آماره‌های RMSE، R^2 و NRMSE به ترتیب برای آموزش برابر با ۰/۹۶، ۰/۴۱ و ۰/۱۱، و برای آزمون برابر با ۰/۹۶، ۰/۴۲ و ۰/۱۱ می‌باشد. در روش RF مقادیر آماره‌های RMSE، R^2 و

جدول ۷- نتایج روش‌های مختلف در برآورد مقدار ET_0 در ایستگاه‌های هواشناسیTable 7 - The results of different methods used for estimating of ET_0 in the meteorological stations

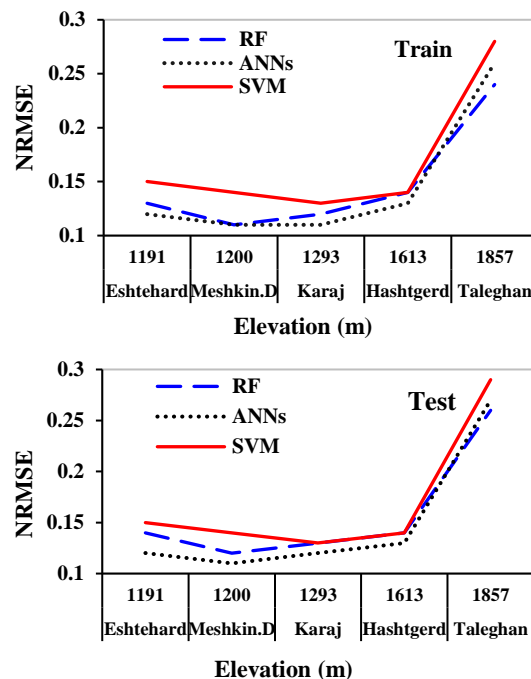
		NRMSE (mm day ⁻¹)	R^2	RMSE (mm day ⁻¹)	d	EF	MBE
Train (n=2680)							
Karaj	RF	0.12	0.96	0.43	0.98	0.96	-0.002
	ANNs	0.11	0.96	0.42	0.99	0.96	-0.0004
	SVM	0.13	0.95	0.47	0.97	0.95	0.02
Eshtehard	RF	0.13	0.95	0.50	0.99	0.95	0.0005
	ANNs	0.12	0.96	0.45	0.99	0.96	0.003
	SVM	0.15	0.95	0.58	0.98	0.93	0.151
Taleghan	RF	0.24	0.85	0.84	0.95	0.85	-0.01
	ANNs	0.26	0.81	0.92	0.94	0.81	-0.005
	SVM	0.28	0.79	0.98	0.93	0.79	0.02
MeshkinDasht	RF	0.11	0.96	0.42	0.99	0.96	-0.001
	ANNs	0.11	0.96	0.41	0.99	0.96	-0.0002
	SVM	0.14	0.94	0.50	0.984	0.94	-0.005
Hashtgerd	RF	0.14	0.95	0.51	0.98	0.95	0.0009
	ANNs	0.13	0.95	0.48	0.99	0.95	-0.001
	SVM	0.14	0.94	0.52	0.98	0.94	0.080
Alborz province	RF	0.14	0.94	0.51	0.98	0.94	-0.002
	ANNs	0.13	0.95	0.48	0.99	0.95	0.003
	SVM	0.15	0.94	0.56	0.98	0.93	0.002
Test (n=1338)							
Karaj	RF	0.13	0.95	0.47	0.97	0.95	0.01
	ANNs	0.12	0.95	0.44	0.99	0.95	0.01
	SVM	0.13	0.94	0.49	0.97	0.94	0.03
Eshtehard	RF	0.14	0.94	0.55	0.98	0.94	0.007
	ANNs	0.12	0.96	0.46	0.99	0.96	0.009
	SVM	0.15	0.94	0.58	0.98	0.93	0.153
Taleghan	RF	0.26	0.80	0.92	0.94	0.80	0.04
	ANNs	0.27	0.79	0.94	0.94	0.80	0.043
	SVM	0.29	0.76	1.00	0.93	0.77	0.07
MeshkinDasht	RF	0.12	0.95	0.45	0.99	0.95	0.019
	ANNs	0.11	0.96	0.42	0.99	0.96	0.017
	SVM	0.14	0.94	0.51	0.98	0.94	0.005
Hashtgerd	RF	0.14	0.94	0.55	0.98	0.93	-0.0124
	ANNs	0.13	0.95	0.48	0.98	0.95	-0.010
	SVM	0.14	0.94	0.52	0.98	0.94	0.069
Alborz province	RF	0.13	0.95	0.50	0.98	0.95	0.08
	ANNs	0.13	0.95	0.48	0.98	0.95	0.09
	SVM	0.14	0.95	0.54	0.98	0.94	0.08

برابر با ۰/۹۲، ۰/۸۰ و ۰/۲۶ می‌باشد. در روش SVM مقادیر آماره‌های RMSE، R^2 و NRMSE به ترتیب برای آموزش برابر با ۰/۹۸، ۰/۷۹ و ۰/۲۸، و برای آزمون برابر با ۰/۷۶، ۰/۱۰۰۴ و ۰/۲۹ می‌باشد (جدول ۷). این نتایج توانایی روش‌های یادگیری ماشین در حل مسائل پیچیده را نشان می‌دهد،

در طالقان، در روش ANNs مقادیر آماره‌های RMSE، R^2 و NRMSE به ترتیب برای آموزش برابر با ۰/۹۲، ۰/۸۱ و ۰/۲۶، و برای آزمون برابر با ۰/۹۴، ۰/۷۹ و ۰/۲۷ می‌باشد. در روش RF مقادیر آماره‌های RMSE، R^2 و NRMSE به ترتیب برای آموزش برابر با ۰/۸۴، ۰/۸۵ و ۰/۲۴، و برای آزمون

ET₀ در هر منطقه می‌باشد (Poormohammadi et al., 2010). در اثر تغییرات اقلیمی، مقدار ET₀ نه تنها بر اساس دمای هوا بلکه بر اساس سایر متغیرهای آب و هوایی دست خوش تغییرات قرار می‌گیرد (Fan et al., 2016; Yang et al., 2017). بنابر این، این متغیرها نیز به شدت تحت تأثیر شرایط منطقه‌ای مانند توپوگرافی و ارتفاع قرار می‌گیرند (Alexandris and Proutsos, 2020). بنابراین، مقدار ET₀ واکنش‌های متفاوتی به متغیرهای آب و هوایی در اقلیم‌های مختلف نشان می‌دهد. نظر به اینکه ارتفاع نقش مهمی در تعدیل تغییرات اقلیمی در مناطق کوهستانی دارد، بعضی از محققان در رابطه با همبستگی دمای هوا و ارتفاع در مناطق کوهستانی گزارش‌هایی ارائه کرده‌اند (You et al., 2020; Cai et al., 2017). پژوهشی نشان داد که مقدار ET₀ بالا (بیش از ۱۲۰۰ میلی‌متر) عمدتاً در ارتفاع پایین با میانگین کمتر از ۱۰۰۰ متر اتفاق می‌افتد (Luo et al., 2021). نتایج تحقیق حاضر نشان داد که اثر ارتفاع بر تغییرات مقدار ET₀ تحت شرایط آب و هوایی متغیر، نامشخص است. نتایج پژوهش حاضر با نتایج Liu and Zhang., (2013) که ET₀ و ETpan با ضریب تعیین (R²) بالای ۰/۸۱ بدست آورده‌اند مطابقت دارد. که نشان دهنده این موضوع می‌باشد که در این تحقیق مقدار ET₀ برآورد شده به دلیل انتخاب درست عوامل موثر بر تخمین مقدار ET₀، قابل اعتماد است. با توجه به جدول ۷، مقادیر آماره MBE در مرحله آموزش و آزمون اکثراً مقادیر کوچک و نزدیک به صفر می‌باشد. بنابراین با توجه به کم بودن مقادیر آماره MBE، نمی‌توان گفت مدل کم‌برآوردی یا بیش برآوردی قابل توجهی داشته است و در نتیجه می‌توان از آن چشم‌پوشی کرد (Singh et al., 2020). آماره EF به جز در ایستگاه طالقان که ضعیف‌ترین نتایج را دارد، در سایر ایستگاه‌ها در هر دو مرحله آموزش و آزمون بالای ۰/۹۳ می‌باشد و هم چنین شاخص d هم به جز در ایستگاه طالقان، در سایر ایستگاه‌ها در هر دو مرحله آموزش و آزمون بین ۰/۹۹-۰/۹۷ می‌باشد، که نشان از صحت و کارایی بالای روش‌های استفاده در این پژوهش در تخمین مقدار ET₀ می‌باشد.

همانطور که در چندین مطالعه در مدل‌سازی ET₀ گزارش شده است (Antonopoulos and Antonopoulos, 2017; Yassin et al., 2016). از طرفی فاکتورهای مختلفی بر عملکرد روش‌های یادگیری ماشین دخیل هستند مانند الگوریتم روش ANNs، یا تنظیمات فرآیندهای روش‌های مورد استفاده، می‌باشد (Ferreira et al., 2019).



شکل ۳- تغییرات خطا با افزایش ارتفاع در تخمین ET₀

Figure 3- Error changes with increasing altitude in ET₀ estimation

شکل ۳ نشان می‌دهد با افزایش ارتفاع از ۱۶۰۰ متر به بالا، میانگین خطا به طور قابل توجهی افزایش می‌یابد. علت آن هم احتمالاً به دلیل تغییرات دمای هوا با افزایش رطوبت و تأثیر آن بر مقدار ET₀ می‌باشد. تغییرات تشعشع خورشیدی به عنوان تابعی از عرض جغرافیایی، در فصول مختلف سال متفاوت است. به طوری که در تمامی فصول با افزایش عرض جغرافیایی از میزان تشعشع ورودی به زمین کاسته می‌شود. بنابراین انتظار می‌رود که مقدار ET₀ در عرض‌های جغرافیایی بالاتر، کمتر از عرض‌های جغرافیایی پایین‌تر به دست آید، از جمله دلایل این امر می‌تواند مرتبط با در نظر گرفتن عواملی مانند عرض جغرافیایی و ارتفاع از سطح دریا در محاسبات مقدار ET₀ باشد. این عوامل از کنترل‌کننده‌های مؤثر مقدار

- on artificial intelligence methods. *Water*, 11(6), 1226.
- Allen, R.G., Pereira, L.S., Howell, T.A., Jensen, M.E. 2011. Evapotranspiration information reporting, I. Factors governing measurement accuracy. *Agricultural Water Management*. 98, 899-920.
- Allen, R. G., Pereira, L. S., Raes, D., Smith, M. 1998. Crop evapotranspiration-guidelines for computing crop water requirements- FAO irrigation and drainage paper NO. 56. FAO, Rome.
- Alves, W.B., Rolim, G.D.S., Aparecido, L.E., de, O. 2017. Reference evapotranspiration forecasting by artificial neural networks. *Engenharia Agrícola*, 37, 1116-1125.
- Antonopoulos, V.Z., Antonopoulos, A. V. 2017. Daily reference evapotranspiration estimates by artificial neural networks technique and empirical equations using limited input climate variables. *Computers and Electronics in Agriculture*, 132, 86-96.
- Aytek A. 2008. Co-active neuro-fuzzy inference system for evapotranspiration modeling. *Soft Computing-A Fusion of Foundations, Methodologies and Applications*, 13(7), 691-700.
- Berry, W. D. 1993. *Understanding Regression Assumptions*. Sage Publications, London.
- Breiman, L. 2001. Random forests. *Machine Learn*. 45, 5-32.
- Cai, D., You, Q., Fraedrich, K., Guan, Y. 2017. Spatiotemporal Temperature Variability over the Tibetan Plateau, Altitudinal Dependence Associated with the Global Warming Hiatus. *Journal of Climatology*, 30, 969-984.
- Chattopadhyay, N., Hulme, M. 1997. Evaporation and potential evapotranspiration in India under conditions of recent and future climate change. *Agricultural and Forest Meteorology*, 87, 55-73.
- Dinpashoh, Y., Jhajharia, D., Fakheri-Fard, A., Singh, V.P., Kahya, E., 2011. Trends in reference crop evapotranspiration over Iran. *Journal of Hydrology*, 399, 422-433.
- Fan, J., Wu, L., Zhang, F., Xiang, Y., Zheng, J. 2016. Climate change effects on reference crop evapotranspiration across different climatic zones of China during 1956-2015. *Journal of Hydrology*, 542, 923-937.
- Fazeli Khiavi, A., Salahi, B., Goodarzi, M. 2020. Assessment effects of climate change on changes in potential evapotranspiration in the Moghan Plain by RCPs. *Watershed Engineering and Management*, 12(4), 977-993. (In Farsi).
- Ferreira, L., da, B., Cunha, F. F., de Oliveira, R. A., Fernandes Filho, E. I. 2019. Estimation of reference evapotranspiration in Brazil with limited meteorological data using ANN and SVM-A new approach. *Journal of Hydrology*, 572, 556-570

نتیجه گیری

نتایج عملکرد سه روش داده کاوی، RF، ANNs و SVM برای تخمین مقدار ET_0 با چهار ترکیب ورودی مختلف در پنج ایستگاه سینوپتیک هواشناسی استان البرز نشان داد که دقت تخمین مقدار ET_0 با روش ANNs بالاتر از روش های دیگر بود. اما دقت نتایج روش RF بسیار به نتایج روش ANNs نزدیک می باشد، از طرفی سرعت برآزش روش ANNs نسبت به روش RF بسیار طولانی می باشد که با در نظر گرفتن همه جوانب می توان گفت روش RF برای تخمین مقدار ET_0 هم می توان از رویکرد مناسبی برخوردار باشد. نتایج این پژوهش نشان داد، مقدار ET_0 نه تنها بر اساس دمای هوا بلکه تحت تأثیر سایر مؤلفه های اقلیمی تغییر می کند، هم چنین به شدت تحت تأثیر شرایط منطقه ای مانند توپوگرافی و ارتفاع می باشد. بنابراین میانگین رطوبت نسبی در مناطقی با ارتفاع زیاد از سطح دریا و میانگین دمای هوا در ارتفاع های کمتر، مهم ترین مؤلفه در برآورد مقدار ET_0 می باشد. نتایج نشان داد مقادیر ET_0 سامانه نیاز آبی گیاهان، همبستگی خوبی با مقادیر ET_0 تخمین زده با روش های ANNs و RF دارد. بنابراین، می توان گفت که حتی اگر تمام اطلاعات عامل ها در یک منطقه در دسترس نباشد، می توان از میانگین دما، میانگین رطوبت نسبی، سرعت باد و ساعت آفتابی برای تخمین ET_0 با دقت و قابلیت اعتبار بالا استفاده کرد. نتیجه نهایی اینکه روش های ANNs و RF با استفاده از داده های اقلیمی محدود می توانند در مطالعات آینده برای توسعه مدل ها جهت تخمین مقدار ET_0 با استفاده از الگوریتم های مختلف یادگیری ماشین استفاده شود.

منابع

- Adab, H., Morbidelli, R., Saltalippi, C. Moradian M., Ghalhari, G. A. F. 2020. Machine learning to estimate surface soil moisture from remote sensing data. *Water*, 12 (11), 3223.
- Alexandris, S., Proutsos., N. 2020. How significant is the effect of the surface characteristics on the Reference Evapotranspiration estimates? *Agricultural Water Management*, 237, 106181.
- Allawi, M. F., Binti Othman, F., Afan, H. A., Ahmed, A. N., Hossain, M., Fai, C. M., El-Shafie, A. 2019. Reservoir evaporation prediction modeling based

- temperature. *Hydrological Processes*, 27, 3941–3948.
- Luo, Y., Gao, P., Mu, X. 2021. Influence of meteorological factors on the potential evapotranspiration in Yanhe River Basin, China. *Water*, 13(9), 1222.
- Mehrazar, A., Massah Bavani, A., Mashal, M., Rahimikhoob, H. 2018. Assessment of climate change impacts on agriculture of the Hashtgerd Plain with emphasis of AR5 models uncertainty. *Irrigation Sciences and Engineering*, 41(3), 45-59. (In Farsi).
- Onyutha, C. 2016. Statistical analyses of potential evapotranspiration changes over the period 1930–2012 in the Nile River riparian countries. *Agricultural and Forest Meteorology*, 226, 80–95.
- Picton, P. 2000. *Neural Networks*, 2nd edn. Palgrave, New York.
- Poormohammadi, S., Malekinezhad, H., Rahimian, M. H. 2010. Investigating the role of physiographical factors on temperature-related parameters affecting evapotranspiration (Case study, Yazd province). *Journal of Arid Biome*, 1(2), 9-19. (In Farsi).
- Raziei, T., Daneshkar Arasteh, P., Saghafian, B. 2005. Annual rainfall trend analysis in arid and semi-arid regions of central and eastern Iran. *Water and Wastewater*, 54, 73-81. (In Farsi).
- Salarian, M., Najafi, M., Nagafi, K., Eslamiyan, S. S., Heidari, M. 2014. The most Appropriate Method to Estimate Potential Evapotranspiration in Meteorological Data Scarce Condition in the Warm and Cold Months of the Year (Case Study of Isfahan). *Iranian Journal of Irrigation and Drainage*, 8(1), 62-73., (In Farsi).
- Sandhu, R., Irmak, S. 2020. Performance assessment of Hybrid-Maize model for rainfed, limited and full irrigation conditions. *Agricultural Water Management*, 242, 106402.
- Sedaghat, A., Shahrestani, M. S., Noroozi, A. A., Nosratabad, A. F., Bayat, H. 2022. Developing pedotransfer functions using Sentinel-2 satellite spectral indices and Machine learning for estimating the surface soil moisture. *Journal of Hydrology*, 127423.
- Seifi, A., Mirlatifi, S.M., Riahi, H. 2011. Developing a Combined Model of Multiple Linear Regression-Principal Component and Factor Analysis (MLR-PCA) for Estimation of Reference Evapotranspiration (Case Study, Kerman Station). *Journal of Water and Soil*, 24(6), 1186-1196.
- Seo, S. 2006. A review and comparison of methods for detecting outliers in univariate data sets, Thesis for Master of Science in Field of Public Health University of Pittsburgh, pp. 1-59.
- Fu, G., Charles, S.P., Yu, J. A. 2009. critical overview of pan evaporation trends over the last 50 years. *Climate Change*, 97, 193–214.
- Gopinathan, K.K. 1988. A general formula for computing the coefficients of the correlation connecting global solar radiation to sunshine duration. *Solar energy*, 41, 499-502.
- Goudarzi, M., Salahi, B., Hosseini, S. A. 2018. Estimation of evapotranspiration rate due to climate change in the Urmia Lake Basin. *Iranian Journal of Watershed Management Science and Engineering*, 12(41), 1-12. (In Farsi).
- Hagan, M.T., H.B. Demuth and M.H. Beale 1996. *Neural Network*. Design PWS Publishing Co.
- Hocking, R.R. 2013. *Methods and Applications of Linear Models, Regression and The Analysis of Variance*. John Wiley & Sons.
- Huo, Z., Feng, S., Kang, S., Dai, X. 2012. Artificial neural network models for reference evapotranspiration in an arid area of northwest China. *Journal of Arid Environment*, 82, 81–90.
- Karimipour, A., Banitalebi, G. 2020. Sensitivity analysis of meteorological data in estimating reference evapotranspiration with the minimum data using wavelet-neuro-fuzzy, ANN and ANFIS models. *Journal of Soil and Water Resources Conservation*, 9(3), 47-72. (In Farsi).
- Kisi, O., Alizamir, M. 2018. Modelling reference evapotranspiration using a new wavelet conjunction heuristic method, Wavelet extreme learning machine vs wavelet neural networks. *Agricultural and Forest Meteorology*, 263, 41-48.
- Kumar M., Raghuvanshi N.S., Singh R., Wallender W.W., Pruitt W.O. 2002. Estimating evapotranspiration using artificial neural network. *Journal of Irrigation and Drainage Engineering (ASCE)*, 128(4), 224–233
- Liang, L., Lijuan, L., Qiang, L. 2010. Temporal variation of reference evapotranspiration during 1961-2005 in the Taer river basin of Northeast China. *Agricultural and Forest Meteorology*, 150, 298-306.
- Lin C.H., Chao C., Chen W.F. 2008. Estimation regional evapotranspiration by adaptive network-based fuzzy inference system for Dan-Shui basin in Taiwan. *Journal of the Chinese Institute of Engineers*, 30(6), 1091-1096.
- Liu, C.M., Zhang, D., Liu, X.M., Zhao, C.S. 2012. Spatial and temporal change in the potential evapotranspiration sensitivity to meteorological factors in China (1960–2007). *Journal of Geographical Sciences*, 22(1), 3-14
- Liu, X., Zhang, D. 2013. Trend analysis of reference evapotranspiration in Northwest China, The roles of changing wind speed and surface air

- Vapnik, V.N. 2000. The nature of statistical learning theory, ser. Statistics for engineering and information science, vol 21. Springer, New York, pp 1003–1008.
- Wen, X., Si, J., He, Z., Wu, J., Shao, H., Yu, H. 2015. Support-Vector-Machine-Based Models for Modeling Daily Reference Evapotranspiration with Limited Climatic Data in Extreme Arid Regions. *Water Resources Management*, 29, 3195–3209.
- Willmott, C.J. 1982. Some comments on the evaluation of model performance. *Bulletin of the American Meteorological Society*, 63, 1309-1313.
- Willmott, C.J., Robeson, S.M., Matsuura, K. 2012. A refined index of model performance. *International Journal of climatology*, 32, 2088-2094.
- Yang, L., Feng, Q., Li, C., Si, J., Wen, X., Yin, Z. 2017. Detecting climate variability impacts on reference and actual evapotranspiration in the Taohe River Basin, NW China. *Hydrology Research*, 48, 596–612.
- Yassin, M.A., Alazba, A.A., Mattar, M.A. 2016. Artificial neural networks versus gene expression programming for estimating reference evapotranspiration in arid climate. *Agricultural Water Management*. 163, 110–124.
- Shiri, J., Nazemi, A.H., Sadraddini, A.A., Landaras, G., Kisi, O., Fard, A. F., Marti, P. 2014. Comparison of heuristic and empirical approaches for estimating reference evapotranspiration from limited inputs in Iran. *Computers and Electronics in Agriculture*, 108, 230- 241.
- Shiri, J. 2017. Evaluation of FAO56-PM, empirical, semi-empirical and gene expression programming approaches for estimating daily reference evapotranspiration in hyper-arid regions of Iran. *Agricultural Water Management*, 188, 101- 114.
- Singh A., Haghverdi, A., Öztürk, H.S., Durner, W. 2020. Developing Pseudo Continuous Pedotransfer Functions for International Soils Measured with the Evaporation Method and the HYPROP System, I. The Soil Water Retention Curve. *Water*, 12, 3425.
- Su, X.L., Song, Y., Niu, J.P., Ji, F. 2015. Sensitivity and attribution of potential evapotranspiration in Jinghuiqu irrigation district. *Natural Resources Research*, 1, 115–123.
- Ungaro, F., Calzolari, C., Busoni, E. 2005. Development of pedotransfer functions using a group method of data handling for the soil of the Pianura Padano-Veneta region of North Italy, water retention properties. *Geoderma*, 124, 293–317.



Comparison of the reference evapotranspiration estimations by data mining methods and Crop Water Requirement System project in Alborz province

A. Sedaghat^{1*}, N. A. Ebrahimi Pak², A. Tafteh³, S. N. Hosseini¹

Received: 10/04/2022

Accepted: 24/04/2023

Abstract

The reference evapotranspiration (ET₀) is an important factor for determining the plant water requirements and irrigation scheduling, which is usually estimated by widely accepted equation of Penman Monteith FAO-56 method. The aim of this study was to evaluate the performance of artificial neural networks (ANNs), random forest (RF) and support vector machine (SVM) methods for estimating the daily ET₀ in Alborz province. A ten-year-data (2010-2020) of five meteorological synoptic stations of namely MeshkinDasht, Hashtgerd, Eshtehard, Taleghan and Karaj were used for estimation of ET₀. The obtained values were compared with the provided data of a national project entitled Crop Water Requirement System. According to the results, the best agreement was found in Meshkin Dasht and Karaj stations. Besides, among the applied approaches, the ANNs method had the highest accuracy comparing to other methods. The values of EF and NRMSE in the ANNs method were determined 0.96 and 0.11, respectively for both training and testing steps in Meshkin Dasht station. While, these values for the RF method were determined 0.96 and 0.11, for training stage and 0.95 and 0.12 for testing stage, respectively. The obtained results in the Karaj station showed that EF and NRMSE in the ANNs method were 0.96 and 0.11, respectively for the training and 0.95 and 0.12 for the testing stage. These values for the RF method were 0.96 and 0.12 for the training stage respectively and 0.95 and 0.13 for the testing. Considering the higher accuracy of the ANNs and RF methods, these approaches can be recommended for estimating the daily ET₀ across Alborz province.

Keywords: Data Mining, Penman Monteith FAO-56 method, Reference Evapotranspiration, Water requirement system



¹ Researcher, Soil and Water Research Institute, Agricultural Research, Education and Extension Organization, Karaj, Iran

(*Corresponding author Email Address: Azadehsedaghat65@gmail.com)

² Associate Professor, Soil and Water Research Institute, Agricultural Research, Education and Extension Organization, Karaj, Iran

³ Assistant Professor, Soil and Water Research Institute, Agricultural Research, Education and Extension Organization, Karaj, Iran

نحوه ارجاع مقاله:

سداقت، آ.، ابراهیمی پاک، ن. ع.، تافته، آ.، حسینی، س. ن. ۱۴۰۲. مقایسه برآوردهای تبخیرتعرق مرجع روزانه با روش‌های داده‌کاوی و سامانه نیاز آبی

گیاهان در استان البرز. نشریه هواشناسی کشاورزی، ۱۱(۲): ۱۶-۲۸. DOI: [10.22125/agmj.2023.336962.1134](https://doi.org/10.22125/agmj.2023.336962.1134)

Sedaghat, A., Ebrahimi Pak, N. A., Tafteh, Hosseini, S. N. 2023. Comparison of the reference evapotranspiration estimations by data mining methods and Crop Water Requirement System project in Alborz province. Journal of Agricultural Meteorology, 11(2): 16-28. DOI: [10.22125/agmj.2023.336962.1134](https://doi.org/10.22125/agmj.2023.336962.1134)