

Estimation of Daily Reference Evapotranspiration with Limited Meteorological Data in Selected Iran's Semi- Arid Climates

B. Bakhtiari^{1*}, A. Mohebbi Dehaghani²
and K. Qaderi³

Abstract

Accurate estimation of evapotranspiration has a great influence on water resources management and planning, especially in arid and semi-arid areas. Different methods have been presented by researchers for evapotranspiration estimation. These include a variety of empirical equations and data-driven methods. In this study to estimate the daily reference evapotranspiration at eight semi-arid climates in Iran, three methods based on the adaptive network-based fuzzy inference system (ANFIS), support vector machines (SVM), and model tree (M5) as well as five empirical equation were used. Meteorological data including maximum and minimum temperatures, relative humidity, wind speed, and the sunshine hours were used. Eleven different combinations of these variables have been used as input variables in data-driven methods for evapotranspiration modeling for the period of 1980 to 2009. Eighty percent of the data were used for the training and twenty percent were used to test the models. The results were compared with those of the standard Penman-Monteith FAO-56 equation. Performance of the methods was evaluated using statistical indices of mean square of error (RMSE), coefficient of determination (R^2), and index of agreement (d). Support vector machines and adaptive networks based on fuzzy inference system methods presented best performance with RMSE between 0.24~1.55 ($\text{mm}\cdot\text{day}^{-1}$) in nine combination of meteorological variables. RMSE of empirical equations varied between 0.71~5.96 ($\text{mm}\cdot\text{day}^{-1}$). Blaney-Criddle and McGunness-Bordne equation presented the highest accuracy in most stations. M5 model has a lower performance compared to ANFIS and SVM methods in the studied climates.

Keywords: Daily reference evapotranspiration, Empirical equations, Semi-arid climates, ANFIS, SVM, M5.

Received: April 11, 2015

Accepted: July 8, 2015

برآورد تبخیر تعرق مرجع روزانه با حداقل داده‌های هواشناسی در اقلیم‌های نیمه خشک منتخب ایران

بهرام بختیاری^{۱*}، علیرضا محبی دهقانی^۲ و کورش قادری^۳

چکیده

تخمین صحیح و دقیق تبخیر-تعرق تأثیر بسزایی در مدیریت و برنامه‌ریزی صحیح منابع آب به ویژه در مناطق نیمه خشک و خشک دارد. روش‌های متعددی برای برآورد تبخیر-تعرق توسط محققان ارائه شده است. از جمله این روش‌ها می‌توان به انواع معادلات تجربی و روش‌های داده محور اشاره کرد. در این مطالعه از سه روش داده محور شبکه‌های تطبیقی مبتنی بر سیستم استنتاج فازی (ANFIS)، مدل درختی (M5) و ماشین‌های بردار پشتیبان (SVM) و پنج معادله تجربی برای تخمین تبخیر-تعرق روزانه در هشت اقلیم نیمه‌خشک ایران استفاده شده است. برای این منظور از داده‌های هواشناسی حداکثر و حداقل دما، رطوبت نسبی، سرعت باد و ساعات آفتابی و ۱۱ ترکیب مختلف این متغیرها بین سال‌های ۱۹۸۰ تا ۲۰۰۹ به عنوان ورودی به روش‌های داده محور برای مدل‌سازی تبخیر-تعرق بهره گرفته شد که ۸۰٪ داده‌ها برای آموزش و ۲۰٪ آن‌ها برای آزمون مدل‌ها استفاده شد. سپس نتایج حاصله با مقادیر به دست آمده از معادله استاندارد پنمن - مانتیث فائو ۵۶ مقایسه شدند. عملکرد روش‌های مورد نظر با استفاده از شاخص‌های آماری میانگین مربعات خطا (RMSE)، ضریب تبیین (R^2) و شاخص توافق (d) مورد بررسی قرار گرفت. روش‌های ماشین بردار پشتیبان و شبکه تطبیقی مبتنی بر سیستم استنتاج فازی در نه ترکیب متغیرهای هواشناسی بهترین عملکرد را با RMSE بین ۰/۲۴ تا ۱/۵۵ میلی‌متر بر روز ارائه دادند. RMSE معادلات تجربی در دامنه بین ۰/۷۱ تا ۵/۹۶ میلی‌متر بر روز متغیر بودند و از میان آن‌ها معادلات بلانی-کریدل و مک گاینس-بردنه در غالب ایستگاه‌ها بالاترین دقت را داشتند. همچنین روش M5 نسبت به دو روش ANFIS و SVM در مواجهه با ورودی‌های مختلف در اقلیم‌های مطالعاتی عملکرد پایین‌تری را از خود نشان داد.

کلمات کلیدی: تبخیر-تعرق مرجع روزانه، معادلات تجربی، اقلیم نیمه‌خشک، ANFIS، SVM، M5.

تاریخ دریافت مقاله: ۹۴/۱/۲۲

تاریخ پذیرش مقاله: ۹۴/۴/۱۷

1- Assistant Prof, Department of Water Engineering, College of Agriculture, Shahid Bahonar University of Kerman, Kerman, Iran. Email: Drbakhtiari@uk.ac.ir

2 -M.Sc. graduate in Water Resources Engineering, Department of Water Engineering, College of Agriculture, Shahid Bahonar University of Kerman, Kerman, Iran.

3- Assistant Professor, Department of Water Engineering, College of Agriculture, Shahid Bahonar University of Kerman, Kerman, Iran.

*- Corresponding Author

۱- استادیار بخش مهندسی آب، دانشکده کشاورزی، دانشگاه شهید باهنر کرمان، کرمان، ایران

۲- دانش‌آموخته کارشناسی ارشد مهندسی منابع آب، دانشکده کشاورزی، دانشگاه شهید باهنر کرمان (عضو انجمن پژوهشگران جوان دانشگاه شهید باهنر کرمان)

۳- استادیار بخش مهندسی آب، دانشکده کشاورزی، دانشگاه شهید باهنر کرمان، کرمان، ایران

*- نویسنده مسئول

پتانسیل بوده است. بختیاری و همکاران (۱۳۸۸) دو مدل تبخیر-تعرق پنمن-مانتیت فائو ۵۶ و پنمن-مانتیت انجمن مهندسين عمران آمریکا ASCE را برای بازه‌های ساعتی در اقلیم گرم خشک کرمان مورد ارزیابی قرار دادند. در این مطالعه داده‌های خرجی از دو مدل با داده‌های به‌دست آمده از لایسیمتر وزنی الکترونی طی ماه‌های فروردین لغایت شهریورماه ۱۳۸۴ مورد مقایسه قرار گرفت. نتایج نشان داد که هر دو مدل در هر ماه مقدار کمتری تبخیر-تعرق را نسبت به داده‌های لایسیمتر در مقیاس ساعتی برآورد کرده‌اند. شهبایی‌فر و همکاران (۱۳۸۹) در مطالعه‌ای به ارزیابی و مقایسه هفت روش تخمین تبخیر-تعرق با داده‌های لایسیمتر در شرایط گلخانه‌ای برای گیاه چمن پرداختند. در این تحقیق کلیه داده‌های هواشناسی به‌صورت دستی در محیط گلخانه موسسه تحقیقات گیاهپزشکی کشور طی مدت‌زمان اسفندماه ۱۳۸۵ تا خردادماه ۱۳۸۶ به‌صورت روزانه برداشت گردید. نتایج حاکی از این بود که معادله پنمن-مانتیت فائو دارای بیش‌ترین دقت و روش بلانی-کریدل کمترین دقت را برای شرایط گلخانه‌ای دارا است. رحیمی خوب و محمودی (۱۳۹۰) با مطالعه بر روی تبخیر-تعرق واقعی حوضه معرف امامه به ارزیابی شبکه‌های عصبی مصنوعی در برآورد این پارامتر با حداقل داده‌های ورودی هواشناسی و هیدرولوژیک پرداختند. میزان تبخیر-تعرق واقعی از طریق معادله بیلان آب حوضه محاسبه شد که شبکه عصبی مصنوعی با ورودی‌های کامل داده‌های هواشناسی بهترین نتایج را ارائه نمود. شاهدهی و زارعی (۱۳۹۰) به مقایسه عملکرد مدل‌های تجربی تبخیر-تعرق بر روی اقلیم مرطوب استان مازندران پرداختند. این محققان از چهار معادله تورنت وایت، پنمن، بلانی-کریدل و هارگریوز-سامانی به عنوان مدل‌های مورد ارزیابی و از خروجی‌های مدل فائو پنمن-مانتیت به‌عنوان مدل مبنا و داده‌های هواشناسی ۱۰ ساله چهار ایستگاه این استان استفاده شد. نتایج نشان دهنده عملکرد بهتر معادله بلانی-کریدل با کمترین خطا و بیش‌ترین همبستگی با مدل پنمن-مانتیت فائو بود. احمدزاده قره‌گویز و همکاران (۱۳۸۹) دقت شبکه‌های عصبی مصنوعی، ANFIS و معادلات تجربی را در برآورد تبخیر-تعرق گیاه مرجع در مناطق بسیار خشک را مورد مطالعه قرار دادند. نتایج حاکی از برتری مدل ANFIS^۴ با ورودی‌های دمای حداکثر، سرعت باد و تابش نسبت به سایر مدل‌ها بود. سلطانی و همکاران (۱۳۹۱) در مطالعه‌ای از آمار روزانه هواشناسی ۲۵ ساله (۲۰۰۸-۱۹۸۴) چهار اقلیم بسیار خشک، خشک، نیمه‌خشک و مرطوب ایران و معادله پنمن-مانتیت به‌عنوان معادله استاندارد تبخیر-تعرق استفاده نمودند. نتایج نشان داد که مدل‌های کاهش داده‌شده پنمن-مانتیت نسبت به شبکه عصبی مصنوعی و روابط هارگریوز و تورک عملکرد بهتری دارند. مرادی و همکاران (۱۳۹۱) با استفاده از معادله استاندارد پنمن-مانتیت فائو تبخیر-تعرق

امروزه رشد جمعیت و نیاز بیش از پیش به منابع آب شیرین باعث چالش جدی در ایجاد منابع آب پایدار شده است. در این راستا متخصصان به دو رویکرد عمده تولید آب و استفاده بهینه از منابع آب توجه زیادی داشته‌اند. برنامه‌ریزی، بهینه‌سازی و شبیه‌سازی در حوضه منابع آب نیازمند شناسایی و برآورد صحیح پارامترهای مختلفی از جمله تبخیر-تعرق می‌باشد. اندازه‌گیری دقیق تبخیر-تعرق نیازمند ابزاری گرانقیمت می‌باشد که امکان استفاده از آن‌ها در همه نقاط وجود ندارد. از این رو محققان همواره به دنبال روابط و روش‌های کاربردی، کم هزینه و با دقت مناسب برای برآورد صحیح مقدار این پارامتر بوده‌اند. روش‌های متعددی برای برآورد صحیح تبخیر-تعرق در سراسر دنیا توسعه داده شده است. از جمله این روش‌ها می‌توان به معادلات تجربی پرستلی-تیلور، بلانی-کریدل، مک‌گاینس-بردنه و روش‌های داده محور از جمله شبکه‌های عصبی مصنوعی، ماشین بردار پشتیبان و مدل‌های درختی اشاره نمود. یکی از مشکلات عمده معادلات تجربی انعطاف پذیری پایین در برخورد با داده‌های مختلف است که منجر به نتایج ضعیف معادلات می‌گردد. روش‌های داده محور با توجه به انعطاف‌پذیری بالا، در مدل‌سازی شرایط مختلف یک پدیده توانایی خوبی را دارا هستند. Kisi and Ciman (2009) از روش ماشین بردار پشتیبان (SVM)^۱ و شبکه عصبی مصنوعی (ANN)^۲ برای مدل‌سازی تبخیر-تعرق در منطقه‌ای در ایالت کالیفرنیا استفاده نمودند. در این تحقیق از داده‌های ET_o بدست آمده از معادله استاندارد پنمن-مانتیت فائو ۵۶ به‌عنوان ورودی‌های SVM و ANN استفاده گردید. نتایج نشان داد که SVM در هر سطح داده ورودی به مدل نتایج بهتری نسبت به ANN دارد. Douglas et al. (2009) در مطالعه‌ای بر روی تپ‌های زمین‌های دارای پوشش فلوریدا، مدل‌های تخمین تبخیر-تعرق پتانسیل تورک، پنمن و پرستلی-تیلور را با یکدیگر مقایسه کردند. در این مقایسه از ۱۸ ایستگاه همدید با مدت زمان‌های متفاوت در چهار منطقه سطح آزاد آب، تالاب، پوشش چمن و باغات مرکبات استفاده شد. حاصل این مقایسه، تأیید روش پرستلی-تیلور در تخمین تبخیر-تعرق پتانسیل در مقابل دو روش دیگر در مقیاس روزانه بود. Adeloje et al. (2012) در تحقیقی پیش‌بینی تبخیر-تعرق مرجع توسط شبکه‌های عصبی خود سازمان دهنده مورد ارزیابی قرار دادند. در این مطالعه از داده‌های روزانه سینوپتیک بین سال‌های ۲۰۰۶-۱۹۹۷ ایستگاه هواشناسی دانشگاه ادینبورگ انگلستان به‌عنوان ورودی‌های مدل‌های شبکه عصبی SOM^۳ و تجربی استفاده شد. نتایج حاکی از توانایی بالای شبکه‌های عصبی مصنوعی و معادله تجربی هارگریوز در تخمین صحیح تبخیر-تعرق

داده‌های مورد نیاز ارائه گردد. همچنین در این مطالعه سعی شده تا عملکرد روش‌های داده محور ANFIS، SVM و M5 در برآورد تبخیر-تعرق با حداقل داده‌های هواشناسی و معادلات تجربی پرکاربرد پرستلی-تیلور، هارگریوز-سامانی، مک گاینس-بردنه، بلانی-کریدل و جنسن-هیز مورد بررسی قرار گرفته و مناسبترین آن‌ها برای مناطق مطالعاتی ارائه گردد.

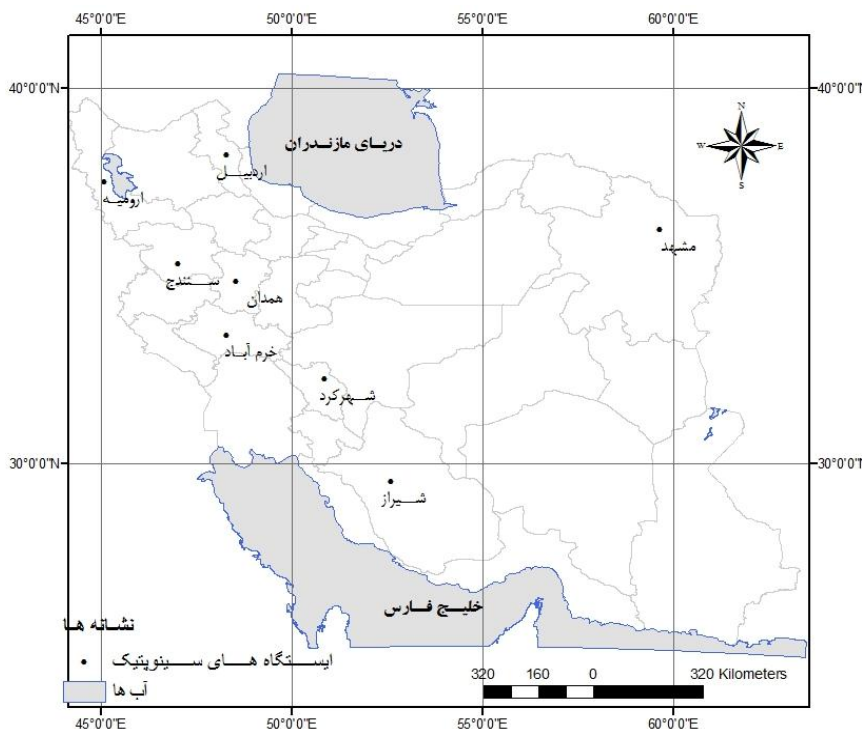
۲- مواد و روش‌ها

۲-۱- مشخصات ایستگاه‌های سینوپتیک مناطق منتخب نیمه خشک

در این مطالعه از داده‌های هواشناسی هشت ایستگاه سینوپتیک که در طبقه بندی اقلیمی دومارتن پیشرفته دارای اقلیم نیمه خشک بودند استفاده گردید. موقعیت ایستگاه‌های سینوپتیک منتخب در شکل ۱ و مشخصات آن‌ها در جدول ۱ ارائه شده است. داده‌های هواشناسی حداکثر و حداقل دما، رطوبت نسبی، سرعت باد و ساعات آفتابی در مقیاس روزانه در دوره ۳۰ ساله (۲۰۰۹-۱۹۸۰) از سازمان هواشناسی کشور جمع‌آوری گردیده است که ۲۴ سال از آن‌ها به عنوان داده‌های آموزش (۲۰۰۳-۱۹۸۰) و ۶ سال به عنوان داده‌های آزمون (۲۰۰۹-۲۰۰۴) قرار داده شد.

ساعتی گیاه مرجع را برای منطقه خشک فریمان با استفاده از داده‌های سال‌های ۲۰۰۸-۲۰۰۹ محاسبه نمود و عملکرد مدل تبخیر-تعرق ASCE و ANFIS را در برابر مدل پنمن-مانیتث فائو مورد آزمون قرارداد. نتایج نشان‌دهنده آن بود که مدل ANFIS با سه ورودی تابش، رطوبت نسبی و دما دارای بهترین دقت و عملکرد است. اسلامیان و همکاران (۲۰۰۹) نیز توانایی روش SVM و ANN در مدل‌سازی تبخیر-تعرق در محیط گلخانه را بررسی کردند. در این مدل‌سازی از داده‌های هواشناسی در درون گلخانه دانشگاه صنعتی اصفهان طی سپتامبر ۲۰۰۷ تا سپتامبر ۲۰۰۸ و داده‌های مربوط به تبخیر-تعرق مرجع از داده‌های تشت تبخیر استفاده شد. نتایج نشان داد که هر دو روش مورد مقایسه با عملکرد خوبی مدل‌سازی انجام دادند اما SVM دارای عملکرد بالاتری بود.

یکی از مشکلاتی که در بسیاری از مناطق از جمله مناطق نیمه خشک وجود دارد، محدودیت دسترسی به انواع متغیرهای هواشناسی است. هدف از انجام این تحقیق بررسی و ارزیابی چندین روش داده محور برای مدل‌سازی و پیش‌بینی مقادیر تبخیر تعرق بر اساس حداقل داده‌های اقلیمی در اقلیم‌های نیمه خشک منتخب ایران بوده است. بنابراین لازم بوده که ترکیبات مختلف ورودی با ساختارهای مختلف روش‌های داده محور مورد بررسی قرار گیرد و در نهایت مناسبترین روش برای مدل‌سازی در مناطق مطالعاتی با حداقل



شکل ۱- موقعیت جغرافیایی ایستگاه‌های مطالعاتی

جدول ۱- مختصات جغرافیایی و نرمال‌های اقلیمی دما، بارش و تبخیر-تعرق در ایستگاه‌های مختلف

ردیف	ایستگاه	طول جغرافیایی دقیقه - درجه (شرقی)	عرض جغرافیایی دقیقه - درجه (شمالی)	ارتفاع از سطح دریا (متر)	میانگین دمای سالانه (سانتی‌گراد)	میانگین بارش سالانه (میلی‌متر)	نوع اقلیم	متوسط تبخیر-تعرق پتانسیل سالانه (میلی‌متر)
۱	همدان	۴۸-۳۲	۳۴-۵۲	۱۷۴۱/۵	۱۱/۳	۳۱۷/۷	نیمه‌خشک فراسرد	۱۱۵۱/۵۹
۲	خرم‌آباد	۴۸-۱۷	۳۳-۲۶	۱۱۴۷/۸	۱۷/۲	۵۰۹	نیمه‌خشک معتدل	۱۳۸۷/۵۳
۳	مشهد	۵۹-۳۸	۳۶-۱۶	۹۹۹/۲	۱۴/۱	۲۵۵/۲	نیمه‌خشک سرد	۱۳۱۴/۰۸
۴	شهرکرد	۵۰-۵۱	۳۲-۱۷	۲۰۴۸/۹	۱۱/۸	۳۲۱/۵	نیمه‌خشک فراسرد	۱۲۴۳/۶۲
۵	شیراز	۵۲-۳۶	۲۹-۳۲	۱۴۸۱	۱۷/۷	۳۴۶	نیمه‌خشک معتدل	۱۷۱۱/۵۱
۶	اردبیل	۴۸-۱۷	۳۸-۱۵	۱۳۳۲	۹	۳۰۳/۹	نیمه‌خشک فراسرد	۱۴۳۵/۲۳
۷	ارومیه	۴۵-۵	۳۷-۳۲	۱۳۱۵/۹	۱۱/۵	۳۴۱	نیمه‌خشک سرد	۹۸۵/۵۴
۸	سندج	۴۷	۳۵-۲۰	۱۳۷۲/۴	۱۳/۴	۲۵۸/۴	نیمه‌خشک سرد	۱۳۹۲/۸۵

۲-۲-۲- روش‌های داده محور

۲-۲-۱- شبکه تطبیقی مبتنی بر سیستم استنتاج فازی (ANFIS)

اگرچه ANN یک ابزار قدرتمند برای مدل‌سازی پدیده‌های مختلف طبیعی است، اما نقص‌هایی هم در آن دیده می‌شود. ترکیب شبکه‌های عصبی مصنوعی و منطق فازی می‌تواند باعث هم‌کوشی در بهبود سرعت، دامنه خطا و تطبیق‌پذیری گردد. ANFIS با ترکیب قواعد پایه سیستم فازی با توانایی یادگیری شبکه‌های عصبی توسعه یافته است. مدل ANFIS ارائه شده یک شبکه عصبی چند لایه بر پایه سیستم فازی با پنج لایه می‌باشد. در این ساختار، گره‌های ورودی و خروجی به ترتیب نشان دهنده ورودی‌ها و پاسخ‌های خروجی است و در لایه‌های مخفی گره‌های عملگر به عنوان توابع عضویت و قواعد وجود دارند. نمونه این معماری در شکل ۲ نشان داده شده است که در آن دایره نشان دهنده گره‌های ثابت و مربع نشان‌دهنده گره‌های تطبیق‌پذیر می‌باشد. در این شکل، دو ورودی x و y و یک خروجی z نشان داده شده است.

در بسیاری از مدل‌های فازی مدل Sugeno بدلیل تفسیرپذیری،

قابلیت محاسباتی بالا و ساخته شده توسط روش‌های بهینه و تطبیق‌پذیر، در دامنه وسیعی بکار برده می‌شود. برای اولین دستورالعمل مدل فازی Sugeno، یک مجموعه قاعده با دو قانون اگر-آنگاه فازی به صورت ذیل می‌توان بررسی کرد (Shing and Jang, 1993):

قاعده ۱: $f_1 = p_1X + q_1Y + r_1$, IF X is A_1 and Y is B_1 , then

قاعده ۲: $f_2 = p_2X + q_2Y + r_2$, IF X is A_2 and Y is B_2 , then

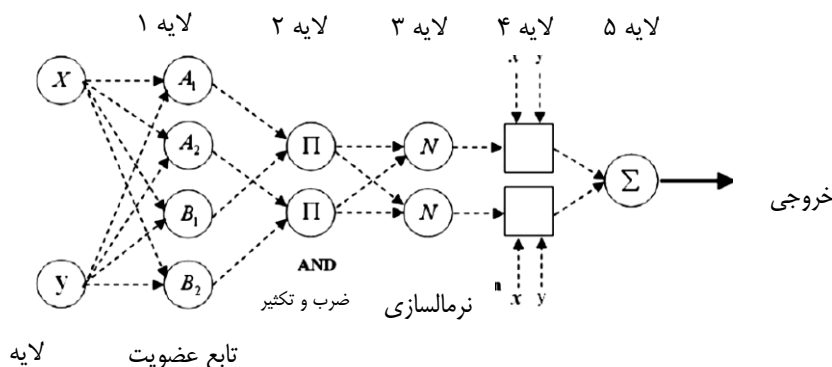
که A_i و B_i مجموعه‌های فازی ابتدایی و p_i و q_i و r_i پارامترهای طراحی هستند که در هنگام فرآیند آموزش، تعیین می‌گردند. معماری ANFIS شامل پنج لایه می‌باشد:

لایه ۱: درجه عضویت هر متغیر ورودی در هر گره از این لایه تعیین می‌شود که این تابع عضویت در رابطه ۱ نشان داده شده است.

$$O_i^1 = \mu_{A_i}(X) \quad i = 1, 2$$

$$O_i^1 = \mu_{B_i}(Y) \quad i = 3, 4$$

که μ_{A_i} و μ_{B_i} می‌توانند هرگونه تابع عضویت فازی قبول کنند. لایه ۲: این لایه شامل گره‌هایی است که ضرب در سیگنال‌های ورودی می‌گردد و نتایج را به خارج می‌فرستد. هر گره در این لایه میزان درجه تطابق ورودی‌ها با قاعده (وزن) را توسط رابطه ۲



شکل ۲- معماری شبکه‌های تطبیقی مبتنی بر سیستم استنتاج فازی (ANFIS)

اندازه‌گیری می‌کند.

$$O_i = \mu_{A_i}(x) \times \mu_{B_i}(y) \quad , \quad i = 1, 2. \quad (2)$$

لایه ۳: آیین‌گره در این لایه نسبت وزن آیین‌قاعده به مجموع وزن‌های قواعد (رابطه ۳) را محاسبه می‌کند که \bar{w}_i مقدار نرمال شده وزن‌ها است.

$$O_i^3 = \bar{w}_i = \frac{w_i}{w_1 + w_2}, \quad i = 1, 2. \quad (3)$$

لایه ۴: گره i در این لایه سهم آیین‌قاعده را در ایجاد خروجی مدل محاسبه می‌کند (رابطه ۴). پارامترهای این لایه را پارامترهای برآیند نیز می‌گویند.

$$O_i^4 = \bar{w}_i z_i = \bar{w}_i (p_i x + q_i y + r_i) \quad i = 1, 2 \quad (4)$$

لایه ۵: این لایه را لایه خروجی نیز می‌گویند و تنها گره آن خروجی نهایی است که مجموعه همه سیگنال‌های ورودی را محاسبه می‌کند (رابطه ۵)

$$O_1^5 = \text{overall output} = \sum_i \bar{w}_i z_i = \frac{\sum_i w_i z_i}{\sum_i w_i} \quad (5)$$

۲-۲-۲- مدل درختی M5

مدل M5 یکی از زیرشاخه‌های ماشین‌های آموزش (آموزش تجربی) مرتبط با ساخت یا بازیابی مدل‌ها در تعداد زیاد نمونه‌های آزمایشی است که به تخمین مسائل رایج مثل داده‌های گم شده و نویزها می‌پردازد (Quinlan, 1992). اولین بار مدل M5 توسط (Quinlan (1992) ارائه گردید و پس از آن توسط وانگ و وایتن (1997) تحت نام M5' احیا و بهبود بخشیده شد. مجموعه داده‌های آموزشی T دارای همبستگی در یک برگ (زیر مجموعه) هستند یا توسط بعضی از آزمون‌ها مجموعه T به یکسری زیرمجموعه متناظر با خروجی‌های آزمون تقسیم می‌گردند و همان مرحله بر روی زیرمجموعه‌ها به صورت بازگشتی اعمال می‌گردد. این تقسیم شدید، اغلب تولید یک ساختار دقیق می‌کند که باید به سمت عقب هرس گردد (به طور مثال جابجایی یک زیرمجموعه با یک برگ). اولین قدم در ساخت یک مدل درختی، محاسبه انحراف معیار مقادیر هدف در مجموعه داده‌های T است. برآیند آزمون، تقسیم شدن مجموعه T است. اگر از انحراف معیار داده‌های هدف $sd(T_i)$ در مجموعه T_i به عنوان شاخص خطا استفاده گردد، انتظار می‌رود که میزان خطای حاصل از آزمون کاهش یابد که این انحراف معیار به شکل رابطه ۶ نوشته می‌شود:

$$SDR = sd(T) - \sum_i \frac{|T_i|}{|T|} sd(T_i) \quad (6)$$

که T نشان دهنده یک دسته نمونه‌ها است که به برگ می‌رسد؛ T_i نشان‌دهنده زیرمجموعه‌ای از نمونه‌ها است که دارای آیین خروجی از مجموعه بالقوه است و sd نشان‌دهنده انحراف معیار است.

برای تخمین خطای یک مدل به دست آمده از داده‌های آموزشی، M5 در ابتدا میانگین باقیمانده‌های مدل را محاسبه می‌نماید. یک مدل خطی چند متغیره برای داده‌های موجود درون هر برگ یا گره درخت تصمیم‌گیری با استفاده از تکنیک‌های رگرسیون استاندارد ساخته می‌شود اما به جای استفاده از همه پارامترهای ورودی، این مدل به پارامترهایی محصور می‌گردد که توسط آزمون‌ها و یا مدل‌های خطی به زیر درخت‌ها در گره‌ها ارجاع داده شده باشند. سپس این مدل‌ها با استفاده از حذف پارامترها برای حداقل کردن خطاهای تخمینی، ساده می‌گردند که به این فرآیند ساده‌سازی می‌گویند. M5 با توجه به کمترین میزان خطا، مدل ساده شده بالادستی یا مدل زیر درخت را به عنوان مدل نهایی برای گره انتخاب می‌کند و مدل‌های دیگر را حذف می‌کند این فرآیند هرس نامیده می‌شود. اگر مدل خطی انتخاب شده باشد زیردرخت موجود در این گره به یک برگ تبدیل می‌گردد. مقدار پیش‌بینی شده در برگ، مقادیر محاسبه شده توسط مدل در آن برگ است. زمانی که مقدار یک متغیر با استفاده از یک مدل درختی پیش‌بینی شد، توسط فرآیند هموارسازی مقدار تعیین شده توسط مدل در مناسب‌ترین برگ اصلاح می‌گردد تا مقادیر پیش‌بینی شده در گره‌ها در طول مسیری از ریشه تا برگ انعکاس یابد. اگر متغیری در ادامه شاخه S_i در زیر درخت S باشد، n_i تعداد موارد آموزشی در S_i ، مقدار پیش‌بینی شده در S_i و $M(S)$ مقدار تعیین شده توسط مدل در S (مقدار پیش‌بینی ذخیره شده) است که توسط رابطه ۷ برآورد می‌شود (Quinlan, 1992).

$$PV(S) = \frac{n_i \times PV(S_i) + k \times M(S)}{n_i + k} \quad (7)$$

که در آن K ثابت هموارسازی است و مقدار آن ۱۵ فرض می‌شود.

۲-۲-۳- ماشین بردار خطی پشتیبان (SVM)

ماشین بردار پشتیبان (SVM)، یک روش طبقه‌بندی گرفته شده از تئوری آموزش آماری است و اولین بار توسط واپنیک^۵ معرفی گردید. SVMها می‌توانند برای مسائل مربوط به طبقه‌بندی و رگرسیون مورد استفاده قرار گیرند. می‌توان ماشین بردار پشتیبان را شبیه به شبکه‌های عصبی مصنوعی، مانند یک شبکه با دو لایه که وزن‌ها در اولین لایه غیر خطی و در دومین لایه خطی هستند ارائه نمود. رگرسیون بردار پشتیبان (SVR) برای توصیف رگرسیون در SVMها استفاده می‌گردد. تخمین رگرسیون با SVR، تخمین تابعی مطابق

باشد و دارای یک فرم تحلیلی شناخته شده باشد. ضرایب α_i^* و α_i با استفاده از حداکثر سازی معادله ۱۲ بدست می‌آیند.

تنها یکی از اعداد α_i^* و α_i باید صفر باشند و داده‌های نقطه‌ای مطابق با این اعداد را بردار پشتیبان می‌گویند (Kisi and Cimen, 2009).

۲-۳- معادلات تجربی

در این مطالعه نتایج پنج معادله تجربی تبخیر- تعرق پرکاربرد بلانی- کریدل، پریستلی- تیلور، جنسن- هیز، مک گاینس- برده و هاگریوز- سامانی با نتایج معادله استاندارد پنمن- ماتیت فائو ۵۶^۶ (PMF-56) مقایسه شد.

در این معادلات، ET_o تبخیر-تعرق گیاه مرجع استاندارد (mm day^{-1})، R_n تشعشع خالص روی سطح گیاه ($\text{MJ m}^{-2} \text{ day}^{-1}$)، R_a تابش خورشیدی در بالای جو ($\text{MJ m}^{-2} \text{ day}^{-1}$)، G شار گرمایی خاک ($\text{MJ m}^{-2} \text{ day}^{-1}$)، T ، T_{min} و T_{max} به ترتیب میانگین حداقل و حداکثر درجه حرارت روزانه ($^{\circ}\text{C}$)، U متوسط سرعت باد روزانه در ارتفاع ۲ متری (m s^{-1})، $(e_a - e_d)$ کمبود فشار بخار (kPa)، Δ شیب منحنی تغییرات فشار بخار ($\text{kPa } ^{\circ}\text{C}^{-1}$)، λ گرمای نهان تبخیر آب (MJ Kg^{-1})، γ ضریب سایکرومتری (رطوبت سنجی) ($\text{kPa } ^{\circ}\text{C}^{-1}$)، RH_{min} حداقل رطوبت نسبی روزانه (%، α ضریب آلیبدو (که معمولاً برابر ۱/۲۶ در نظر گرفته می‌شود)، n تعداد ساعات آفتابی روزانه، N حداکثر ساعات آفتابی روزانه و عدد ۹۰۰ فاکتور یا ضریب تعدیل است.

معادلات استفاده شده در این تحقیق به شرح معادلات ۱۳ تا ۲۰ می‌باشند.

الف- معادله پنمن-ماتیت فائو ۵۶ (Allen et al., 1998)

$$ET_o = \frac{0.408 \Delta (R_n - G) + \gamma \frac{900}{T + 273} U (e_a - e_d)}{\Delta + \gamma (1 + 0.34 U_2)} \quad (13)$$

ب - معادله بلانی- کریدل^۷ (۱۹۵۰)

$$ET_o = a + b [P(0.46T + 8.13)] \quad (14)$$

$$a = 0.0043(RH_{min}) - \left(\frac{n}{N}\right) - 1.4 \quad (15)$$

$$R(\alpha^* \alpha) = -\epsilon \sum_{i=1}^N (\alpha_i^* + \alpha_i) + \sum y_i (\alpha_i^* - \alpha_i) - \frac{1}{2} \sum_{i,j=1}^N (\alpha_i^* + \alpha_i) \times (\alpha_j^* + \alpha_j) k(x_i, x_j) \quad (12)$$

$$\text{subject to } \sum_{i=1}^N (\alpha_i^* - \alpha_i) = 0, 0 \leq \alpha_i, \alpha_i^* \leq C$$

دسته داده‌های گرفته شده از $\{(x_i, y_i)\}_n$ است که x_i بردارهای ورودی، y_i مقادیر خروجی و n تعداد کل دسته داده‌ها می‌باشد. تابع رگرسیون خطی استفاده شده تابع شماره ۸ می‌باشد (Kisi and Cimen, 2009).

$$f(x) = \omega \cdot \phi(x) + b \quad (8)$$

که $\phi(x)$ تابع غیر خطی که توسط آن x به فضای ویژه نگاشت می‌گردد، و b و ω نشان دهنده بردار وزن و ضریبی است که باید از داده‌ها تخمین زده شود. ضرایب b و ω با استفاده از کوچک‌سازی جمع ریسک‌های تجربی (اولین مولفه تابع ۹) و مولفه مرکب (دومین مولفه تابع ۹) تخمین زده می‌شود.

$$R = c \sum_{i=1}^n L_{\epsilon}(f(x_i), y_i) \frac{1}{2} \|\omega\|^2 \quad (9)$$

$$L_{\epsilon}(f(x_i), y_i) = \begin{cases} 0 & \text{for } |f(x_i) - y_i| < \epsilon \\ |f(x_i) - y_i| - \epsilon & \text{otherwise} \end{cases}$$

که c یک مقدار ثابت مثبت است (مقداری ضمنی به عنوان پارامتر کنترل ظرفیت اضافی) که بین مدل مرکب و مقادیری خطا بیشتر از ϵ مجاز تعادل ایجاد می‌کند، $\|\omega\|^2$ مولفه تنظیم است که نشان دهنده قاعده اقلیدسی است و L_{ϵ} به عنوان ϵ -غیرحساس تابع تلفات نامیده می‌شود و که میزان ریسک تجربی را اندازه‌گیری می‌کند. L_{ϵ} دارای این مزیت است که برای توصیف بردار رگرسیون نیازی به فهم داده‌های ورودی ندارد. به عنوان مثال رابطه ۱۲ نشان می‌دهد اگر اختلاف بین مقدار پیش‌بینی $f(x)$ و مقدار اندازه‌گیری شده y_i کمتر از ϵ باشد مقدار تابع تلفات برابر صفر است. انتخاب ϵ راحت‌تر از انتخاب C است و آن اغلب به عنوان نسبتی از مقادیر خروجی y_i تعریف می‌گردد. بنابراین، تابع رگرسیون غیره با استفاده از رابطه ۱۰ و ۱۱ بدست می‌آید (Kisi and Cimen, 2009).

$$f(x, \alpha, \alpha^*) = \sum_{i=1}^N (\alpha_i - \alpha^*) k(x_i, x) + b \quad (10)$$

با شرط $\alpha_i, \alpha_i^* \geq 0$ ، $\alpha_i, \alpha_i^* = 0$ ، $i = 1, \dots, N$ ، رابطه ۱۱ تابع کرنل $k(x_i, x)$ ضریب نقطه‌ای در فضای ویژه D -بعدی را توصیف می‌کند.

$$k(x_i, x) = \sum_{i=1}^D \phi_j(x) \phi_i(y) \quad (11)$$

باید به این نکته توجه داشت که مقادیر ویژه ϕ_j نیازی به محاسبه ندارند؛ ترجیحاً آنچه که نیاز است یک تابع کرنل است که بسیار ساده

$$b = 0.82 - 0.0041(RH_{\min}) + 1.07\left(\frac{n}{N}\right) + 0.066(U_{day}) + 0.006(RH_{\min}) - 0.0006(RH_{\min})(U) \quad (16)$$

RMSE نرمال حاصل از یک صفت کمتر از ۱۰٪ باشد کارکرد مدل عالی، بین ۱۰ تا ۲۰٪ خوب، بین ۲۰ تا ۳۰٪ متوسط و بیشتر از ۳۰٪ ضعیف ارزیابی میگردد. d شاخص توافق ویلموت بوده که نزدیک بودن آن به یک نیز بیانگر نزدیکی مقادیر شبیه‌سازی شده به مشاهده شده و در حقیقت کارکرد بهتر مدل است.

۳- نتایج و بحث

۳-۱- نتایج معادلات تجربی

نتایج مربوط به عملکرد معادلات تجربی برای ایستگاه‌های مطالعاتی در جدول ۲ ارائه گردید. با توجه به نتایج موجود در جدول ۲ رابطه مک گاینس- بردنه در اقلیم‌های نیمه خشک مشهد، خرم‌آباد، سنندج و شیراز با RMSE به ترتیب ۰/۹۶، ۱/۱۷، ۰/۸۶ و ۱/۰۱ میلی‌متر بر روز و شاخص توافق بین ۰/۹۶، ۰/۹۵، ۰/۹۷ و ۰/۹۶ نتایجی به مراتب بهتر از سایر معادلات از خود نشان داده است. معادله بلانی-کریدل در مناطق همدان، ارومیه و شهرکرد با RMSE به ترتیب ۰/۷۱، ۱/۰۶ و ۱/۰۸ میلی‌متر بر روز و شاخص توافق به ترتیب ۰/۹۰، ۰/۹۷، ۰/۹۴ عملکرد بهتری نسبت با سایر معادلات از خود نشان داد. در اقلیم نیمه خشک فراسرد اردبیل معادله جنسن-هیز با RMSE برابر ۱/۶۹ میلی‌متر بر روز و شاخص توافق ۰/۹۰ نتایجی با دقت و توافق بالاتری نسبت به سایر معادلات ارائه کرد. در غالب اقلیم‌های نیمه‌خشک مورد بررسی، معادلات مک گاینس- بردنه و بلانی کریدل نتایجی با دقتی نزدیک به یکدیگر نشان دادند. با توجه به نتایج موجود در جدول ۲ رابطه مک گاینس- بردنه در اقلیم‌های نیمه خشک مشهد، خرم‌آباد، سنندج و شیراز با RMSE به ترتیب ۰/۹۶، ۱/۱۷، ۰/۸۶ و ۱/۰۱ میلی‌متر بر روز و شاخص توافق بین ۰/۹۶، ۰/۹۵، ۰/۹۷ و ۰/۹۶ به مراتب بهتر از سایر معادلات از خود نشان داده است. معادله بلانی-کریدل در مناطق همدان، ارومیه و شهرکرد با RMSE به ترتیب ۰/۷۱، ۱/۰۶ و ۱/۰۸ میلی‌متر بر روز و شاخص توافق بین ۰/۹۰، ۰/۹۷، ۰/۹۴ عملکرد بهتری نسبت با سایر معادلات از خود نشان داد. در اقلیم نیمه خشک فراسرد اردبیل معادله جنسن-هیز با RMSE برابر ۱/۶۹ میلی‌متر بر روز و شاخص توافق ۰/۹۰ جواب‌هایی با دقت و توافق بالاتری نسبت به سایر معادلات ارائه کرد. در غالب اقلیم‌های نیمه خشک مورد بررسی معادلات مک گاینس- بردنه و بلانی کریدل نتایجی با دقتی نزدیک به یکدیگر نشان دادند. شکل ۳ نمونه‌ای از تغییرات شاخص‌های آماری به دست آمده حاصل از روش‌های تجربی مطالعاتی را نسبت به روش استاندارد در ایستگاه

ج- معادله پریستلی-تیلور^۹ (Oudin, 2005)

$$ET_o = \alpha \frac{\Delta (R_n - G)}{\Delta + \gamma} \frac{1}{\lambda} \quad (17)$$

د- معادله هارگریوز-سامانی^۹ (۱۹۸۵)

$$ET_o = .0023(T_{\max} - T_{\min})^{0.5} (T_{\text{mean}} + 17.8) R_a \quad (18)$$

ه- معادله جنسن-هیز^{۱۰} (Oudin, 2005):

$$ET_o = \frac{1}{\lambda} 0.025(T + 3) R_a \quad (19)$$

و- معادله مک گاینس-بردنه^{۱۱} (Oudin, 2005):

$$ET_o = \frac{1}{\lambda} R_a \left(\frac{T + 5}{68} \right) \quad (20)$$

۲-۴- شاخص‌های استاندارد آماری

جهت مقایسه عملکرد روش‌های مختلف داده محور و معادلات تجربی به کار گرفته شده برای تخمین تبخیر-تعرق از شاخص‌های استاندارد آماری ضریب تبیین (R^2)، ریشه میانگین مربعات خطا (RMSE) و شاخص توافق (d) که به ترتیب در روابط ۲۱، ۲۲ و ۲۳ به آن‌ها اشاره شده است، استفاده گردید. در روابط ۲۱، ۲۲ و ۲۳، ET_{PMF56i} آمین برآورد مدل PMF-56 (میلیمتر در روز)، \overline{ET}_{PMF56} میانگین برآوردهای مدل PMF-56 (میلی‌متر بر روز)، $ET_{Model-i}$ آمین برآورد مدل مورد آزمون (میلی‌متر بر روز) و \overline{ET}_{Model} میانگین برآوردهای مدل مورد آزمون (میلی‌متر بر روز) می‌باشد.

$$R^2 = \frac{\sum_{i=1}^N (ET_{PMF56i} - \overline{ET}_{PMF56})(ET_{Model-i} - \overline{ET}_{Model})^2}{\sum_{i=1}^N (ET_{PMF56i} - \overline{ET}_{PMF56}) \sum_{i=1}^N (ET_{Model-i} - \overline{ET}_{Model})} \quad (21)$$

$$RMSE = \frac{1}{n} \sqrt{\sum_{i=1}^N (ET_{PMF56i} - ET_{Model-i})^2} \quad (22)$$

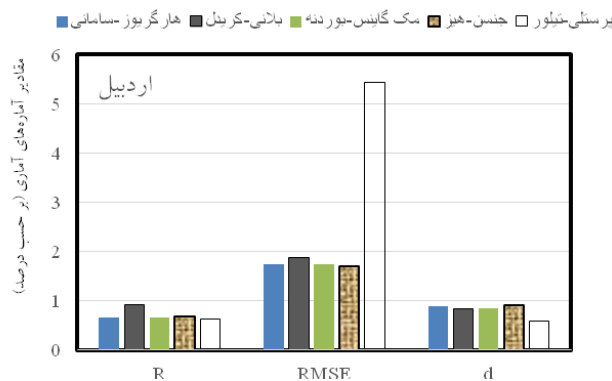
$$d = 1 - \left[\frac{\sum_{i=1}^N (ET_{PMF56i} - ET_{Model-i})^2}{\sum_{i=1}^N ((ET_{Model-i} - \overline{ET}_{PMF56}) + (ET_{Model-i} - \overline{ET}_{Model}))} \right] \quad (23)$$

R^2 بیانگر میزان احتمال هم‌بستگی میان داده‌های مشاهداتی و محاسباتی در مدل‌سازی می‌باشد. مقدار حداقل این آماره صفر و بهترین مقدار آن برابر با یک است. این ضریب در واقع نتایج تقریبی پارامتر مورد نظر در آینده را براساس مدل ریاضی تعریف شده که منطبق بر داده‌های موجود است، بیان می‌دارد. RMSE تفاوت میان مقدار پیش‌بینی شده توسط مدل و مقدار واقعی می‌باشد. RMSE یک آماره مناسب برای مقایسه خطاهای پیش‌بینی توسط یک مجموعه داده است. هر چه مقدار این آماره کمتر باشد (نزدیک به صفر) مدل تطابق بهتری نسبت به داده‌های مشاهداتی دارد. اگر

اردبیل نشان می‌دهد. همانگونه که در شکل نشان داده شده است، در نشان داده است. اقلیم اردبیل، معادله جنسن-هیز بهترین تطابق را با روش استاندارد

جدول ۲- نتایج عملکرد معادلات تجربی تبخیر-تعرق در اقلیم‌های مورد مطالعه

ایستگاه	معادلات شاخص‌ها					
	پرستلی-تیلور	جنسن-هیز	هارگریوز-سامانی	بلانی-کریدل	مک گاینس-بوردنه	
اردبیل	R ²	۰/۶۲	۰/۶۸	۰/۶۷	۰/۹۳	۰/۶۷
	RMSE	۵/۴۳	۱/۶۹	۱/۷۵	۱/۸۸	۱/۷۴
	d	۰/۵۹	۰/۹۰	۰/۸۹	۰/۸۳	۰/۸۵
همدان	R ²	۰/۷۵	۰/۸۰	۰/۵۹	۰/۸۵	۰/۸۱
	RMSE	۳/۴۳	۲/۲۲	۴/۲۴	۱/۰۶	۱/۰۶
	d	۰/۷۱	۰/۸۵	۰/۶۵	۰/۹۵	۰/۹۴
خرم‌آباد	R ²	۰/۶۷	۰/۸۱	۰/۷۸	۰/۸۶	۰/۸۲
	RMSE	۴/۶۲	۲/۶۸	۳/۰۰	۱/۲۱	۱/۱۷
	d	۰/۷۰	۰/۸۵	۰/۷۹	۰/۹۶	۰/۹۵
مشهد	R ²	۰/۸۲	۰/۹۰	۰/۸۷	۰/۸۹	۰/۹۱
	RMSE	۳/۲۶	۲/۶۷	۱/۳۷	۱/۱۸	۰/۹۸
	d	۰/۷۶	۰/۸۲	۰/۸۷	۰/۹۵	۰/۹۶
سنندج	R ²	۰/۸۲	۰/۸۹	۰/۵۳	۰/۸۹	۰/۹۰
	RMSE	۳/۴۹	۲/۲۵	۴/۹۸	۱/۱۰	۰/۸۶
	d	۰/۷۴	۰/۸۸	۰/۶۲	۰/۹۶	۰/۹۷
ارومیه	R ²	۰/۸۳	۰/۸۸	۰/۸۰	۰/۹۵	۰/۸۹
	RMSE	۲/۹۰	۱/۰۸	۱/۳۰	۰/۷۱	۰/۸۵
	d	۰/۷۷	۰/۸۶	۰/۹۲	۰/۹۷	۰/۹۶
شهرکرد	R ²	۰/۷۹	۰/۷۹	۰/۵۸	۰/۸۸	۰/۸۱
	RMSE	۳/۶۲	۱/۷۷	۵/۹۶	۱/۰۸	۱/۱۲
	d	۰/۶۹	۰/۸۹	۰/۵۳	۰/۹۴	۰/۹۳
شیراز	R ²	۰/۸۷	۰/۸۶	۰/۶۷	۰/۹۰	۰/۸۷
	RMSE	۳/۶۳	۲/۹۲	۲/۸۶	۱/۰۶	۱/۰۱
	d	۰/۷۱	۰/۸۰	۰/۷۸	۰/۹۶	۰/۹۶



شکل ۳- تغییرات شاخص‌های آماری جهت انتخاب مناسبترین مدل برآورد ET_c در مقایسه با روش PMF-56 در اقلیم نیمه خشک فراسرد اردبیل

۳-۲- نتایج روش‌های داده‌محور

پس از تعیین ترکیب متغیرهای ورودی (جدول ۳) به روش‌های داده محور، عملکرد نه مدل برتر داده محور در جدول ۴ ارائه گردید. لازم به ذکر است که نتایج ارائه شده در جدول ۵ نتایج مربوط به روش M5 دیده نمی‌شود. دلیل این موضوع عملکرد پایین‌تر روش M5 نسبت به دو روش ANFIS و SVM در مواجهه با ورودی‌های مختلف است.

جدول ۳- ترکیب‌های متغیرهای هواشناسی ورودی

مدل	نوع داده‌های ورودی
۱	T_{MEAN} (متوسط دما)
۲	RH (رطوبت نسبی)
۳	R_S (تابش)
۴	U (سرعت باد)
۵	T_{MEAN} -RH (متوسط دما-رطوبت نسبی)
۶	T_{MEAN} - R_S (متوسط دما-تابش)
۷	T_{MEAN} -U (متوسط دما-سرعت باد)
۸	T_{MEAN} -RH- R_S (متوسط دما-رطوبت نسبی-تابش)
۹	T_{MEAN} -RH-U (متوسط دما-رطوبت نسبی-سرعت باد)
۱۰	T_{MEAN} - R_S -U (متوسط دما-تابش-سرعت باد)
۱۱	T_{MEAN} -RH- R_S -U (متوسط دما-رطوبت نسبی-تابش-سرعت باد)

با توجه به نتایج جدول ۴ در کلیه مناطق مورد مطالعه بهترین ترکیب، ترکیب ورودی با چهار متغیر هواشناسی تعیین شد که بالاترین دقت و توافق ($0.62 < RMSE < 0.74$ و $0.98 < d < 1.00$) را با داده‌های معادله استاندارد دارا می‌باشد. در اقلیم‌های شیراز، شهرکرد، سنندج، ارومیه، مشهد، خرم‌آباد و همدان ترکیب متغیرهای متوسط دما، سرعت باد و تابش خالص روزانه به عنوان ورودی با RMSE بین 0.29 تا 0.4 میلی‌متر بر روز بهترین نتیجه را در بین ترکیب‌هایی با سه ورودی را نشان داد و رتبه دوم را در میان مدل‌های داده محور کسب کرد. در ایستگاه اردبیل سه متغیر متوسط دما، رطوبت نسبی و تابش خالص خورشیدی روزانه بهترین نتیجه را با RMSE برابر $1/11$ میلی‌متر بر روز در میان مدل‌هایی با سه ترکیب از خود نشان دادند. در همه ایستگاه‌ها به غیر از اردبیل و ارومیه مدل‌های داده محور با دو ورودی هواشناسی متوسط دما و سرعت باد روزانه با RMSE بین 0.60 تا 0.87 میلی‌متر بر روز رتبه برتر در میان مدل‌هایی با دو ورودی را کسب کردند. در اقلیم نیمه‌خشک فراسرد اردبیل مدل، ANFIS با دو ورودی هواشناسی

متوسط دما و رطوبت نسبی روزانه نسبت به سایر مدل‌ها با همین ورودی‌ها با RMSE برابر 0.87 میلی‌متر بر روز و R^2 برابر 0.89 بهترین عملکرد را نشان داد. در بین مدل‌هایی که دارای تنها یک متغیر ورودی بودند اقلیم‌های نیمه‌خشک اربیل، همدان، خرم‌آباد، ارومیه و سنندج با تنها ورودی هواشناسی متوسط دمای روزانه با مقدار شاخص RMSE بین $1/01$ تا $1/45$ میلی‌متر بر روز بهترین نتایج را نشان دادند. مدل‌های ANFIS و SVM در اقلیم نیمه‌خشک فراسرد شهرکرد که تنها متغیر ورودیشان به ترتیب دمای متوسط روزانه و تابش خالص آفتاب بود با RMSE برابر $1/23$ میلی‌متر بر روز دارای خطای مشابه هستند اما مدل ANFIS با همبستگی بیشتر (R^2 برابر 0.77) می‌تواند برآوردهایی به مراتب نزدیک‌تر به برآوردهای مدل PMF-56 تولید نماید. در اقلیم نیمه‌خشک معتدل شیراز مدل‌های SVM تنها با یک ورودی دمای متوسط و یا تابش خالص خورشیدی روزانه با RMSE برابر $1/07$ میلی‌متر بر روز دارای خطای یکسانی هستند اما مدل SVM با ورودی متوسط دما و شاخص توافق برابر 0.94 نتایجی نزدیک‌تر به مدل استاندارد در بین مدل‌های تک ورودی نشان داد. در ایستگاه مشهد بر خلاف سایر مناطق نیمه‌خشک مورد مطالعه، تابش خالص روزانه به عنوان تنها ورودی به مدل ANFIS با RMSE برابر 0.96 میلی‌متر بر روز به عنوان بهترین مدل داده محور با یک متغیر هواشناسی تعیین گردید. با توجه به اینکه در مناطق مطالعاتی مقادیر آماره‌های خطا برای سناریوهای مختلف مدل M5 نسبت به سایر مدل‌های داده‌محور مورد بررسی، معنی دارتر بوده است لذا این مدل در ارجحیت استفاده قرار نمی‌گیرد. به طور مثال یکی از سناریوهای مدل M5 با تنها ورودی هواشناسی متوسط سرعت باد روزانه و RMSE برابر $1/94$ میلی‌متر در روز و R^2 برابر 0.14 نامطلوب‌ترین نتایج را در اقلیم نیمه‌خشک سرد ارومیه ارائه داده است. به همین ترتیب مقادیر شاخص توافق (d) و سایر آماره‌های مورد آزمون در این روش پایین‌تر از سایر روش‌ها به دست آمده است.

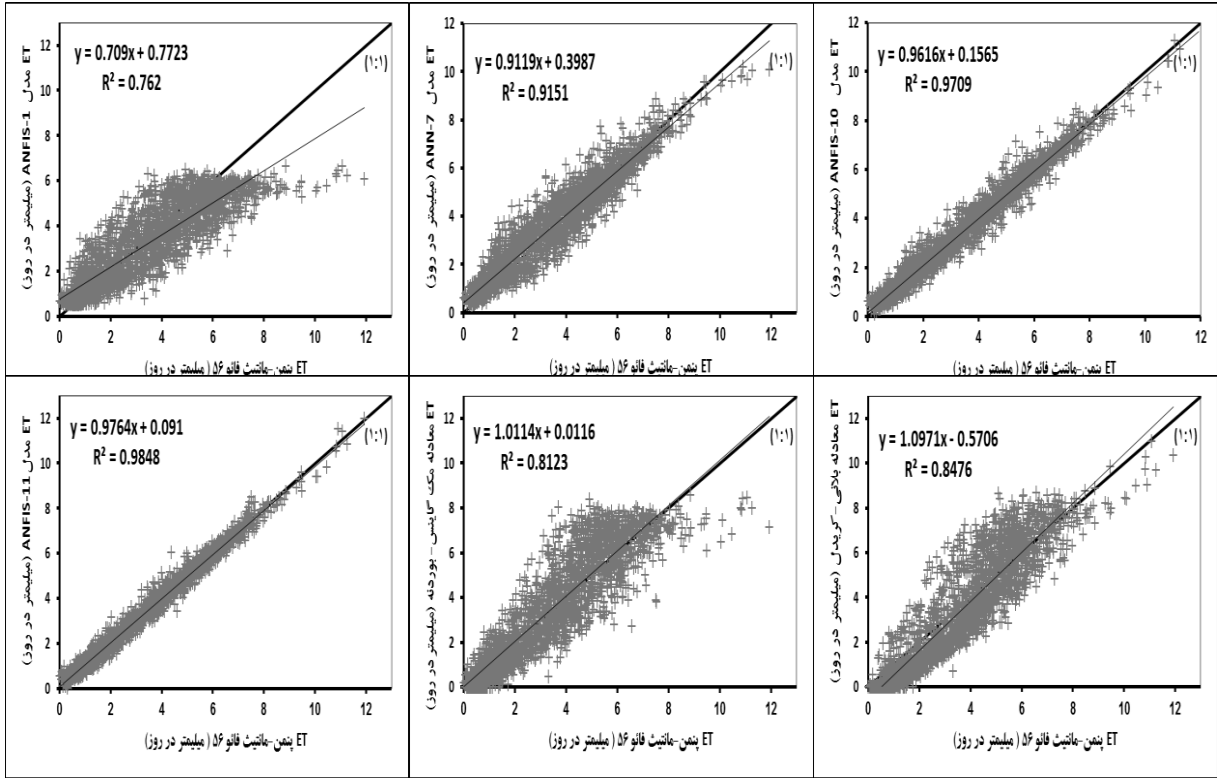
در تمامی اقلیم‌های نیمه‌خشک همدان، خرم‌آباد، ارومیه، شهرکرد و شیراز موثرترین متغیر هواشناسی در برآورد تبخیر-تعرق پتانسیل توسط روش‌های داده‌محور، متغیر دما است. پس از متغیر دما به ترتیب تأثیرگذاری متغیرهای سرعت باد، تابش خالص آفتاب و رطوبت نسبی مهمترین عوامل موثر بر دقت روش‌های داده‌محورند. در اقلیم مشهد مهمترین عوامل تأثیرگذار بر تعیین تبخیر-تعرق پتانسیل عامل تابش خالص می‌باشد و پس از آن به ترتیب تأثیرگذاری دما، سرعت باد و رطوبت نسبی می‌باشند.

تحقیق حاضر) و همچنین اقلیم بسیار خشک برتری داشته است. همچنین نتایج پژوهش حاضر با نتایج بیات‌ورکشی و همکاران (۱۳۸۸) در اقلیم نیمه‌خشک فراسرد همدان همخوانی داشته است. در شکل‌های ۴، ۵ و ۶ مقایسه برترین مدل‌های داده‌محور با تعداد مختلف ورودی و دو معادله تجربی برتر با نتایج معادله پنمن-مانتیت فائو ۵۶ ارائه شده است.

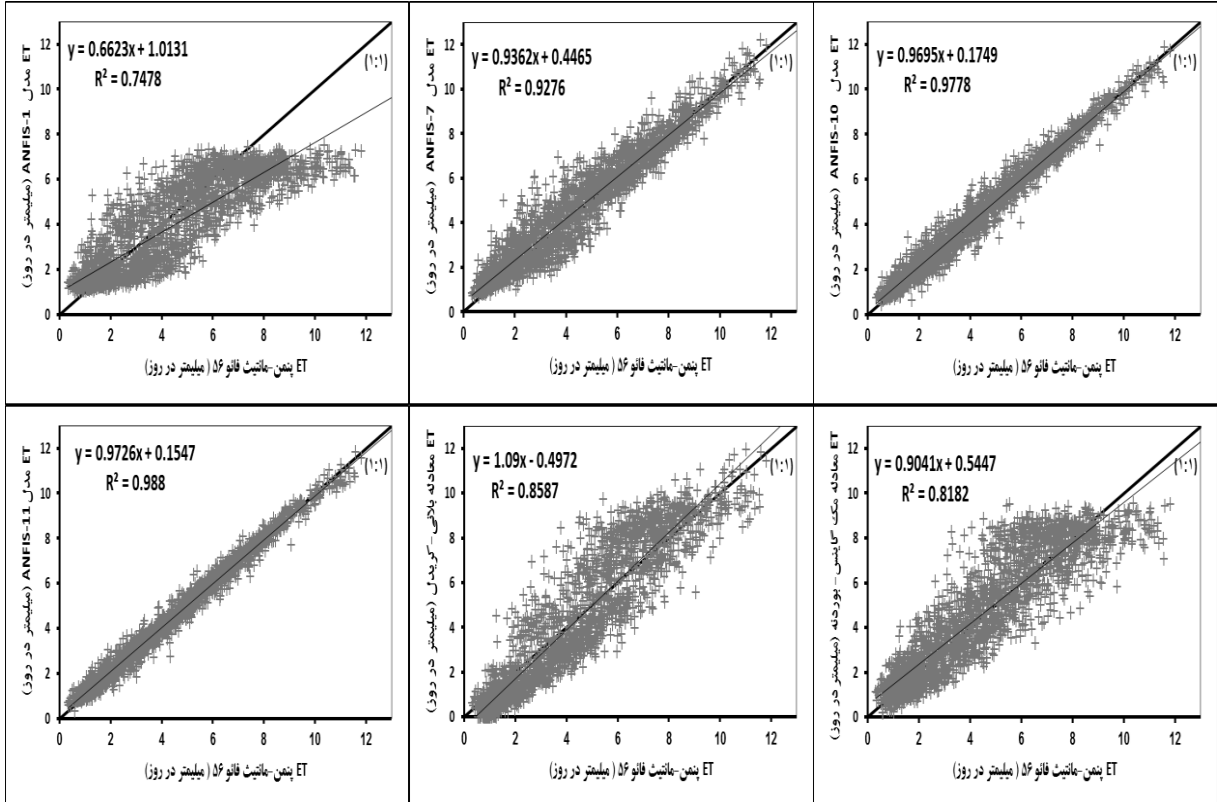
در اقلیم نیمه خشک اردبیل مهمترین متغیر هواشناسی ورودی به مدل‌های داده‌محور، متغیر دما و پس از آن به ترتیب اهمیت رطوبت نسبی، تابش خالص خورشیدی و سرعت باد هستند. مقایسه نتایج حاصل از این تحقیق با نتایج حاصل از مطالعه احمدزاده قره‌گویز و همکاران (۱۳۸۹) که برای اقلیم بسیارخشک انجام دادند نشان می‌دهد که روش ANFIS نسبت به سایر روش‌های تجربی در مناطق نیمه‌خشک با زیر اقلیم‌های فراسرد، سرد و معتدل (حاصل از

جدول ۴- نتایج مدل‌های داده‌محور برتر هر ایستگاه

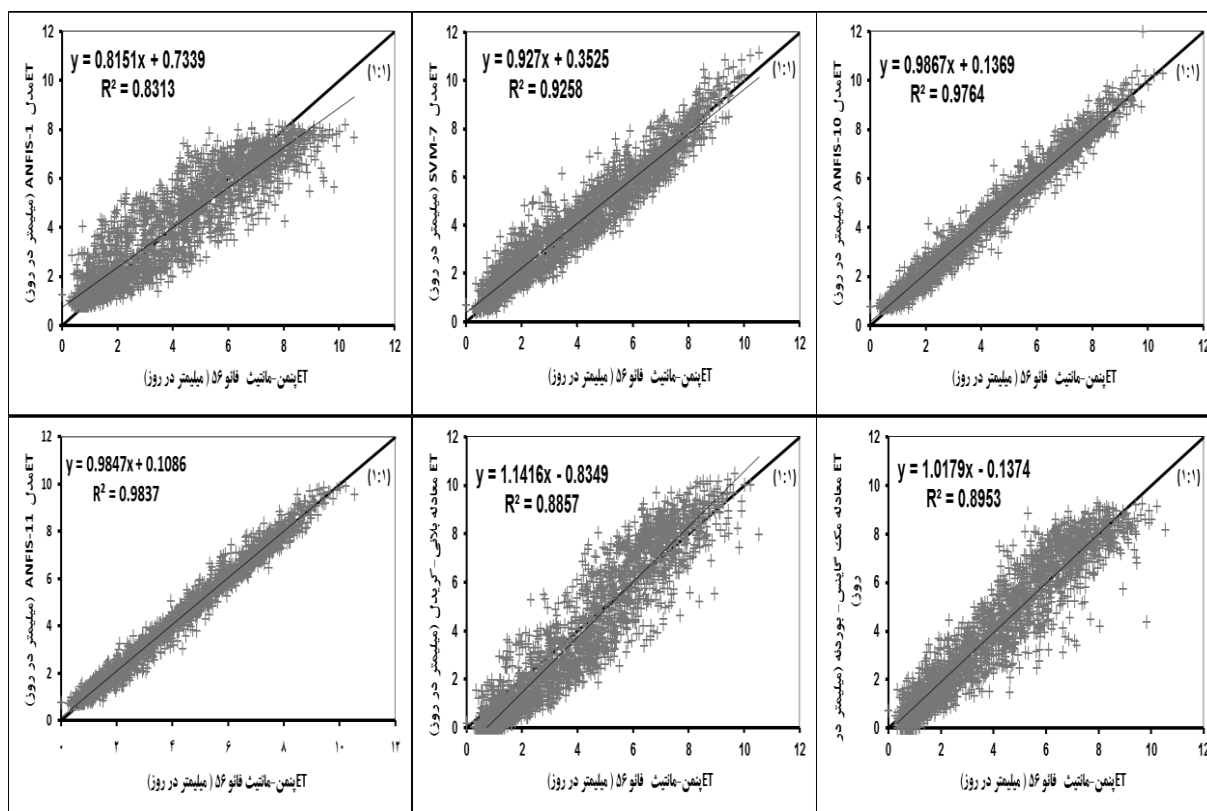
ایستگاه	مدل عملکرد	(۱)	(۲)	(۳)	(۴)	(۵)	(۶)	(۷)	(۸)	(۹)	(۱۰)	(۱۱)
		اردبیل	Model R ² RMSE d	SVM ۰/۷۰ ۱/۴۵ ۰/۹۰	ANFIS ۰/۸۱ ۱/۵۰ ۰/۸۸	SVM ۰/۷۲ ۱/۴۰ ۰/۹۱	SVM ۰/۸۹ ۱/۲۱ ۰/۹۳	ANFIS ۰/۸۷ ۱/۲۱ ۰/۹۷	ANFIS ۰/۹۴ ۱/۶۳ ۰/۹۸	ANFIS ۰/۸۹ ۱/۱۱ ۰/۹۵	ANFIS ۰/۸۷ ۱/۱۱ ۰/۹۵	ANFIS ۰/۸۷ ۱/۱۱ ۰/۹۵
همدان	Model R ² RMSE d	ANFIS ۰/۷۶ ۱/۰۹ ۰/۹۲	ANFIS ۰/۷۰ ۱/۲۶ ۰/۸۸	ANFIS ۰/۹۱ ۱/۶۵ ۰/۹۸	ANFIS ۰/۸۱ ۱/۰۱ ۰/۹۱	SVM ۰/۸۸ ۱/۰۳ ۰/۹۳	ANFIS ۰/۹۱ ۱/۰۳ ۰/۹۳	ANFIS ۰/۹۹ ۱/۳۷ ۰/۹۹	ANFIS ۰/۹۹ ۱/۳۷ ۰/۹۹	ANFIS ۰/۹۴ ۱/۵۲ ۰/۹۸	ANFIS ۰/۹۴ ۱/۵۲ ۰/۹۸	ANFIS ۰/۹۸ ۱/۲۷ ۱
خرم‌آباد	Model R ² RMSE d	ANFIS ۰/۷۵ ۱/۴۴ ۰/۹۰	SVM ۰/۷۴ ۱/۵۵ ۰/۹۰	ANFIS ۰/۹۳ ۱/۸۷ ۰/۹۸	SVM ۰/۸۴ ۱/۲۱ ۰/۹۴	SVM ۰/۷۷ ۱/۳۸ ۰/۹۱	SVM ۰/۸۴ ۱/۲۲ ۰/۹۴	ANFIS ۰/۹۸ ۱/۴۰ ۰/۹۹	ANFIS ۰/۹۸ ۱/۴۰ ۰/۹۹	ANFIS ۰/۹۵ ۱/۶۰ ۰/۹۹	ANFIS ۰/۹۵ ۱/۶۰ ۰/۹۹	ANFIS ۰/۹۹ ۱/۳۰ ۱/۰۰
مشهد	Model R ² RMSE d	ANFIS ۰/۷۹ ۱/۰۹ ۰/۹۴	ANFIS ۰/۸۴ ۱/۰۹ ۰/۹۵	ANFIS ۰/۹۲ ۱/۶۸ ۰/۹۸	ANFIS ۰/۹۱ ۱/۷۴ ۰/۹۷	ANFIS ۰/۸۰ ۱/۰۸ ۰/۹۴	ANFIS ۰/۹۱ ۱/۷۳ ۰/۹۷	ANFIS ۰/۹۸ ۱/۳۷ ۰/۹۹	ANFIS ۰/۹۸ ۱/۳۷ ۰/۹۹	ANFIS ۰/۹۵ ۱/۵۶ ۰/۹۹	ANFIS ۰/۹۵ ۱/۵۶ ۰/۹۹	ANFIS ۰/۹۹ ۱/۲۴ ۱/۰۰
ارومیه	Model R ² RMSE d	SVM ۰/۸۵ ۱/۰۹ ۰/۹۴	SVM ۰/۸۲ ۱/۰۶ ۰/۹۲	ANFIS ۰/۹۱ ۱/۶۲ ۰/۹۷	ANFIS ۰/۹۱ ۱/۶۲ ۰/۹۷	ANFIS ۰/۹۲ ۱/۸۸ ۰/۹۴	SVM ۰/۹۴ ۱/۲۹ ۰/۹۶	ANFIS ۰/۹۷ ۱/۳۷ ۰/۹۹	ANFIS ۰/۹۷ ۱/۳۷ ۰/۹۹	ANFIS ۰/۹۵ ۱/۴۴ ۰/۹۹	ANFIS ۰/۹۵ ۱/۴۴ ۰/۹۹	SVM ۰/۹۸ ۱/۲۶ ۱/۰۰
سنندج	Model R ² RMSE d	ANFIS ۰/۸۳ ۱/۰۱ ۰/۹۵	ANFIS ۰/۷۸ ۱/۱۶ ۰/۹۳	ANFIS ۰/۹۳ ۱/۶۷ ۰/۹۸	ANFIS ۰/۸۸ ۱/۸۴ ۰/۹۷	ANFIS ۰/۸۵ ۱/۹۷ ۰/۹۶	ANFIS ۰/۸۹ ۱/۸۲ ۰/۹۷	ANFIS ۰/۹۸ ۱/۳۸ ۰/۹۹	ANFIS ۰/۹۸ ۱/۳۸ ۰/۹۹	ANFIS ۰/۹۴ ۱/۵۹ ۰/۹۹	ANFIS ۰/۹۴ ۱/۵۹ ۰/۹۹	ANFIS ۰/۹۸ ۱/۳۲ ۱/۰۰
شهرکرد	Model R ² RMSE d	ANFIS ۰/۷۷ ۱/۲۳ ۰/۸۹	ANFIS ۰/۷۳ ۱/۲۳ ۰/۹۰	ANFIS ۰/۷۳ ۱/۲۳ ۰/۹۰	ANFIS ۰/۸۹ ۱/۷۳ ۰/۹۷	ANFIS ۰/۸۱ ۱/۱۹ ۰/۹۰	SVM ۰/۸۴ ۱/۰۲ ۰/۹۴	SVM ۰/۹۷ ۱/۴۰ ۰/۹۹	SVM ۰/۹۷ ۱/۴۰ ۰/۹۹	ANFIS ۰/۹۳ ۱/۶۰ ۰/۹۸	ANFIS ۰/۹۳ ۱/۶۰ ۰/۹۸	SVM ۰/۹۸ ۱/۳۷ ۰/۹۹
شیراز	Model R ² RMSE d	SVM ۰/۸۰ ۱/۰۷ ۰/۹۴	SVM ۰/۸۰ ۱/۰۷ ۰/۹۴	ANFIS ۰/۹۶ ۱/۶۰ ۰/۹۸	ANFIS ۰/۸۹ ۱/۸۷ ۰/۹۶	ANFIS ۰/۸۳ ۱/۹۸ ۰/۹۵	ANFIS ۰/۸۹ ۱/۸۲ ۰/۹۷	ANFIS ۰/۹۸ ۱/۲۹ ۱/۰۰	ANFIS ۰/۹۸ ۱/۲۹ ۱/۰۰	SVM ۰/۹۶ ۱/۴۸ ۰/۹۹	SVM ۰/۹۶ ۱/۴۸ ۰/۹۹	ANFIS ۰/۹۹ ۱/۲۵ ۱/۰۰



شکل ۴- مقایسه ET مدل‌های برتر با PMF-56 در اقلیم نیمه‌خشک فراسرد همدان



شکل ۵- مقایسه ET مدل‌های برتر با PMF-56 در اقلیم نیمه‌خشک معتدل خرم‌آباد



شکل ۶- مقایسه ET_0 مدل‌های برتر با PMF-56 در اقلیم نیمه‌خشک سرد سنج

۴- نتیجه‌گیری

همانطور که گفته شد در بسیار از نقاط برای برآورد پارامتر تبخیر-تعرق پتانسیل به روش مستقیم، محدودیت‌های مختلفی وجود دارد. برای برآورد آسانتر این پارامتر روش‌ها و معادلات مختلفی ارائه و با استفاده از شاخص‌های آماری استاندارد مقایسه گردیده است. همانطور که در بخش نتایج نیز اشاره گردید، روش‌های داده محور ANFIS و SVM با تعداد متغیر ورودی مختلف، دقت نسبتاً مناسبی در برآورد ET_0 برای اقلیم‌های نیمه‌خشک مورد مطالعه در این پژوهش ارائه دادند که می‌تواند در اموری همچون مدل‌سازی هیدرولوژی و برنامه‌ریزی آبیاری با محدودیت داده‌های اقلیمی استفاده شود. از طرفی روش‌های داده محور به مراتب نتایجی دقیقتر نسبت به معادلات تجربی در ایستگاه‌های مورد بررسی در اختیار قرار می‌دهند. با توجه به نتایج به دست آمده در این پژوهش، نامطلوب‌ترین مدل داده محور، در اقلیم نیمه خشک فراسرد اردبیل مدل ماشین بردار خطی با ورودی متغیر هواشناسی سرعت باد است. همچنین مؤثرترین عوامل جوی بر روی عملکرد مدل‌ها به ترتیب اهمیت، میانگین دما، میانگین رطوبت نسبی، سرعت باد و تشعشع خالص خورشیدی در این اقلیم است. با توجه به نتایج شاخص‌های آماری برای مدل‌های برتر، مهم‌ترین عامل در تعیین و تخمین ET_0

در اقلیم نیمه خشک فرا سرد همدان، متغیر متوسط دمای روزانه بوده و متغیرهای هواشناسی متوسط سرعت باد روزانه، تابش خالص خورشیدی روزانه و متوسط رطوبت نسبی روزانه در درجات بعدی اهمیت قرار دارند. برای اقلیم نیمه‌خشک معتدل خرم‌آباد در میان مدل‌های داده‌محور مورد آزمون، مدل ANFIS-10 و ANFIS-11 با RMSE به ترتیب برابر 0.3 و 0.4 میلی‌متر بر روز و ضریب تبیین نزدیک به یک، دقت بالاتری را در تخمین تبخیر-تعرق پتانسیل نشان داده است. در اقلیم نیمه خشک سرد مشهد مدل ANFIS-10 با RMSE برابر 0.37 میلی‌متر بر روز و ضریب تبیین 0.98 مطلوب‌ترین برآورد را ارائه نمود. از طرفی تابش خالص خورشیدی به سطح زمین موثرترین متغیر هواشناسی در تخمین دقیق‌تر تبخیر-تعرق پتانسیل در اقلیم نیمه خشک سردسیر مشهد تعیین گردید و پس از آن تابش، دما و سرعت باد موثرترین عوامل می‌باشند. در اقلیم ارومیه نیز مدل داده محور SVM-11 با چهار متغیر هواشناسی ورودی و با RMSE برابر 0.26 میلی‌متر بر روز و R^2 برابر 0.98 بالاترین دقت و عملکرد را در میان مدل‌های داده محور دارا است. متغیر دما عامل موثر در تخمین دقیق‌تر ET_0 توسط مدل‌های داده محور در این اقلیم نیمه خشک سرد بوده و پس از دما، سرعت باد و تابش خالص خورشیدی و رطوبت نسبی درجات بعدی تأثیرگذاری را

(مطالعه موردی اقلیم کرمان). علوم و فنون کشاورزی و منابع طبیعی، علوم آب و خاک، سال ۱۳، شماره ۵۰: ۲۶-۱۳.

بیات‌ورکشی م، زارع‌ایبانه ح، معروفی ص، سبزی‌پرور ع، سلطانی ف (۱۳۸۸) شبیه‌سازی تبخیرتقرق گیاه مرجع به روش هوش مصنوعی و روش‌های تجربی در مقایسه با اندازه‌گیری‌های لایسمتری در اقلیم نیمه‌خشک سرد همدان. مجله پژوهش‌های حفاظت آب و خاک، جلد ۱۶، شماره چهار: ۱۰۰-۷۹.

رحیمی خوب ع، محمودی ع (۱۳۹۰) برآورد تبخیر-تقرق واقعی از سطح حوضه آبخیز با استفاده از مدل‌های شبکه عصبی و حداقل داده‌های هواشناسی - مطالعه موردی حوضه معرفه امامه. تحقیقات منابع آب ایران، جلد ۷، شماره ۴: ۶۱-۵۱.

شاهدی ک، زارعی م (۱۳۹۰) ارزیابی روش‌های برآورد تبخیر-تقرق پتانسیل در استان مازندران. فصلنامه علمی پژوهشی مهندسی آبیاری و آب سال اول، شماره ۳: ۲۱-۱۱.

شهبایی فر م، عساری م، کوچک زاده م، عزیزی زهان ع (۱۳۸۶) ارزیابی شش روش محاسباتی تبخیر-تقرق سطح مرجع با داده‌های لایسمتری در شرایط گلخانه‌ای. اولین کارگاه فنی ارتقاء کارایی مصرف آب با کشت محصولات گلخانه‌ای کمیته آبیاری و زهکشی ایران-تهران.

علیزاده ا (۱۳۸۷) اصول هیدرولوژی کاربردی. انتشارات دانشگاه امام رضا (علیه‌السلام)، چاپ ۲۴، ۸۷۰ صفحه.

مرادی ح، تمنا م، انصاری ح، نادریان فر م (۱۳۹۱) سیستم‌های استنتاج فازی جهت برآورد تبخیرتقرق مرجع ساعتی (مطالعه موردی: منطقه فریمان). مجله پژوهش‌های حفاظت آب و خاک. جلد نوزدهم، شماره ۱: ۱۶۸-۱۵۳.

Allen RG, Pereira LS, Raes D and Smith M (1998) Crop evapotranspiration guidelines for computing crop water requirements. FAO Irrigation and Drainage, Paper No. 56., FAO, Rome.

Bhattacharya B and Solomatine DP (2005) Machine learning in sedimentation modeling. Neural Networks, 19: 208-214.

Bhattacharya B and Solomatine DP (2005) Neural networks and M5 model trees in modelling water level-discharge relationship. Neuro computing, 63: 381-396.

Blaney HF and Criddle WD (1950) Determining water requirements in irrigated areas from climatological and irrigation data. Soil conservation service

دارند. در اقلیم نیمه خشک سرد سنج بهترین عملکرد مربوط به مدل ANFIS-10 با سه متغیر هواشناسی ورودی بوده است. در این بین مدل SVM-4 با تک ورودی سرعت باد روزانه نامطلوب‌ترین عملکرد را نشان داد. در اقلیم نیمه خشک فرا سرد شهرکرد مهم‌ترین عامل تأثیرگذار در بین متغیرهای هواشناسی در دقت مدل‌های داده محور، متغیر دما بوده و پس از آن ترکیب دما با متغیرهای سرعت باد، تابش خالص خورشیدی و رطوبت نسبی بیشترین اهمیت را داشته‌اند. در صورت تمایل بر استفاده از معادلات تجربی می‌توان از دو معادله مک‌گاینس-بردنه و بلانی‌کریدل جهت برآورد تبخیر-تقرق پتانسیل در اقلیم‌های نیمه‌خشک مورد نظر استفاده نمود. پیشنهاد می‌شود مدل‌های ANFIS و SVM به عنوان یک مدل تخمین داده‌های ET برای مدلسازی هیدرولوژی مناطق مورد مطالعه و یا مناطقی با ویژگی‌های هیدرومتئورولوژی مشابه مورد استفاده قرار گیرند. همچنین بهبود و ارتقاء عملکرد روش ANFIS را می‌توان با استفاده از توابع عضویت متنوع‌تر و بیشتر مورد آزمون قرار داد. از طرفی با توجه به اینکه مقدار حداکثری ET در برخی زمینه‌های آبیاری و هیدرولوژی دارای اهمیت است می‌توان دقت مدلسازی مدل‌های داده محور و معادلات تجربی در نقاط حداکثری (Peak) تبخیر-تقرق پتانسیل را مورد بررسی قرار داد.

پی‌نوشت‌ها

- 1-Support Vector Machines
- 2-Artificial Neural Networks
- 3-Self-Organizing Map
- 4-Adaptive Network based Fuzzy Inference System
- 5-Vapnik
- 6-Penman-Manteith FAO 56
- 7-Blaney - Criddle
- 8-Priestley - Taylor
- 9-Hargreaves - Samani
- 10-Jensen-Haise
- 11-McGuinness - Bordne

۵- مراجع

احمدزاده قره‌گیوز ک، میرلطیفی س م، محمدی ک (۱۳۸۹) مقایسه سیستم‌های هوشمند (ANN و ANFIS) در تخمین میزان تبخیر-تقرق گیاه مرجع در مناطق بسیار خشک. نشریه آب و خاک، جلد ۲۴، شماره چهار: ۶۸۹-۶۷۹.

بختیاری ب، لیاقت ع، خلیلی ع، خانجانی م (۱۳۸۸) ارزیابی دو مدل ترکیبی برآورد تبخیر-تقرق مرجع چمن در بازه زمانی ساعتی

- Kisi O and Cimen M (2009) Evapotranspiration modelling using support vector machines. *Hydrological Sciences Journal*, 54(5): 918-928.
- Oudina L, Hervieu F, Michela C, Perrina C, Andre'assiana V, Anctilb F and Loumagnea C (2005) Which potential evapotranspiration input for a lumped rainfall-runoff model? Part 2—Towards a simple and efficient potential evapotranspiration model for rainfall-runoff modeling. *Journal of Hydrology*, 303(1-4): 290-306.
- Quinlan JR (1992) Learning with continuous classes. In: Adams, A. sterling, L. (eds) *Proc. AI'92, 5th Australian Joint Conference on Artificial Intelligence*. World Scientific, Singapore, pp 343-348.
- Shing J and Jang R (1993) ANFIS: Adaptive-Network-Based Fuzzy Inference System. *Man and Cybernetics*, 23(3): 665-685.
- technical paper 96, Soil conservation service. US Department of Agriculture, Washington.
- Ditthakit P and Chinnara Sri C (2012) Estimation of pan coefficient using M5 model tree. *American Journal of Environmental Sciences*, 8 (2): 95-103.
- Douglas EM, Jacobs JM, Sumner DV and Ray RL (2009) A comparison of models for estimating potential evapotranspiration for Florida land cover types. *Journal of Hydrology*, 373: 366-376.
- Eslamian SS, Abedi-Koupai J, Amiri MJ and Gohari SA (2009) Estimation of daily reference evapotranspiration using support vector machines and, artificial neural networks in greenhouse. *Research Journal of Environmental Sciences*, 3 (4): 439-447.
- Hargreaves GH and Samani ZA (1985) Reference crop evapotranspiration from temperature. *Applied Engineering in Agriculture*, 1(2):96-99