

برآورد تبخیر و تعرق مرجع روزانه در مناطق مرطوب با استفاده از روش‌های داده‌محور رگرسیون فرآیند گاوسی،
رگرسیون بردار پشتیبان و جنگل تصادفی
سعید صمدیان‌فرد، مبارک سالاری‌فر، سحر جاویدان و فاطمه میکائیلی

دوره ۶، شماره ۴، زمستان ۱۳۹۹، صفحات ۳۷۳ - ۳۶۰

Vol. 6(4), Winter 2020, 360 – 373

DOI: 10.22034/jewe.2020.241690.1394

Estimation of Daily Reference Evapotranspiration
in Humid Climates Using Data-Driven Methods of
Gaussian Process Regression, Support Vector
Regression and Random Forest

Samadianfard S., Salarifar M., Javidan S. and
Mikaeili F.



www.jewe.ir

OPEN ACCESS

ارجاع به این مقاله:

صمدیان‌فرد س.، سالاری‌فر م.، جاویدان س. و میکائیلی ف. (۱۳۹۹). برآورد تبخیر و تعرق مرجع روزانه با استفاده از روش‌های داده‌محور رگرسیون فرآیند گاوسی، رگرسیون بردار پشتیبان و جنگل تصادفی. محیط‌زیست و مهندسی آب، دوره ۶، شماره ۴، صفحات: ۳۶۰-۳۷۳.

Citing this paper: Samadianfard S., Salarifar M., Javidan S. and Mikaeili F. (2020). Estimation of daily reference evapotranspiration in humid climates using data-driven methods of Gaussian process regression, support vector regression and random forest. Environ. Water Eng., 6(4), 360–373. DOI: 10.22034/jewe.2020.241690.1394.

برآورد تبخیر و تعرق مرجع روزانه در مناطق مرطوب با استفاده از روش‌های داده‌محور رگرسیون فرآیند گاوسی، رگرسیون بردار پشتیبان و جنگل تصادفی

سعید صمدیان فرد^{۱*}، مبارک سالاری فر^۲، سحر جاویدان^۲ و فاطمه میکائیلی^۲

استادیار، گروه مهندسی آب، دانشکده کشاورزی، دانشگاه تبریز، تبریز، ایران
دانشجوی کارشناسی ارشد، گروه مهندسی آب، دانشکده کشاورزی، دانشگاه تبریز، تبریز، ایران

*نویسنده مسئول: s.samadian@tabrizu.ac.ir

مقاله اصلی

تاریخ دریافت: [۱۳۹۹/۰۵/۰۸]

تاریخ بازنگری: [۱۳۹۹/۰۸/۰۳]

تاریخ پذیرش: [۱۳۹۹/۰۸/۰۵]

چکیده

تخمین دقیق تبخیر و تعرق گیاه مرجع در برنامه‌ریزی‌های آبیاری اهمیت ویژه‌ای دارد. همچنین، عدم دسترسی به داده‌های لایسیمیتری باعث شده است پژوهش‌گران به استفاده از روش‌های غیرمستقیم از جمله روش‌های داده‌محور روی آورند. در پژوهش حاضر، توانایی روش‌های داده‌محور رگرسیون فرآیند گاوسی (GPR)، رگرسیون بردار پشتیبان (SVR) و جنگل تصادفی (RF) در تخمین تبخیر و تعرق گیاه مرجع مورد بررسی قرار گرفت. بدین منظور، داده‌های هواشناسی دمای میانگین، سرعت باد، رطوبت نسبی و ساعات آفتابی در بازه زمانی ۹۷-۱۳۹۲ در ایستگاه شمالی کشور از جمله آستارا، بندر انزلی، رشت، رامسر، نوشهر، ساری، بندر ترکمن، گرگان، گنبدکاووس جمع‌آوری شد. تبخیر و تعرق محاسبه‌شده با استفاده از روش فائو-پنمن-مونتیث به‌عنوان خروجی‌های هدف در نظر گرفته‌شده و چهار سناریو ترکیبی از پارامترهای هواشناسی برای واسنجی و صحت‌سنجی روش‌های مورد مطالعه، مدنظر قرار گرفتند. دقت روش‌های مذکور با استفاده از پارامترهای آماری ضریب همبستگی، شاخص پراکندگی و ضریب وایلموت مورد مقایسه قرار گرفت. نتایج نشان داد که مدل GPR4 با شاخص پراکندگی در محدوده ۰/۱۳۲ تا ۰/۱۷۹ در ایستگاه‌های آستارا، بندر انزلی، رشت، رامسر، نوشهر و ساری، مدل SVR4 با شاخص پراکندگی ۰/۱۱۶ تا ۰/۱۲۰ در ایستگاه‌های بندر ترکمن و گنبدکاووس و روش هارگریوز-سامانی با شاخص پراکندگی ۰/۵۰۹ در ایستگاه گرگان برآوردهای به‌مراتب دقیق‌تری از تبخیر و تعرق گیاه مرجع داشته‌اند.

واژه‌های کلیدی: تبخیر و تعرق؛ روش‌های داده‌محور؛ شمال کشور؛ فائو-پنمن-مونتیث؛ هارگریوز-سامانی.

۱- مقدمه

برای برآورد میزان آب موردنیاز در طرح‌های آبیاری، تعیین آب مصرفی گیاهان باغی و زراعی، برنامه‌ریزی آبیاری، تقسیم آب و کنترل سیلاب، تبخیر و تعرق از مؤلفه‌های مهم و حیاتی به شمار می‌رود (Samadianfard and Panahi 2018). تبخیر و تعرق به دو روش مستقیم و غیرمستقیم اندازه‌گیری می‌شود. استفاده از روش اندازه‌گیری مستقیم (لایسیمتر)^۱ به دلیل هزینه‌بر بودن و زمان‌بر بودن در همه مناطق امکان‌پذیر نمی‌باشد (Khoshhal et al. 2015). با توجه به نتایج به‌دست‌آمده از تحقیقات انجام‌گرفته در نقاط مختلف جهان، رابطه فائو-پنمن-مونتیث^۲ به‌عنوان رابطه‌ای استاندارد موردقبول اکثر محققین قرارگرفته است (Azimi et al. 2016). (Zareabyaneh et al. 2011) به مقایسه‌ی روش فائو-پنمن-مونتیث و تشت تبخیر کلاس A با داده‌های لایسیمتری در برآورد تبخیر و تعرق گیاه برنج در منطقه آمل پرداختند و نتایج مطالعه آن‌ها نشان داد که استفاده از روش فائو-پنمن-مونتیث در محاسبه تبخیر و تعرق گیاه مرجع و اعمال ضریب گیاهی تبخیر و تعرق برنج در این منطقه، چهار الی پنج درصد برآورد بیشتری نسبت به مقادیر اندازه‌گیری لایسیمتری دارد. با این حال، رابطه فائو-پنمن-مونتیث نیازمند تعداد بیشتری از داده‌های هواشناسی همچون تابش، دما، رطوبت و سرعت باد می‌باشد که در آن گیاه مرجع، گیاه فرضی (چمن یا یونجه) با ارتفاع ۱۲ cm، ضریب بازتاب ۰.۲۳ و مقاومت روزانه ۷۰ s/m می‌باشد (Samadianfard and Panahi 2018). (Trajkovic 2007) به‌علت این‌که روش فائو-پنمن-مونتیث نیازمند برخی پارامترهای هواشناسی است که در تمامی مناطق و ایستگاه‌ها در دسترس نمی‌باشد، استفاده از روش تجربی هارگریوز را به‌جای روش مذکور در مناطق مرطوب امکان‌سنجی نمود و به این نتیجه رسید که استفاده از روش هارگریوز با ضرایب تعدیل‌شده مقادیر تبخیر و تعرق

پتانسیل را فقط ۱٪ بیش‌تر از روش فائو-پنمن-مونتیث برآورد می‌کند و این روش به خاطر اینکه فقط از داده دما استفاده می‌کند، قابل‌استفاده بوده و توصیه می‌گردد.

در زمینه برآورد تبخیر و تعرق، پژوهش‌های بسیاری در سطح ملی و بین‌المللی انجام پذیرفته است. Hashemi et al. (2019) روش‌های پایه تابشی در برآورد تبخیر و تعرق گیاه مرجع را مقایسه کرده و به این نتیجه رسیدند که در شهرهای بناب و کلیبر روش پرستلی تیلور^۳ به‌ترتیب با داشتن خطای ۰/۵ و ۰/۳۸ نتایج بهتری ارائه داد و در شهر مرند روش ایرماک^۴ با دارا بودن خطای ۰/۵ عملکرد مناسب‌تری داشته است. (Izadyar et al. 2019) به برآورد تبخیر و تعرق گیاه مرجع با استفاده از روش‌های دما پایه و انتقال جرم در برخی ایستگاه‌های هواشناسی استان آذربایجان شرقی پرداخته و تبخیر و تعرق گیاه مرجع را با استفاده از روش‌های تجربی هارگریوز-سامانی، WMO^۵ و فائو-پنمن-مونتیث برآورد کرده و با مقایسه پارامترهای خطا از قبیل جذر میانگین مربعات خطا و ضریب همبستگی روش هارگریوز-سامانی^۶ را به‌عنوان روش برتر در ایستگاه‌های مورد مطالعه پیشنهاد کردند. Behmanesh et al. (2014) مدل‌های سری زمانی خطی و دو-خطی^۷ در پیش‌بینی تبخیر و تعرق گیاه مرجع در ایستگاه سینوپتیک ارومیه را ارزیابی و با مدل فائو-پنمن-مونتیث مقایسه کردند. آن‌ها به این نتیجه رسیدند که مدل دو-خطی توانایی بیشتری در پیش‌بینی مقادیر ماهانه تبخیر و تعرق گیاه مرجع در مقایسه با مدل خطی داشت. (Dameneh and Sanaeinejad 2018) تبخیر و تعرق پتانسیل در منطقه فریمان را با استفاده از روش پرستلی تیلور و فن سنجش از دور ارزیابی نموده و به این نتیجه رسیدند که در اقلیم نیمه‌خشک، روش پرستلی تیلور از دقت بالایی برخوردار است.

³ Priestley taylor equation

⁴ Irmak equation

⁵ World Meteorological Organization

⁶ Hargreaves samani equation

⁷ Bilinear models

¹ Lysimeter

² Fao Penman Monteith equation

رسیدند که هر دو مدل مذکور مبتنی بر دما می‌توانند به‌طور دقیق تبخیر و تعرق روزانه را تخمین بزنند. (2019) Carter and liang ۱۰ روش داده‌محور و فن سنجش از دور را در برآورد تبخیر و تعرق گیاه مرجع ارزیابی کرده و دقت و زمان اجرای هر یک از ۱۰ روش را ارزیابی نمودند. آن‌ها دریافتند که مناسب‌ترین نتایج با استفاده از مدل درخت رگرسیون تجمعی بوت استرپ^۸ و مدل شبکه عصبی با سه لایه پنهان به دست می‌آید. (2019) Granata سه ترکیب مختلف برآورد تبخیر و تعرق را با استفاده از الگوریتم‌های رگرسیون بردار پشتیبان^۹، جنگل تصادفی، بگینگ^{۱۰} و مدل M5 ارزیابی نمود. مدل M5 در ترکیب ۱ که متغیرهای ورودی آن تابش خورشیدی خالص، شار گرما، میزان رطوبت خاک، سرعت باد، میانگین رطوبت نسبی و میانگین دما بود، بهترین نتیجه را ارائه داد. (Huang et al. 2020) مدل‌های جنگل تصادفی، رگرسیون بردار پشتیبان و مدل اصلاح‌شده شاتل واورت-والاس را در برآورد تبخیر و تعرق گیاه مرجع مورد بررسی قرار داده و نتیجه گرفتند که محاسبات مقاومت سطح خاک با استفاده از روش تکرارشونده مونت کارلو، محاسبات مقاومت روزانه گیاهی، غلظت و انتشار کربن دی‌اکسید، عملکرد مدل اصلی شاتل واورت-والاس را بهبود می‌بخشد.

در پژوهش حاضر سعی شد که برای نه ایستگاه با توزیع پراکندگی مناسب در مناطق مرطوب ایران و با در نظر گرفتن ۱۲ ترکیب ورودی مختلف از داده‌های هواشناسی، مقادیر تبخیر و تعرق گیاه مرجع روزانه به‌وسیله روش‌های داده‌محور رگرسیون فرآیند گاوسی^{۱۱} (GPR)، رگرسیون بردار پشتیبان (SVR) و جنگل تصادفی (RF) مدل‌سازی شود. این تحقیق بر پایه ارزیابی این روش‌ها در تخمین مقادیر تبخیر و تعرق گیاه مرجع روزانه در مناطق مختلف

(Shabani et al. 2017) تلفیقی از روش شبکه عصبی و روش‌های محاسباتی را جهت تخمین دقیق‌تر تبخیر و تعرق مرجع بکار برده و به این نتیجه رسیدند که تنها روش جنسن-هیز^۱ منجر به تخمین تبخیر و تعرق مرجع با دقت بالا شده است. (2019) Kaviyani et al. تبخیر و تعرق واقعی گیاهان زراعی را به کمک الگوریتم‌های بیلان انرژی در دشت قزوین برآورد نموده و الگوریتم SSEB^۲ را با کمترین میزان جذر میانگین خطا و بیشترین همبستگی با داده‌های لایسیمتری، به‌عنوان الگوریتم برتر معرفی کردند. (2019) Siasar and Honar مدل‌های ماشین بردار پشتیبان^۳ و جنگل تصادفی^۴ را برای برآورد تبخیر و تعرق گیاه مرجع روزانه در شمال استان سیستان و بلوچستان بکار برده و مدل جنگل تصادفی را با بیشترین دقت و کمترین میانگین قدر مطلق خطا برای برآورد تبخیر و تعرق در این منطقه توصیه نمودند. (Sima and Khorsand Movaghar 2019) مدل‌های بیلان انرژی مبتنی بر داده‌های سنجش‌ازدور در برآورد تبخیر از سطح دریاچه‌های شور را مقایسه نموده و به این نتیجه رسیدند که تبخیر به‌دست‌آمده از مدل SEBAL^۵ نوع اول کمترین مقدار و تبخیر به‌دست‌آمده از مدل SEBS^۶ بیشترین مقدار را دارد. (2015) Saremi and Farhadi Bansouleh به تعیین پارامترهای مؤثر در برآورد تبخیر و تعرق گیاه مرجع با استفاده از شبکه عصبی مصنوعی پرداختند. ایشان نتیجه گرفتند که افزایش تعداد اطلاعات در لایه ورودی لزوماً منجر به بهبود نتایج مدل‌های هوشمند نمی‌شود. (2017) Cui et al. مدل‌های جنگل تصادفی و شبکه عصبی رگرسیون تعمیم‌یافته (GRNN)^۷ را با استفاده از دو ترکیب ورودی شامل داده‌های کامل و فقط دما، در برآورد تبخیر و تعرق گیاه مرجع روزانه ارزیابی نموده و به این نتیجه

⁷ Generalize Regression Neural Networks

⁸ Bootstrap

⁹ Support Vector Regression

¹⁰ Bagging method

¹¹ Gaussian process regression

¹ Jensen haise equation

² Simplified Surface Energy Balance

³ Support vector machine

⁴ Random forest

⁵ Surface Energy Balance Algorithm for Land

⁶ Surface Energy Balance System

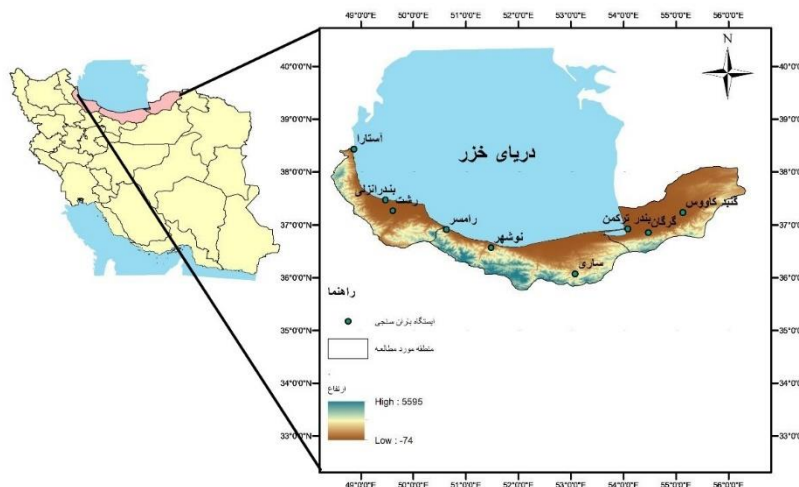
شمالی کشور پایه‌ریزی شد. در نهایت، با بررسی و آنالیز آماری، دقیق‌ترین روش پیشنهاد شد.

انتخاب دوره آماری ۶ ساله، بررسی روند تغییرات احتمالی مقادیر تبخیر و تعرق در سال‌های اخیر و مدل‌سازی دقیق‌تر و هر چه بهتر روند مذکور بود. مقادیر محاسبه‌شده تبخیر و تعرق گیاه مرجع با روش فائو-پنمن-مونتیث (ET_0) که از جمله روش‌های دقیق و تأییدشده توسط محققین مختلف در تخمین تبخیر و تعرق در اقلیم‌های مختلف می‌باشند، به‌عنوان مقادیر استاندارد و هدف برای واسنجی و صحت‌سنجی سناریوهای متفاوت روش‌های مورد مطالعه رگرسیون فرآیند گاوسی (GPR)، رگرسیون بردار پشتیبان (SVR) و جنگل تصادفی (RF) مورد استفاده قرار گرفت. موقعیت مکانی ایستگاه‌های مورد مطالعه در شکل (۱) و مشخصات جغرافیایی آن‌ها در جدول (۱) آورده شده است.

۲- مواد و روش‌ها

۱-۲- منطقه مورد مطالعه

برای بررسی و برآورد مقادیر تبخیر و تعرق، نه ایستگاه شمالی کشور در استان‌های گیلان، مازندران و گلستان از جمله آستارا، بندر انزلی، رشت، رامسر، نوشهر، ساری، بندر ترکمن، گرگان، گنبدکاووس انتخاب شدند. بدین منظور، در یک دوره آماری شش‌ساله (۱۳۹۲-۱۳۹۷) جمع‌آوری داده‌های هواشناسی دمای میانگین (T)، سرعت باد (U)، رطوبت نسبی (RH) و ساعات آفتابی (N) از سازمان هواشناسی کشور صورت پذیرفت. همچنین با توجه به تغییرات الگوی تبخیر و تعرق در دهه‌های اخیر، هدف از



شکل ۱- پراکنش ایستگاه‌های مورد مطالعه در منطقه شمالی ایران

Fig. 1 Distribution of studied stations in the northern region of Iran

جدول ۱- مشخصات جغرافیایی ایستگاه‌های مورد مطالعه

Table 1 Geographical profile table of the studied station

Station	Longitude	Latitude	Elevation from sea level (m)
Astara	48° 52'	38° 26'	-22
Bandar Anzali	49° 28'	37° 28'	-16
Rasht	49° 36'	37° 15'	0
Ramsar	50° 38'	36° 55'	72
Nushahr	51° 29'	36° 39'	-12
Sari	53° 05'	36° 04'	54
Bandar Torkaman	54° 05'	36° 55'	-25
Gorgan	54° 28'	36° 51'	105
Gonbad Kavus	55° 08'	37° 14'	52

روش‌های مورد مطالعه رگرسیون فرآیند گاوسی (GPR)، رگرسیون بردار پشتیبان (SVR) و جنگل تصادفی (RF) اختصاص یافت. چهار ترکیب مختلف از داده‌های هواشناسی به عنوان ورودی مدل‌ها و مقادیر محاسبه شده متناظر، به عنوان خروجی در نظر گرفته شد (جدول ۲).

لازم به ذکر است که تمامی محاسبات واسنجی و صحت‌سنجی با استفاده از نرم‌افزار Weka که توسط Witten در سال ۱۹۹۳ در دانشگاه وایکاتو نیوزلند توسعه داده شده، انجام پذیرفت.

۲-۳- روش فائو-پنمن-مونتیث

Allen et al. (1998) رابطه (۱) را که به نام رابطه فائو-پنمن-مونتیث معروف است به عنوان دقیق‌ترین روش برآورد نیاز آبی گیاه ارائه نمودند و مقادیر تبخیر و تعرق گیاه مرجع با روش فائو-پنمن-مونتیث و با استفاده از رابطه (۱) محاسبه گردید.

$$ET_0 = \frac{0.408\Delta(R_n - G) + \gamma\left(\frac{900}{T + 273}\right)u_2(e_s - e_a)}{\Delta + \gamma(1 + 0.34u_2)} \quad (1)$$

$$m(x) = E(f(x)) \quad (2)$$

$$k(x, x') = E(f(x) - m(x))(f(x') - m(x')) \quad (3)$$

که، $k(x, x')$ تابع کواریانس بوده و در نقاط x و x' محاسبه می‌شود. فرآیند گاوسی $f(x)$ به صورت رابطه (۴) بیان می‌شود.

$$f(x) \sim GP(m(x), k(x, x')) \quad (4)$$

که معمولاً جهت ساده‌سازی، مقدار تابع میانگین برابر با صفر در نظر گرفته می‌شود (Saghebian and Roushangar 2019).

۲-۵- روش رگرسیون بردار پشتیبان

ماشین بردار پشتیبان یکی از روش‌های یادگیری است که توسط Boser et al. (1992) بر پایه تئوری یادگیری آماری معرفی شد. آن‌ها طی سال‌های بعد، تئوری ابر

جدول ۲- پارامترهای دخیل در هر سناریو در روش‌های SVR، GPR و RF

Table 2 Parameters involved in each scenario in the methods of GPR, RF, and SVR

Scenario Number	Input Parameters
1	T
2	T, RH
3	T, RH, U
4	T, RH, U, N

۲-۲- پارامترهای ورودی مدل

در پژوهش حاضر، امکان‌سنجی استفاده از پارامترهای هواشناسی دما، رطوبت نسبی، سرعت باد، ساعات آفتابی در ترکیب‌های مختلف به منظور تخمین تبخیر و تعرق گیاه مرجع مدنظر قرار گرفت. بنابراین، داده‌های خام هواشناسی روزانه بین سال‌های ۱۳۹۲-۹۷ جمع‌آوری و مقادیر تبخیر و تعرق گیاه مرجع با روش فائو-پنمن-مونتیث و با استفاده از نرم‌افزار اکسل، محاسبه شد. از بین کل داده‌های موجود، حدود ۷۰٪ جهت واسنجی و ۳۰٪ بقیه جهت صحت‌سنجی

که در آن ET_0 تبخیر و تعرق گیاه مرجع (mm/day)، R_n تابش خالص در سطح پوشش گیاهی ($MJ/m.day$)، G شار گرما به داخل خاک ($MJ/m.day$)، T متوسط دمای هوا در ارتفاع دو متری از سطح زمین ($^{\circ}C$)، u_2 سرعت باد در ارتفاع دو متری از سطح زمین (m/s)، e_a فشار بخار واقعی (KPa)، e_s فشار بخار اشباع هوا (KPa)، $e_s - e_a$ کمبود فشار بخار اشباع (KPa)، Δ شیب منحنی فشار بخار ($KPa/^{\circ}C$) و γ ضریب رطوبتی ($KP/^{\circ}C$) می‌باشد.

۲-۴- روش رگرسیون فرآیند گاوسی

فرآیند گاوسی مجموعه‌ای از متغیرهای تصادفی است که تعداد محدودی از آن‌ها با توزیع‌های گاوسی ادغام شده‌اند. توزیع گاوسی در واقع توزیع بین متغیرهای تصادفی بوده، درحالی‌که فرآیند گاوسی بیانگر توزیع بین توابع می‌باشد. فرآیند گاوسی $f(x)$ توسط توابع میانگین $m(x)$ و کواریانس به صورت روابط (۲) و (۳) تعریف می‌شود.

خطایی استفاده کردند که خطاهایی که در یک ε -
Insensitive به نام فاصله معین از مقادیر واقعی می‌باشند
را نادیده می‌گیرد. این تابع به صورت رابطه (۵) تعریف
می‌شود.

$$L(y, f(x, a)) = |y - f(x, a)|_{\varepsilon} = \begin{cases} 0 & \text{if } |y - f(x, a)| \leq \varepsilon \\ |y - f(x, a)| - \varepsilon & \text{if } |y - f(x, a)| > \varepsilon \end{cases} \quad (5)$$

مدل جنگل تصادفی برای بسیاری از مجموعه داده‌ها،
دسته‌بندی را با سرعت بالایی انجام می‌دهد و برخلاف مدل‌های
کلاسیک مانند رگرسیون که تنها بر پایه یک مدل تکیه‌دارند، با
استفاده از صدها و هزاران درخت از اطلاعات بیشتری در داده‌ها
استفاده می‌کند تا بتوان استنباط بهتری از متغیرها داشت. این
الگوریتم حاوی چندین درخت تصمیم است که خروجی آن از
خروجی‌های درختان انفرادی به دست می‌آید. در ساخت هر
درخت به طور تصادفی تعدادی از متغیرهای ورودی شرکت
می‌کنند و هنگامی که تعداد مشاهدات در مقایسه با تعداد
پیش‌بینی‌کننده‌ها نسبتاً کم باشد یک روش پیش‌بینی کارآمد
به حساب می‌آید (Breiman 2001). برای یک مجموعه
طبقه‌بندی شده خاص $h_1(x), h_2(x), \dots, h_k(x)$ ، توابع حاشیه
را می‌توان به صورت رابطه (۷) نوشت.

$$m(x, y) = \text{av}_k I(h_k(x)=y) - \max_{j \neq y} \text{av}_k I(h_k(x)=j) \quad (7)$$

دمای حداقل و حداکثر و تابش فرازمینی نیاز می‌باشد
(Wang et al. 2008). روش هارگریوز-سامانی با وجود
توصیه در بعضی مناطق در همه اقلیم‌ها کارایی ندارد. این
موضوع با توجه به اینکه روش مذکور دمایی بوده و اثر
عواملی همچون رطوبت نسبی، باد و مدت و شدت تابش
آفتاب در آن دیده نشده سازگاری دارد (Sharifian et al.
2012). معادله هارگریوز-سامانی برای برآورد تبخیر و تعرق
گیاه مرجع به شکل رابطه (۸) ارائه شده است.

$$ET_0 = 0.0023 \times Ra \times (T + 17.8) \times \sqrt{RT} \quad (8)$$

که، ET_0 تبخیر و تعرق گیاه مرجع (mm/day)، Ra تابش
برون زمینی (mm/day)، T متوسط دمای هوا ($^{\circ}C$)، TR
اختلاف دمای حداکثر و حداقل ($^{\circ}C$) می‌باشد.

صفحه‌ی بهینه را به عنوان طبقه‌بندی‌کننده خطی ارائه داده
و طبقه‌بندی‌کننده‌های غیرخطی را با کمک توابع کرنل
معرفی نمودند. (Basak et al. 2007) برای کاربرد
ماشین‌های بردار پشتیبان در مسائل رگرسیون از تابع

این تابع خطا (رابطه ۵) مقادیر خطای کم‌تر از ε را در نظر
نمی‌گیرد و لازم است وابستگی تابع 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، مطابق رابطه (۶)، مشخص می‌شود.

$$y = f(x) + \text{noise} \quad (6)$$

بنابراین موضوع اصلی، پیدا کردن فرم تابع f است که بتواند
به صورت صحیح، موارد جدیدی را که SVM تاکنون تجربه
نکرده است، پیش‌بینی کند (Siasar and Honar 2019).

۲-۶- روش جنگل تصادفی

که، I تابع شاخص است و پیش‌بینی نهایی با میانگین نتایج
حاصل از همه درختان مشخص می‌شود (Huang et al. 2020).

۲-۷- روش هارگریوز-سامانی

بسیاری از روش‌های برآورد تبخیر و تعرق گیاه مرجع به
داده‌های متعدد هواشناسی نیازمند می‌باشند، اما برخی از
این داده‌ها در دسترس نیستند و به فرض دسترسی، از دقت
کافی برخوردار نمی‌باشند؛ بنابراین روش‌های مبتنی بر
حداقل داده‌های اقلیمی گسترش یافتند که روش هارگریوز-
سامانی از این جمله می‌باشد (Najafi 2006). روش
هارگریوز-سامانی یکی از ساده‌ترین رابطه‌ها برای برآورد
تبخیر و تعرق گیاه مرجع است، چرا که از دقت نسبتاً خوبی
در مناطق خشک و نیمه‌خشک برخوردار می‌باشد و در آن به

$$R = \frac{\text{cov}(O_i, P_i)_i}{\sqrt{\text{var}(O_i) \times \text{var}(P_i)}} \quad (9)$$

$$WI = \left| 1 - \frac{\sum_{i=1}^n (O_i - P_i)^2}{\sum_{i=1}^n (|P_i - \bar{O}_i| + |O_i - \bar{O}_i|)^2} \right|, \quad 0 \leq WI \leq 1 \quad (10)$$

ضریب همبستگی ۰/۹۸۱ و ۰/۹۸۲، شاخص پراکندگی ۰/۱۱۶ و ۰/۱۲۰، ضریب ویلموت ۰/۹۹ و ۰/۹۹ و در ایستگاه گرگان روش هارگریوز-سامانی با ضریب همبستگی ۰/۷۶۲، شاخص پراکندگی ۰/۴۴۷ و ضریب ویلموت ۰/۸۵۷، دقت بیشتری نسبت به سایر روش‌ها دارند. در حالت کلی در صورت وجود پارامترهای دما، رطوبت نسبی، سرعت باد و ساعات آفتابی مدل GPR4 با توجه به اینکه در اغلب ایستگاه‌ها عملکرد بهتری داشته است، به‌عنوان روش برتر معرفی می‌شود.

از طرف دیگر، ترکیب شماره ۱ با توجه به اینکه تنها از پارامتر دما استفاده می‌نماید، در مدل‌های مذکور به جز در ایستگاه‌های رامسر، بندر ترکمن و نوشهر، نسبت به روش هارگریوز-سامانی عملکرد ضعیفی ارائه کرده است. در نتیجه در ایستگاه‌های مورد مطالعه، در صورت کمبود پارامترهای ورودی بهتر است از مدل هارگریوز-سامانی استفاده شود. نتایج حاصل از پژوهش حاضر تطابق کاملی با نتایج مطالعه Samadianfard and Panahi (2018) که در ایستگاه‌های مناطق خشک و نیمه‌خشک به انجام رسید، داشته و ضرورت استفاده از پارامترهای هواشناسی دما، رطوبت نسبی، سرعت باد و ساعات آفتابی در برآورد دقیق تبخیر و تعرق مرجع روزانه را به اثبات رسانید.

در نهایت، ترکیب‌های ۳ و ۴ برای هر سه روش GPR، SVR و RF به‌عنوان سناریوی برتر در پیش‌بینی تبخیر و تعرق گیاه مرجع انتخاب شده و نمودارهای تغییرات مقادیر تبخیر و تعرق مرجع پیش‌بینی شده برای دو سناریو برتر

۸-۲- ارزیابی دقت

برای ارزیابی دقت روش‌های مورد مطالعه در برآورد تبخیر و تعرق گیاه مرجع، از پارامترهای ضریب همبستگی (R)، ضریب ویلموت (WI) و شاخص پراکندگی (SI)، مطابق روابط (۹) الی (۱۱) استفاده شد.

$$SI = \frac{\sqrt{\frac{1}{n} \sum_{i=1}^n (P_i - O_i)^2}}{\bar{O}} \quad (11)$$

که، P_i تبخیر و تعرق گیاه مرجع تخمین زده شده، O_i تبخیر و تعرق گیاه مرجع محاسبه شده با روش فائو-پنمن-مونتیث و n تعداد داده‌ها می‌باشند.

۳- یافته‌ها و بحث

همان‌طوری که قبلاً نیز اشاره گردید، امکان‌سنجی استفاده از پارامترهای هواشناسی دما، رطوبت نسبی، سرعت باد، ساعات آفتابی در ترکیب‌های مختلف به‌منظور تخمین تبخیر و تعرق گیاه مرجع مورد بررسی قرار گرفت. بدین منظور، داده‌های خام هواشناسی روزانه بین سال‌های ۹۷-۱۳۹۲ جمع‌آوری و مقادیر تبخیر و تعرق گیاه مرجع با روش فائو-پنمن-مونتیث و با استفاده از نرم‌افزار اکسل، محاسبه شد. مشخصات آماری پارامترهای مذکور در ایستگاه‌های مورد مطالعه در جدول (۳) ارائه شده است.

پس از انجام محاسبات، دقت روش‌های داده‌محور مذکور و رابطه تجربی هارگریوز بر اساس معیارهای آماری (روابط (۹) تا (۱۱)) محاسبه شده و در جدول (۴) ارائه شده است.

با توجه به جدول (۴)، در ایستگاه‌های آستارا، بندر انزلی، رشت، رامسر و نوشهر مدل GPR4 با ضریب همبستگی بین ۰/۹۶۱ تا ۰/۹۷۹، شاخص پراکندگی بین ۰/۱۳۲ تا ۰/۱۷۹ و ضریب ویلموت بین ۰/۹۸۰ تا ۰/۹۸۹، در ایستگاه‌های بندر ترکمن و گنبدکاووس مدل SVR4 با

روش‌های داده‌محور و روش تجربی هارگریوز در شکل (۲) نشان داده شده‌اند.

جدول ۳- مشخصات آماری پارامترهای هواشناسی مورد استفاده

Table 3 Statistical characteristics of implemented meteorological parameters

Station	Variable	Mean	Minimum	Maximum	Standard Deviation	Coefficient Of Variation	Skewness
Astara	T	16.1	-2.7	30.2	7.9	0.5	-0.06
	RH	80.12	28.4	98.6	10.5	0.13	-1.003
	U	1.06	0.00	6.83	0.69	0.65	1.65
	N	5.5	0.00	13.5	4.43	0.81	0.13
Bandar anzali	ET ₀	2.43	0.35	7.87	1.73	0.71	0.64
	T	17.1	-1.3	29.2	7.3	0.4	-0.07
	RH	82.36	31.2	99.6	8.98	0.1	-1.03
	U	2.47	0.00	10.94	1.38	0.56	1.7
Rasht	N	5.51	0.00	14.00	4.51	0.82	0.09
	ET ₀	2.6	0.41	7.33	1.78	0.68	0.6
	T	17.1	-1.9	30.5	7.67	0.4	-0.12
	RH	82.7	25.2	100.0	10.2	0.12	-1.3
Ramsar	U	1.15	0.00	6.73	0.71	0.61	1.8
	N	4.91	0.00	13.4	4.27	0.86	0.18
	ET ₀	2.4	-2.73	6.93	1.72	0.72	0.53
	T	17.40	-1.50	30.50	7.401	0.425	-0.013
Noshahr	RH	80.19	48.60	98.90	8.172	0.102	-0.408
	U	1.37	0.00	6.26	0.761	0.555	1.572
	N	13.50	-2.40	24.90	6.658	0.493	-0.135
	ET ₀	3.83	0.43	9.25	2.650	0.692	0.338
Sari	T	17.24	1.60	29.90	7.248	0.420	-0.042
	RH	80.88	35.80	99.60	8.388	0.104	-0.566
	U	1.46	0.00	5.61	0.701	0.481	1.279
	N	5.52	0.00	13.40	4.217	0.764	-0.007
Bandartorkaman	ET ₀	2.48	0.31	6.89	1.581	0.636	0.615
	T	18.59	-1.70	115.20	8.123	0.437	0.660
	RH	79.36	35.60	99.40	9.201	0.116	0.459
	U	0.90	0.00	4.02	0.554	0.619	1.285
Gorgan	N	5.85	0.00	13.00	4.083	0.698	-0.164
	ET ₀	2.63	0.44	29.79	1.794	0.683	2.002
	T	18.47	-4.30	34.50	8.104	0.439	-0.139
	RH	73.38	37.50	98.00	9.678	0.132	-0.142
Gonbadkavus	U	2.48	0.00	13.46	1.429	0.577	1.431
	N	6.66	0.00	13.30	4.093	0.615	-0.385
	ET ₀	3.37	0.44	12.61	2.153	0.639	0.437
	T	18.29	-4.50	33.60	8.565	0.468	-0.090
Gorgan	RH	72.25	38.00	100.00	10.257	0.142	0.219
	U	3.43	0.00	25.93	4.721	1.378	3.049
	N	6.64	0.00	13.30	4.162	0.627	-0.342
	ET ₀	3.74	0.33	20.93	2.575	0.687	1.543
Gonbadkavus	T	19.24	-6.85	36.80	8.857	0.460	-0.099
	RH	65.45	21.50	97.00	14.029	0.214	0.052
	U	1.23	0.00	6.88	0.719	0.583	1.077
	N	7.03	0.00	13.60	4.150	0.590	-0.489
	ET ₀	3.37	0.33	10.28	2.156	0.639	0.406

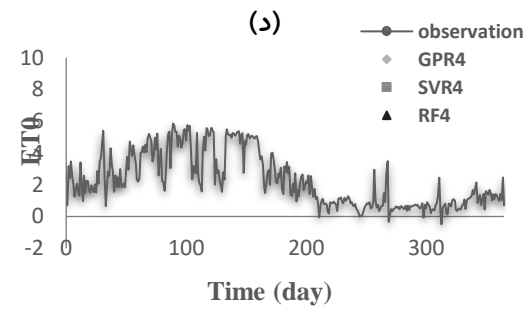
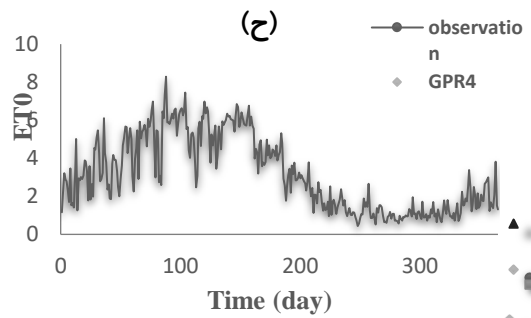
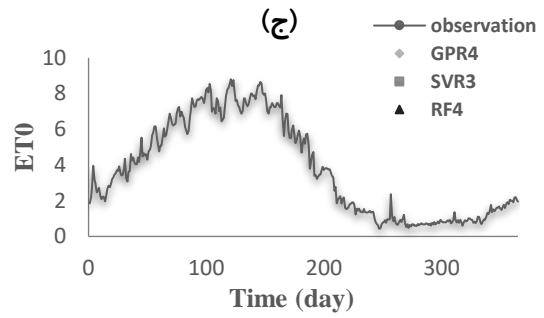
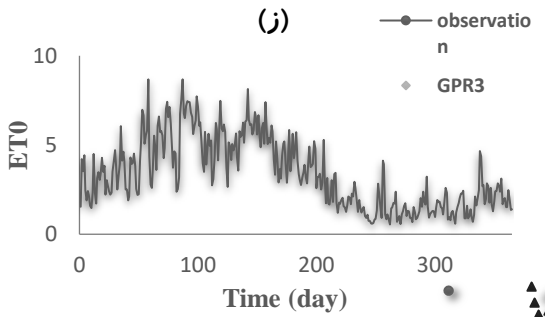
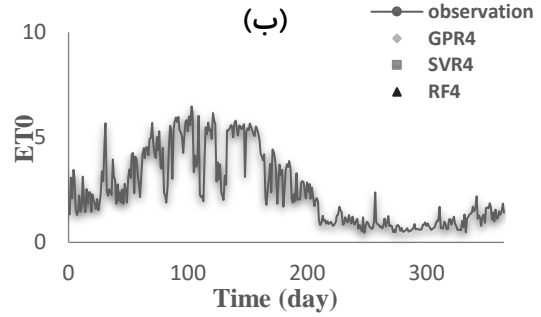
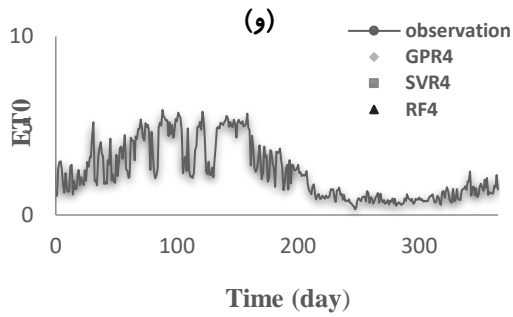
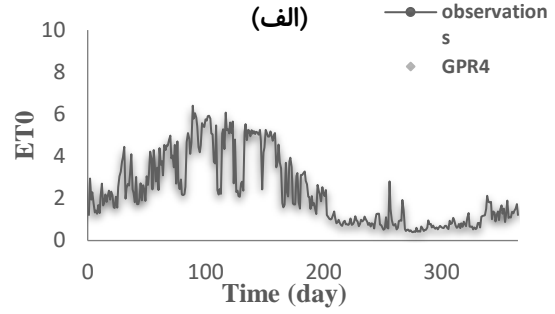
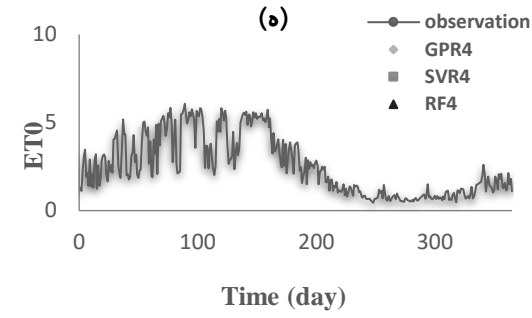
در ایستگاه‌های آستارا، بندر انزلی، رشت، رامسر و نوشهر و مدل SVR4 در ایستگاه‌های بندر ترکمن و گنبدکاووس برآوردهای دقیق‌تری از مقدار تبخیر و تعرق گیاه مرجع داشته‌اند. دقت بالای مدل‌های مذکور احتمالاً ناشی از استفاده این مدل‌ها از تمامی پارامترهای مؤثر در پدیده تبخیر و تعرق باشد. بنابراین می‌توان چنین نتیجه گرفت که پارامترهای هواشناسی دمایی میانگین، سرعت باد، رطوبت نسبی و ساعات آفتابی نقش غیرقابل‌انکاری در برآورد دقیق مقدار تبخیر و تعرق دارند.

همان‌طور که اشاره شد دقت بالای برآوردهای حاصل‌شده از مدل GPR4 در ایستگاه‌های آستارا، بندر انزلی، رشت، رامسر و نوشهر، مدل SVR4 در ایستگاه بندر ترکمن و گنبدکاووس و روش هارگریوز-سامانی در ایستگاه گرگان از شکل (۲) نیز قابل‌تشخیص می‌باشد؛ زیرا انطباق نقاط برآورد شده از روش‌های مذکور نسبت به روش‌های دیگر بر نقاط مشاهداتی بیشتر است که نشان از نزدیکی نتایج حاصله با روش فائو-پنمن-مونتیث (مشاهداتی) دارد. در نهایت، نتایج آنالیزهای انجام‌یافته نشان داد که مدل GPR4

جدول ۴- نتایج سناریوهای تعریف‌شده برای روش‌های SVR، GPR، RF و روش تجربی هارگریوز

Table 4 calculations performed for scenarios defined for methods SVR, GPR, RF and Hargreaves experimental methods

Station		Model												
		GPR 1	GPR2	GPR 3	GPR 4	SVR 1	SVR 2	SVR 3	SVR 4	RF1	RF2	RF3	RF4	HS
Astara	R	0.86	0.92	0.97	0.97	0.86	0.91	0.96	0.97	0.83	0.89	0.96	0.97	0.96
	SI	0.34	0.27	0.18	0.16	0.35	0.29	0.19	0.16	0.39	0.31	0.19	0.16	0.23
	W I	0.92	0.95	0.98	0.99	0.93	0.95	0.98	0.99	0.91	0.94	0.98	0.99	0.97
Bandar	R	0.86	0.92	0.97	0.98	0.88	0.92	0.97	0.98	0.84	0.90	0.96	0.97	0.96
	SI	0.39	0.25	0.16	0.13	0.31	0.26	0.18	0.13	0.35	0.28	0.18	0.15	0.21
	W I	0.84	0.96	0.98	0.99	0.94	0.96	0.98	0.99	0.92	0.95	0.98	0.99	0.97
Anzali	R	0.84	0.89	0.97	0.97	0.84	0.90	0.96	0.97	0.82	0.91	0.96	0.97	0.95
	SI	0.43	0.31	0.18	0.17	0.37	0.31	0.18	0.17	0.40	0.28	0.20	0.18	0.32
	W I	0.82	0.94	0.98	0.99	0.90	0.94	0.98	0.99	0.90	0.95	0.98	0.98	0.95
Ramsar	R	0.94	0.93	0.96	0.96	0.94	0.95	0.96	0.87	0.94	0.94	0.95	0.95	0.92
	SI	0.34	0.34	0.19	0.18	0.23	0.21	0.20	0.47	0.22	0.22	0.21	0.20	0.51
	W I	0.89	0.89	0.98	0.98	0.96	0.97	0.98	0.83	0.97	0.97	0.97	0.98	0.79
Nushahr	R	0.83	0.90	0.96	0.97	0.97	0.90	0.96	0.97	0.79	0.87	0.96	0.97	0.95
	SI	0.40	0.27	0.16	0.14	0.15	0.28	0.17	0.15	0.38	0.30	0.18	0.15	0.21
	W I	0.80	0.94	0.98	0.99	0.99	0.94	0.98	0.99	0.88	0.93	0.98	0.98	0.97
Sari	R	0.86	0.90	0.96	0.97	0.86	0.90	0.96	0.97	0.82	0.91	0.96	0.97	0.96
	SI	0.57	0.28	0.17	0.15	0.32	0.28	0.17	0.15	0.35	0.27	0.18	0.16	0.30
	W I	0.34	0.94	0.98	0.99	0.92	0.94	0.98	0.99	0.90	0.95	0.98	0.98	0.95
Bandar Torkaman	R	0.90	0.92	0.97	0.98	0.90	0.92	0.96	0.98	0.85	0.87	0.93	0.96	0.92
	SI	0.26	0.23	0.16	0.12	0.26	0.24	0.16	0.12	0.33	0.31	0.24	0.18	0.31
	W I	0.95	0.96	0.98	1.00	0.95	0.96	0.98	0.99	0.92	0.93	0.96	0.98	0.92
Gorgan	R	0.71	0.75	0.67	0.47	0.71	0.75	0.76	0.48	0.66	0.71	0.76	0.87	0.76
	SI	0.76	0.76	0.69	0.92	0.78	0.78	0.73	0.92	0.78	0.77	0.69	0.51	0.45
	W I	0.68	0.71	0.75	0.50	0.70	0.71	0.75	0.51	0.68	0.71	0.76	0.87	0.86
Gonbad	R	0.90	0.92	0.95	0.98	0.90	0.92	0.95	0.98	0.86	0.89	0.95	0.98	0.96
	SI	0.27	0.25	0.19	0.12	0.27	0.25	0.19	0.12	0.31	0.28	0.19	0.13	0.26
	W I	0.94	0.95	0.97	0.99	0.95	0.95	0.98	0.99	0.93	0.94	0.97	0.99	0.96



مرجع با مدل SVR1 با ضریب ویلموت ۰/۹۹۳ گزارش شده در پژوهش آن‌ها ارائه نمود.

۴- نتیجه گیری

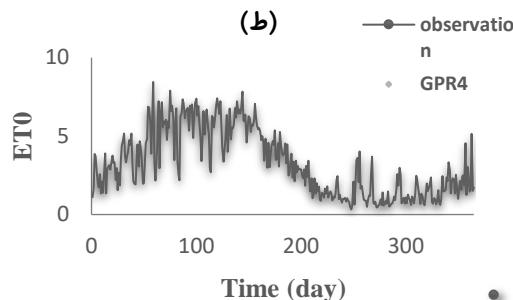
تخمین دقیق تبخیر و تعرق گیاه مرجع در برنامه‌ریزی‌های آبیاری از اهمیت ویژه‌ای برخوردار است. لذا در پژوهش حاضر سعی گردید تا میزان تبخیر و تعرق گیاه مرجع با روش‌های داده‌محور رگرسیون فرآیند گاوسی، ماشین بردار پشتیبان و جنگل تصادفی و روش تجربی هارگریوز-سامانی برای نه ایستگاه هواشناسی در شمال ایران برآورد گردیده و دقیق‌ترین روش پیشنهاد شود. در نهایت، با انجام تحلیل‌های آماری، نتایج نشان داد:

۱- در شهر آستارا، بندر انزلی، رشت، رامسر و نوشهر مدل GPR4 با شاخص پراکندگی در محدوده ۰/۱۳۲ تا ۰/۱۷۹ دقت بیشتری نسبت به مدل‌های دیگر در تخمین تبخیر و تعرق دارد.

۲- مدل SVR4 در شهر بندر ترکمن و گنبدکاووس با شاخص پراکندگی بین ۰/۱۱۶ و ۰/۱۲۰ مدل مناسب‌تری برای تخمین تبخیر و تعرق است.

۳- در شهر گرگان روش هارگریوز-سامانی برآوردهای به‌مراتب دقیق‌تری از تبخیر و تعرق گیاه مرجع دارا می‌باشد.

۴- از نتایج کلی پژوهش حاضر، چنین استنباط می‌گردد که مدل‌های GPR4 و SVR4 توانایی به‌مراتب بالاتری در تخمین هر چه دقیق‌تر تبخیر و تعرق در مناطق مرطوب داشته و برای برنامه‌ریزی‌های آبیاری و مدیریت منابع آب پیشنهاد می‌گردند.



شکل ۲- نمودارهای تغییرات زمانی مقادیر تبخیر و تعرق مرجع با استفاده از روش فائو-پنمن-مونته‌یت و روش‌های SVR، GPR، RF و روش تجربی هارگریوز الف- بندرانزلی، ب- آستارا، ج- رشت، د- رامسر، ه- نوشهر، و- ساری، ز- بندر ترکمن، ح- گرگان، ط- گنبد کاوس

Fig. 2 Charts of temporal variations reference evapotranspiration values using the FAO-56 Penman-Monteith and Hargreaves experimental methods: a) Bandar Anzali, b) Astara, c) Rasht, d) Ramsar, e) Nowshahr, f) Sari, h) Gorgan, g) Bandar Torkman, i) Gonbad Kavus

علاوه بر این، مقایسه نتایج به‌دست‌آمده با نتایج حاصل از پژوهش Huang et al. (2020) نشان داد که دقت مدل RF در پژوهش حاضر با ضریب همبستگی ۰/۹۸۸ در پیش‌بینی تبخیر و تعرق گیاه مرجع به‌مراتب بیشتر از مدل RF با ضریب همبستگی ۰/۹۷۵ گزارش شده در پژوهش آن‌ها است. (Samadianfard and Panahi (2018) تبخیر و تعرق مرجع روزانه را در ایستگاه‌های مناطق خشک و نیمه‌خشک با استفاده از روش‌های داده‌کاوی رگرسیون بردار پشتیبان و مدل M5 برآورد نمودند. مقایسه نتایج به‌دست‌آمده نشان داد که مدل SVR4، به‌عنوان یکی از دقیق‌ترین مدل‌های نتیجه شده در پژوهش حاضر با داشتن ضریب ویلموت ۰/۹۹۰ پیش‌بینی نزدیکی از تبخیر و تعرق

evapotranspiration using remote sensing, artificial neural network and comparison with the experimental method (Penman-Monteith-FAO). J. R. S. GIS Nat. Resour., 6(4), 61-75.

Basak D., Pal S. and Patranabis D. C. (2007). Support vector regression. Neu. Inform. Process., 11, 203-225.

References

- Allen R. G., Pereira L. S., Raes D. and M. Smith. (1998). Crop Evapotranspiration. Guidelines for Computing Crop Water Requirements. Irrigation and Drainage Paper No. 56, FAO, Rome, Italy, 300pp.
- Azimi A., Rangzan K., Kaboulizade M. and Khoramian M. (2016). Estimating of

- Behmanesh J., Azadtala Tape N., Montaseri M. and Besharat S. (2014). Evaluation of linear and nonlinear series models in predicting evatranspiration of reference plant in Urimia synoptic station. *J. Wat. Res. Agri. B.*, 28(1), 85-96.
- Boser B. E., Guyon I. M. and Vapnik V. N. (1992). A training algorithm for optimal margin classifiers. In Haussler D., Editor, 5th Annual ACM workshop on COLT, pages 144-152, Pittsburgh, PA.
- Breiman L. (2001). Application and analysis of random forests and machine learning. *J. Wat. Manag.*, 15(1), 5-32.
- Carter C. and Liang S. (2019). Evaluation of ten machine learning methods for estimating terrestrial evapotranspiration from remote sensing. *Int J. Appl. Earth Obs. Geoinform.*, 78, 86-92.
- Cui N., Feng Y., Gong D., Zhang Q. and Zhao L. (2017). Evaluation of random forests and generalized regression neural networks for daily reference evapotranspiration modelling. *Agri. Wat. Manag.*, 193, 163-173.
- Granata F. (2019). Evapotranspiration evaluation models based on machine learning algorithms—A comparative study. *Agri. Wat. Manag.*, 217, 303-315.
- Hashemi S., Izadyar M., and Samadianfard S. (2019). Comparison of radiant basis methods in estimating evapotranspiration. 4th International congress on natural resources development, agriculture, environment and tourism of Iran. Tabriz University of Islamic Art In collaboration with Shiraz University and Yasuj University.
- Huang J., Chen H. and McBean E. (2020). Partitioning of daily evapotranspiration using a modified shuttleworthwallace model, random Forest and support vector regression, for a cabbage farmland. *Agri. Wat. Manag.*, 228, 1-12.
- Kaviyani A., Bahmanabadi B., Daneshkararaste P. and Nazari R. (2018). Estimation of actual evatranspiration of crops using energy balance algorithms in Qazvin plain. *Echo. Hydrol.*, 5(4), 1103-1117.
- Khorsand Movaghar M. and Sima S. (2019). Comparison of the remote sensing based-energy balance models for estimating evaporation from salin lakes. *J. Geos. Inform. Technol.*, 2, 155-175.
- Khoshhal J., Zareh H. and Joshani A. (2015). Different methods for estimating reference evatranspiration by FAO evaporation pan method in the east and southeast of the country. *Quart. J. Nat. Geogr.*, 8(28), 1-16.
- Moqbeli Dameneh M. and Sanaeinejad S.H. (2018). Estimate of potential evapotranspiration in Freiman using the priestleytaylor method and remote sensing technique. *R. S. GIS Nat. Resour.*, 3, 72-84.
- Najafi P. (2006). Application of Hargreaves of Samani and Jensen – hayes modes in the evaluation of alfalfa evatranspiration in Isfahan. *Ecol. Crop.*, 2(5), 57-68.
- Saghebian M., and Roushangar K. (2019). Prediction of total and bedform roughness coefficient in alluvial channels based on experimental data via Gaussian process regression method. *Iran. J. Irrig. Drain.*, 2, 437-499.
- Samadianfard S. and Panahi S. (2018). Estimating daily reference evapotranspiration using data mining methods of support vector regression and M5 model tree. *J. Watershed Manag. Res.*, 18, 157-167.
- Saremi M. and Farhadi Bansouleh B. (2015). Determiation of effective parameterse in estimating reference crop evapotranspiration using artificial neural networks. *Iran. J. Irrig. Drain.*, 4, 614-623.
- Shabani A., Sepaskhah A. R., Bahrami M. and Razzaghi F. (2017). Combined

- application of artificial neural network and computational methods to estimate the reference evapotranspiration. *Iran. Wat. Resour. Res.*, 1, 152-162.
- Sharifian H., Dehgani, A. and Karimi rad A. (2012). Presentation of correction coefficient for Hargreaves Samsni methods as estimating evatranspiration of reference plant (case study of Gorgan synoptic station). *Soil Wat. Conserv. Res. Agri. Sci. Nat. Resour.*, 19(3), 227-236.
- Siasar H. and Honar T. (2019). Application of support vector machine, CHAID and random forest models, in estimated daily reference evapotranspiration in northern Sistan and Baluchestan Province. *Iran. J. Irrig. Drain.*, 2, 378-388.
- Trajkovic S. (2007). Hargreaves versus Penman-Monteith under humid conditions. *J. Irrig. Drain. Eng.*, 133(1), 38-42.
- Wang Y. M., Traore S. and Kerh T. (2008). Neural network approach for estimating reference evapotranspiration from limited climatic data in Burkina Faso. *WSEAS Transact. Comput.* (7), 704-713.
- Izadyar M., Hashemi S. and Samadianfard S. (2018). Estimation of refrence crop evapotranspiration by using basal temperature and mass transfer methods. 4th International congress on Natural Resources Development, Agriculture, Environment and Tourism of Iran, Tabriz University of Islamic Art in collaboration with Shiraz University and Yasuj University.

Estimation of Daily Reference Evapotranspiration in Humid Climates Using Data-Driven Methods of Gaussian Process Regression, Support Vector Regression and Random Forest

Saeed Samadianfard^{1*}, Mobarak Salarifar², Sahar Javidan² and Fatemeh Mikaeili²

¹Assist. Professor, Department of Water Engineering, Faculty of Agriculture, University of Tabriz, Tabriz, Iran

²M.Sc. Student, Department of Water Engineering, Faculty of Agriculture, University of Tabriz, Tabriz, Iran

*Corresponding author: s.samadian@tabrizu.ac.ir

Original Paper

Received: June 29, 2020

Revised: October 24, 2020

Accepted: October 26, 2020

Abstract

Accurate estimation of reference evapotranspiration has great importance in irrigation scheduling. Moreover, the lack of availability of lysimetric data has led researchers to use indirect methods, including data-driven approaches. In the present study, the ability of Gaussian process regression (GPR), support vector regression (SVR) and random forest (RF) data-driven methods was investigated to estimate the evapotranspiration of the reference plant. For this purpose, meteorological data on average temperature, wind speed, relative humidity and sunny hours in the period 2013-18 were collected in nine northern stations of Iran including Astara, Bandar Anzali, Rasht, Ramsar, Nowshahr, Sari, Turkmen port, Gorgan, and Gonbad Kavous. Evapotranspiration calculated using FAO-Penman-Monteith method was considered as the target output and four combined scenarios of meteorological parameters were considered to calibrate and validate the studied methods. The accuracy of the mentioned methods was compared using the statistical parameters of correlation coefficient, scatter index, and Wilmott's coefficient. The results showed that GPR4 model with scatter index in the range of 0.132 to 0.179 in Astara, Bandar Anzali, Rasht, Ramsar, Nowshahr and Sari stations, SVR4 model with dispersion index of 0.116 to 0.120 in Turkmen and Gonbad Kavous stations and the Hargreaves-Samani method with a scatter index of 0.509 at Gorgan station had much more accurate estimates of the evapotranspiration of the reference plant.

Keywords: Data Driven Methods; FAO Penman-Monteith Method; Hargreaves-Samani Equation; North of Iran; Reference Evapotranspiration.