

ارزیابی مدل‌های هوشمند داده محور در برآورد تبخیر- تعرق مرجع روزانه در چند اقلیم سواحل جنوبی دریای خزر

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

تاریخ دریافت: ۱۳۹۲/۱۰/۱ تاریخ پذیرش: ۱۳۹۳/۶/۵

چکیده

در بسیاری از مطالعات تعیین میزان دقیق متغیر تبخیر-تعرق از اهمیت ویژه‌ای برخوردار است. مدل‌های هوشمند داده محور دارای توانایی بالایی در مدل‌سازی پدیده‌های پیچیده و غیرخطی هستند. در این مطالعه از چهار روش داده محور شامل شبکه‌های عصبی مصنوعی (ANN)، ماشین‌های بردار پشتیبان (SVM)، مدل درختی M5 و شبکه‌های تطبیقی مبتنی بر سیستم استنتاج فازی (ANFIS) برای مدل‌سازی تبخیر-تعرق مرجع (ET₀) در سه ایستگاه از اقلیم‌های سواحل جنوبی دریای خزر با استفاده از داده‌های بلند مدت روزانه هواشناسی بهره گرفته شد. ۱۱ ترکیب مختلف متغیرهای هواشناسی به عنوان ورودی به مدل‌های داده محور انتخاب گردید و از شاخص‌های استاندارد آماری ضریب تبیین (R²)، ریشه میانگین مربعات خطا (RMSE) و شاخص توافق (DI) برای ارزیابی مدل‌ها استفاده شد. نتایج نشان داد که مدل ANFIS-11 با RMSE بین ۰/۲ تا ۰/۳۸ میلی‌متر بر روز دقیق‌ترین و مدل‌هایی که ورودی آن‌ها فقط سرعت باد است با R² بین ۰/۰۲ تا ۰/۳۹ و RMSE بین ۱/۳۵ تا ۱/۶۸ میلی‌متر بر روز ضعیف‌ترین برآوردها را در اقلیم‌های مورد مطالعه دارند.

واژه‌های کلیدی: تبخیر-تعرق مرجع روزانه، SVM، M5، ANFIS، ANN

مقدمه

در مدیریت و تخصیص بهینه و صحیح منابع آب برای بخش کشاورزی پارامترهای مختلفی دخیل هستند که از جمله می‌توان به تبخیر-تعرق اشاره نمود. روش‌های گوناگونی برای تعیین دقیق این مؤلفه وجود دارد اما این روش‌ها غالباً بسیار هزینه‌بر می‌باشد و به دلیل نیاز به تجهیزات خاص، در همه نقاط قابل استفاده نیست. از این رو در سالیان اخیر تلاش بر این بوده تا روش‌ها و روابط با کاربری آسان و ارزان برای برآورد دقیق تبخیر-تعرق ایجاد گردد. از جمله این روش‌ها می‌توان به روش‌های هوشمند داده محور اشاره نمود. این روش‌ها به دلیل انعطاف‌پذیری بالا می‌توانند بسیاری از پدیده‌های غیر خطی را با دقتی بالا مدل‌سازی نمایند. آدلوی و همکاران (۲۰۱۲) در تحقیقی، پیش‌بینی تبخیر-تعرق مرجع توسط شبکه‌های عصبی خود سازمان دهنده را مورد ارزیابی قرار دادند. در این مطالعه از داده‌های

روزانه ایستگاه هواشناسی دانشگاه ادینبورگ انگلستان به‌عنوان ورودی مدل‌های شبکه عصبی مصنوعی و تجربی استفاده شد. نتایج نشان داد که شبکه عصبی مصنوعی (ANN) با RMSE برابر ۰/۰۰۸ میلی‌متر در روز دارای بهترین عملکرد بوده و مدل تجربی هارگریوز از بین سایر مدل‌های تجربی، بهترین دقت را نشان داد (Adeloye et al., 2012). هو و همکاران (۲۰۱۲) در ایستگاهی در شمال شرقی چین با استفاده از ۵۰ سال داده هواشناسی و معادله پنمن - ماتیت و شبکه‌های عصبی مصنوعی، تبخیر-تعرق مرجع را پیش‌بینی کردند. نتایج نشان داد که شبکه‌های عصبی مصنوعی با دقت بالایی می‌توانند تبخیر-تعرق را تخمین بزنند (Hou et al., 2012). کوبانر (۲۰۱۱) با استفاده از دو ساختار متفاوت سیستم‌های تطبیقی مبتنی بر سیستم استنتاج فازی (ANFIS) و داده‌های روزانه هواشناسی لس‌آنجلس، تبخیر-تعرق مرجع را پیش‌بینی کرد. نتایج حاکی از آن بود که S-ANFIS^۳ دارای عملکرد دقیق‌تری نسبت به G-ANFIS^۴ است (Cobaner, 2011). پولیدو کالو و همکاران (۲۰۰۹) پیش‌بینی نیاز آبی را با استفاده از روش‌های ANFIS و ANN و رگرسیون چند

۱- استادیار بخش مهندسی آب، دانشکده کشاورزی، دانشگاه شهید باهنر کرمان
۲- کارشناسی ارشد مهندسی منابع آب، بخش مهندسی آب، دانشکده کشاورزی، دانشگاه شهید باهنر کرمان

۳- استادیار بخش مهندسی آب، دانشکده کشاورزی، دانشگاه شهید باهنر کرمان
* - نویسنده مسئول: (Email: Drbakhtiari@uk.ac.ir)

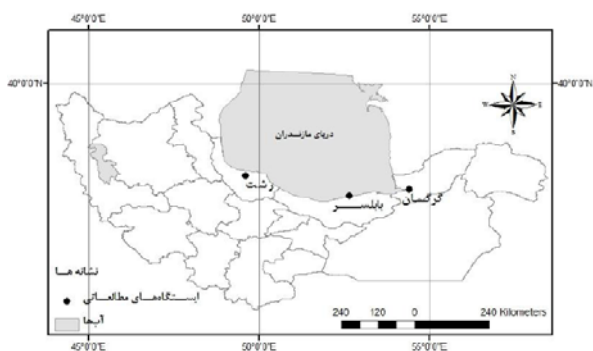
3- Subtractive clustering based Fuzzy Inference System
4- grid partition based fuzzy inference system

ورودی‌های مختلف مورد آزمون قرار گرفته و بهترین روش برای اقلیم‌های مدنظر انتخاب گردد.

مواد و روش‌ها

منطقه مورد مطالعه

در این مطالعه از داده‌های بلند مدت روزانه (۱۹۸۰-۲۰۰۹) رطوبت نسبی، میانگین دما و ساعات آفتابی روزانه سه ایستگاه سینوپتیک گرگان، رشت و بابلسر در حاشیه دریای مازندران استفاده گردید. در جدول ۱ موقعیت جغرافیایی و نرمال‌های اقلیمی ایستگاه‌های مطالعاتی ارائه و در شکل ۱ پراکنش آن‌ها نشان داده شده است. در این مطالعه ۸۰٪ به عنوان داده‌های آموزش و ۲۰٪ دیگر به عنوان داده‌های آزمون مدل‌ها استفاده شد.



شکل ۱- موقعیت جغرافیایی ایستگاه‌های مطالعاتی

روش‌های داده محور

شبکه‌های عصبی مصنوعی

می‌توان یک نرون عصبی و عملکرد آن را توسط مدل‌های ریاضی، مدل‌سازی نمود. شکل ۲ یک نرون عصبی مصنوعی را نشان می‌دهد.

متغیره، بهبود بخشیدند. در این پژوهش داده‌های روزانه هواشناسی و مشخصه‌های محصولات و نیاز آبی بین سال‌های ۱۹۸۸-۱۹۹۱ در استان Codoba در اسپانیا انتخاب شد. پس از مقایسه نیازهای آبی به دست آمده با تقاضاهای آبی مصرف‌کنندگان، روش ANFIS از بین روش‌های دیگر عملکرد بهتری را در پیش‌بینی از خود نشان داد (Pulido-Calov et al., 2009). مرادی و همکاران (۱۳۹۱) با استفاده از معادله استاندارد پنمن-مانتیت فائو، تبخیر-تعرق مرجع در بازه زمانی ساعتی را برای منطقه خشک فریمان با استفاده از داده‌های ۱۰ دقیقه‌ای هواشناسی سال‌های ۲۰۰۸-۲۰۰۹ محاسبه نمودند و عملکرد مدل‌های ASCE و ANFIS را در برابر مدل پنمن-مانتیت فائو مورد آزمون قرار دادند. نتایج نشان‌دهنده آن بود که مدل ANFIS با سه ورودی تابش، رطوبت نسبی و دما دارای بهترین دقت و عملکرد است. احمدزاده قره‌گویز و همکاران (۱۳۸۹) دقت شبکه‌های عصبی مصنوعی و ANFIS در برآورد تبخیر-تعرق گیاه مرجع در برابر مدل‌های تجربی در مناطق بسیار خشک را مورد مطالعه قرار دادند. در این مطالعه از داده‌های روزانه هواشناسی سه ایستگاه همدید اصفهان، کرمان و یزد استفاده گردید. نتایج نشان‌دهنده برتری مدل ANFIS با ورودی‌های دمای حداکثر، سرعت باد و تابش نسبت به سایر مدل‌ها بود. اسلامیان و همکاران (۲۰۰۹) نیز توانایی روش SVM و ANN را در مدل‌سازی تبخیر-تعرق در محیط گلخانه بررسی کردند.

در این مدل‌سازی از داده‌های هواشناسی در درون گلخانه دانشگاه صنعتی اصفهان و داده‌های مربوط به تبخیر-تعرق مرجع از داده‌های تست تبخیر استفاده شد. نتایج نشان داد هر دو روش مورد مقایسه با عملکرد خوبی مدل‌سازی را انجام دادند اما SVM دارای عملکرد بالاتری بود (Eslamian et al., 2009).

در یک جمع‌بندی مشخص گردید که روش‌های هوشمند داده محور پرکاربرد SVM، M5، ANFIS و ANN در برآورد تبخیر-تعرق در اقلیم‌های مدیترانه‌ای، مرطوب و فرا مرطوب ارزیابی نگردیده‌اند. از این رو در این مطالعه تلاش گردید تا این روش با

جدول ۱- وضعیت جغرافیایی و نرمال‌های هواشناسی ایستگاه‌های مطالعاتی

ردیف	ایستگاه	طول جغرافیایی دقیقه-درجه (شرقی)	عرض جغرافیایی دقیقه-درجه (شمالی)	ارتفاع از سطح دریا (متر)	میانگین دمای سالانه (سانتی-گراد)	میانگین بارش سالانه (میلی-متر)	نوع اقلیم	متوسط تبخیر-تعرق پتانسیل سالانه (میلی‌متر)
۱	گرگان	۵۴ ۲۴	۳۶ ۵۴	۱۳/۳	۱۷/۸	۶۰۱	مدیترانه‌ای معتدل	۸۸۵/۹۳
۲	بابلسر	۵۲ ۲۹	۳۶ ۴۳	-۲۱	۱۷/۰	۸۹۴/۴	مرطوب معتدل	۷۹۷/۴۷
۳	رشت	۴۹ ۳۷	۳۷ ۱۹	-۸/۶	۱۶/۳۶	۱۳۵۹	فرا مرطوب معتدل	۷۲۹/۴

هر نرون می‌تواند تابع فشرده‌سازی مستقلی داشته باشد (منهاج، ۱۳۷۹). در این پژوهش از آرایش $\{۱۰، ۱۰، ۵\}$ با تابع انتقال تانژانت سیگموئید استفاده گردید.

ماشین بردار پشتیبان

ماشین بردار پشتیبان (SVM^۴)، یک روش طبقه‌بندی گرفته شده از تئوری آموزش آماری است و اولین بار توسط واپنیک^۵ معرفی گردید. SVMها می‌توانند برای هر دو منظور مسائل طبقه‌بندی و رگرسیون مورد استفاده قرار گیرد. می‌توان ماشین بردار پشتیبان را شبیه به شبکه‌های عصبی مصنوعی، مانند یک شبکه با دو لایه که وزن‌ها در اولین لایه غیرخطی و در دومین لایه خطی هستند ارائه نمود. تخمین رگرسیون با SVR^۶، تخمین تابعی مطابق دسته داده‌های گرفته شده از $\{(x_i, y_i)\}_n$ است که بردارهای ورودی، y_i مقادیر خروجی و n تعداد کل دسته داده‌ها می‌باشد. تابع رگرسیون خطی استفاده شده تابع شماره ۲ است (Cimen and Kisi., 2009).

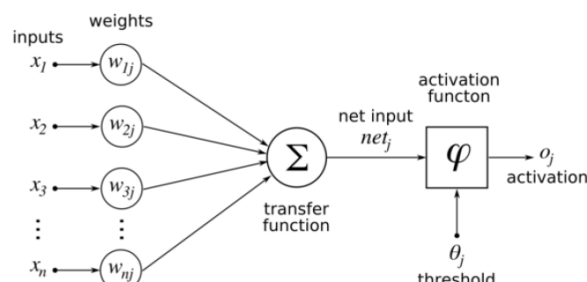
$$f(x) = \omega \cdot \phi(x) + b \quad (۲)$$

که $\phi(x)$ تابع غیرخطی که توسط آن x به فضای ویژه نگاشت می‌گردد، و b و ω به ترتیب نشان دهنده بردار وزن و ضریبی است که باید از داده‌ها برآورد شود. ضرایب b و ω توسط کوچک‌سازی جمع ریسک‌های تجربی (اولین مؤلفه تابع ۳) و مؤلفه مرکب (دومین مؤلفه تابع ۳) تخمین زده می‌شود.

$$R = c \sum_{i=1}^n L_\varepsilon(f(x_i), y_i) + \frac{1}{2} \|\omega\|^2 \quad (۳)$$

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

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، $\|\omega\|^2$ مؤلفه تنظیم است که نشان دهنده قاعده اقلیدسی است و L_ε به عنوان ε -غیرحساس تابع تلفات نامیده می‌شود که میزان ریسک تجربی را اندازه‌گیری می‌کند. L_ε دارای این مزیت است که برای توصیف بردار رگرسیون ضریب نگاشت نیازی به داده‌های ورودی ندارد. به عنوان مثال معادله ۳ 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 اغلب به عنوان نسبتی از مقادیر خروجی y_i تعریف می‌گردد. بنابراین، تابع رگرسیون غیره با استفاده از معادلات ۴ و ۵ به‌دست می‌آید (Cimen and Kisi.,



شکل ۲- ساختار یک نرون مصنوعی (منهاج، ۱۳۷۹)

سیگنال‌های ورودی X_1 تا X_n معادل سیگنال‌های عصبی ورودی و وزن‌های W_{i1} تا W_{in} معادل مقادیر اتصالات سیناپسی ورودی‌های نرون می‌باشند که جمعاً ورودی نرون را تشکیل داده است.

تابع جمع‌کننده $\sum_{j=1}^n x_j w_{ij}$ تمامی عملیات هسته سلول را انجام می‌دهد. خروجی نرون توسط معادله (۱) مشخص می‌شود.

$$y_i = \text{ActivationFunction} \left(\sum_{j=1}^n x_j w_{ij} \right) \quad (۱)$$

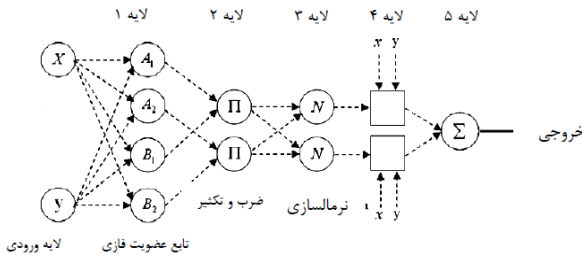
در حقیقت در طراحی یک نرون مصنوعی فقط کافی است وزن‌های W_{ij} مشخص شود تا شبکه عصبی بتواند خروجی مورد نظر را از ورودی خاص تولید کند. روش‌های مختلف یادگیری وجود دارد که می‌تواند بر اساس زوج مرتب‌های (خروجی، ورودی) مقدار وزن‌ها را به‌دست آورد. یک ساختار شبکه‌ای از تعدادی عناصر مرتبط به هم به نام نرون که هر نرون دارای ورودی‌ها و خروجی‌هایی است و یک عمل نسبتاً ساده و محلی^۱ را انجام می‌دهد. شبکه‌های عصبی عموماً عملکرد خود را طی یک پروسه یادگیری فرا می‌گیرد.

شبکه‌هایی که از یک نرون ساخته شده‌اند دارای محدودیت‌هایی می‌باشند برای حل این مشکل شبکه‌های دیگری پیشنهاد شد که در اینجا از یکی از آنها که دارای بیش‌ترین کاربرد است استفاده شد. مدل عمومی شبکه‌های پرسپترون^۲، شبکه جلو رونده^۲ با روال تعلیم انتشار به‌عقب^۳ است. شبکه‌های جلو رونده شبکه‌هایی هستند که ورودی‌های لایه اول نرون‌های آن به لایه‌های بعدی متصل بوده و در هر سطح این مسئله صادق بوده تا به لایه خروجی برسد. روال انتشار به‌عقب بدین معنی است که پس از مشخص شدن خروجی شبکه، ابتدا وزن‌های لایه آخر تصحیح شده و بعد به ترتیب اوزان لایه‌های قبلی تصحیح می‌شوند. نرون‌های لایه ورودی عملی را انجام نمی‌دهند و اوزان آنها ثابت و برابر یک است. انتشار عملگر رو به جلو است. تمامی نرون‌ها به‌غیر از لایه ورودی جمع‌کننده بوده و

4- Support Vector Machines
5- Vapnik
6- Support Vector Regression

1- Local
2- Feed Forward
3- Back Propagation Training

دستورالعمل مدل فازی Sugeno، یک مجموعه قاعده با دو قانون اگر-آنگاه فازی، به صورت ذیل می‌توان بررسی کرد (Shing and Jang., 1993):



شکل ۳- معماری شبکه‌های تطبیقی مبتنی بر سیستم استنتاج فازی (ANFIS)

قاعده ۱: IF X is A₁ and Y is B₁, then f₁=p₁X + q₁Y + r₁
 قاعده ۲: IF X is A₂ and Y is B₂, then f₂=p₂X + q₂Y + r₂
 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 ۷ 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} هرگونه تابع عضویت فازی را قبول می‌کنند.
 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 ۸ اندازه‌گیری می‌کند.

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

لایه ۳: آمین گره در این لایه نسبت وزن آمین قاعده به مجموع وزن‌های قواعد (رابطه ۹) را محاسبه می‌کند که W_i مقدار نرمال شده وزن‌ها است.

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

لایه ۴: گره i در این لایه سهم آمین قاعده را در ایجاد خروجی مدل محاسبه می‌کند (رابطه ۱۰). پارامترهای این لایه را پارامترهای برآیند نیز می‌گویند.

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

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

(2009).

$$f(x, \alpha, \alpha^*) = \sum_{i=1}^N (\alpha_i - \alpha_i^*) k(x_i, x) + b \quad (4)$$

با شرط $\alpha_i, \alpha_i^* \geq 0$ و $\alpha_i, \alpha_i^* = 0$ ، $i = 1, \dots, N$ ، رابطه ۵ تابع کرنل $k(x_i, x)$ ضرب نقطه‌ای در فضای ویژه D-بعدی را توصیف می‌کند.

$$k(x_i, x) = \sum_{j=1}^D \phi_j(x) \phi_j(y) \quad (5)$$

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 بوده و دارای یک فرم تحلیلی شناخته شده می‌باشد. ضرایب α_i و α_i^* با استفاده از حداکثر سازی معادله ۶ به دست می‌آیند.

$$R(\alpha^* \alpha) = -\text{sum}_{i=1}^N (\alpha_i^* + \alpha_i) + \sum y_i (\alpha_i^* - \alpha_i) - \frac{1}{2} \sum_{i,j=1}^N (\alpha_i^* + \alpha_i) \times (\alpha_j^* + \alpha_j) k(x_i, x_j) \quad (6)$$

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

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). در این مطالعه پس از آزمون کرنل‌های مختلف، کرنل RBF^۱ به عنوان تابع نگاشت استفاده شد.

شبکه‌های تطبیقی مبتنی بر سیستم استنتاج فازی

اگرچه شبکه‌های عصبی مصنوعی 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^۲ با ترکیب قواعد پایه سیستم فازی با توانایی یادگیری شبکه‌های عصبی توسعه یافته است. مدل ANFIS ارائه شده‌ی 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 ۳ نشان داده شده است که در آن دایره نشان دهنده گره‌های ثابت و مربع نشان دهنده 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 به دلیل تفسیرپذیری، قابلیت محاسباتی بالا و ساخته شده توسط روش‌های بهینه و تطبیق‌پذیر، در دامنه وسیعی به کار برده می‌شود. برای اولین

1- Radial Base Function

2- Adaptive Networks based Fuzzy Inference System

(معادله ۱۱)

$$O_i^s = \text{overall output} = \sum_i \bar{w}_i z_i = \frac{\sum_i w_i z_i}{\sum_i w_i} \quad (11)$$

مدل درختی M5

مدل M5 یکی از زیرشاخه‌های ماشین‌های آموزش مرتبط با ساخت یا بازبینی مدل‌ها در تعداد زیاد نمونه‌های آزمایشی است که به تخمین مسائل رایج مثل داده‌های گم شده و نویزها می‌پردازد (Quinlan., 1992). این مدل اولین بار توسط (Quinlan 1992) ارائه گردید. در شکل ۴ اجزای تشکیل دهنده مدل درختی M5 ارائه شده است.

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 صورت معادله ۱۲ نوشته می‌شود.

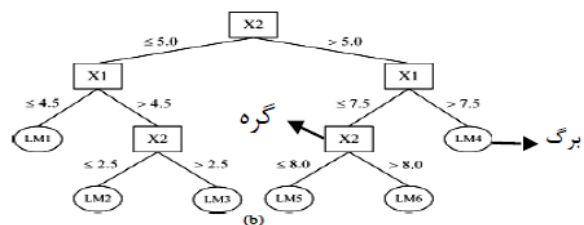
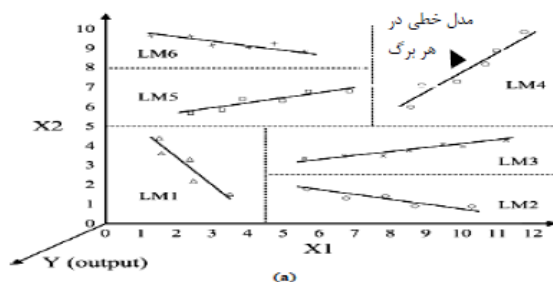
$$SDR = sd(T) - \sum_i \left| \frac{T_i}{T} \right| sd(T_i) \quad (12)$$

که T نشان دهنده یک دسته نمونه‌ها است که به برگ می‌رسد، T_i نشان دهنده زیرمجموعه‌ای از نمونه‌ها است که دارای آامین خروجی از مجموعه بالقوه است و sd نشان دهنده 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 معادله ۱۳ برآورد می‌شود.

$$\hat{f}(S) = \frac{n_i \times PV(S_i) + k \times M(S)}{n_i + k} \quad (13)$$

در اینجا K ثابت هموارسازی است که مقدار آن ۱۵ فرض می‌شود (Quinlan., 1992).



شکل ۴- نمونه الگوریتم مدل درختی M5 با شش مدل رگرسیون خطی (Ditthakit and Chinnarasri, 2012)

$$DI = 1 - \left[\frac{\sum_{i=1}^N (ET_{PM-i} - ET_{M-i})^2}{\sum_{i=1}^N ((ET_{M-i} - \overline{ET}_{PM-i}) + (ET_{M-i} - \overline{ET}_{M-i}))} \right] \quad (17)$$

نتایج و بحث

در این مطالعه از ۱۱ ترکیب مختلف متغیرهای هواشناسی به عنوان ورودی‌های مدل‌های داده محور استفاده شد که در جدول ۲ ارائه گردیده است.

جدول ۲- ترکیب‌های داده‌های هواشناسی ورودی به مدل‌ها

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)

اقلیم معتدل مدیترانه‌ای گرگان

در بررسی مدل‌های داده محور با ورودی‌ها متنوع هواشناسی، مدل ANFIS-11 با چهار مؤلفه هواشناسی ورودی، مناسب‌ترین برآوردها را با کم‌ترین خطا (RMSE برابر ۰/۳۸ میلی‌متر بر روز و ضریب تبیین برابر ۰/۹۶ داشته است. پس از مدل ANFIS-11، مدل‌های ANFIS-10 با سه متغیر ورودی و ANFIS-7 با دو متغیر ورودی و ANFIS-1 با یک متغیر ورودی بهترین پاسخ‌ها را در ترکیب‌های مختلف نشان دادند (جدول ۳). در میان این مدل‌ها، مدل ANN-4 با متغیر ورودی سرعت باد و RMSE برابر ۱/۶۸ میلی‌متر بر روز و R^2 برابر ۰/۲۴، نامطلوب‌ترین جواب‌ها را در مقایسه با برآوردهای PMF-56 تولید کرد. در شکل ۵ عملکرد مدل‌های برتر نسبت به معادله PMF-56 نشان داده شده است.

در اقلیم معتدل مدیترانه‌ای گرگان (با توجه به جدول ۳) دما تأثیرگذارترین عامل هواشناسی و پس از آن به ترتیب متغیرهای سرعت باد، تابش و در آخر رطوبت نسبی در تخمین تبخیر-تعرق پتانسیل موثرترین می‌باشند.

معادله استاندارد تبخیر-تعرق پتانسیل

در این مطالعه از خروجی معادله پنمن-مانتیث فائو ۵۶ (PMF-56) به عنوان مقدار تبخیر-تعرق معیار در بررسی عملکرد مدل‌های داده محور استفاده شد. سازمان خوار و بار جهانی، روش پنمن - مانتیث را به دلیل اتکا به پارامترهای فیزیکی، بیش‌تر از سایر روش‌ها و تطابق بیش‌تر و بهتر با داده‌های لایسیمتر بیش از سایر روش‌های تخمین تبخیر-تعرق توصیه نموده است (خیرابی و همکاران، ۱۳۷۶). روش پنمن - مانتیث در معادله (۱۷) نشان داده شده است (Allen et al., 1998).

$$T_0 = \frac{0.408 \Delta(R_n - G) + \gamma \frac{900}{T + 273} U_2 (e_s - e_d)}{\Delta + \gamma(1 + 0.34U_2)} \quad (14)$$

که در آن ET_0 تبخیر-تعرق گیاه مرجع استاندارد ($mm \ day^{-1}$)، R_n تشعشع خالص از سطح گیاه ($MJ \ m^{-2} \ day^{-1}$)، G جریان گرمایی خاک ($MJ \ m^{-2} \ day^{-1}$)، T میانگین درجه حرارت ($^{\circ}C$)، U_2 متوسط سرعت باد در طول شبانه‌روز در ارتفاع ۲ متری ($m \ s^{-1}$)، $(e_s - e_d)$ کمبود فشار بخار (kPa)، Δ شیب منحنی تغییرات فشار بخار ($kPa \ ^{\circ}C^{-1}$)، γ ثابت سایکرومتری ($kPa \ ^{\circ}C^{-1}$)، عدد ۹۰۰ فاکتور یا ضریب تعدیل است.

شاخص‌های آماری استاندارد

جهت مقایسه عملکرد مدل‌های مختلف داده محور در تخمین تبخیر-تعرق از شاخص‌های استاندارد آماری ضریب تبیین (R^2)، ریشه میانگین مربعات خطا ($RMSE$) و شاخص توافق (DI) که به ترتیب در روابط ۱۵، ۱۶ و ۱۷ به آن‌ها اشاره شده است، استفاده گردید (رضایی و میدی، ۱۳۸۴ و Willmott, 1981). در این معادلات، ET_{PM} آمین برآورد مدل PMF-56 (میلی‌متر بر روز)، \overline{ET}_{PM} میانگین برآوردهای مدل PMF-56 (میلی‌متر بر روز)، ET_{M-i} آمین برآورد مدل مورد آزمون (میلی‌متر بر روز) و \overline{ET}_M میانگین برآوردهای مدل مورد آزمون (میلی‌متر بر روز) می‌باشد.

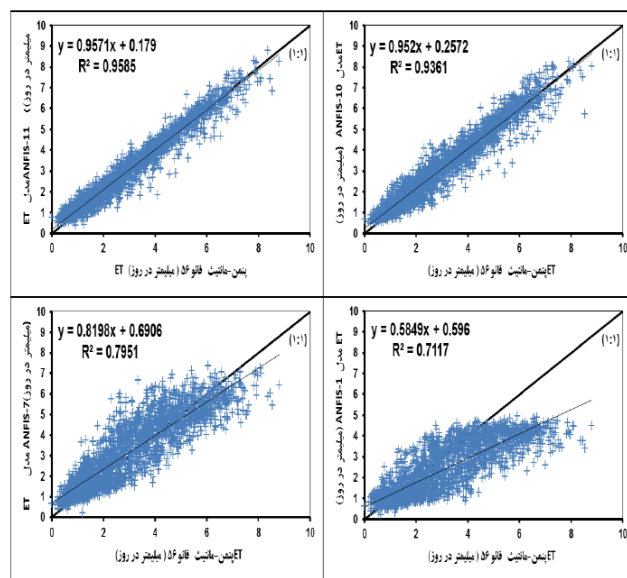
$$R^2 = \frac{(\sum_{i=1}^N (ET_{PM-i} - \overline{ET}_{PM})(ET_{M-i} - \overline{ET}_M))^2}{(\sum_{i=1}^N (ET_{PM-i} - \overline{ET}_{PM})^2)(\sum_{i=1}^N (ET_{M-i} - \overline{ET}_M)^2)} \quad (15)$$

$$RMSE = \frac{1}{n} \sqrt{\sum_{i=1}^N (ET_{PM-i} - ET_{M-i})^2} \quad (16)$$

- 1- Penman-Montieth Fao 56
- 2- Root Mean Square of Errors
- 3- Index of Agreement

جدول ۳ - عملکرد مدل‌های داده محور در اقلیم معتدل مدیترانه‌ای گرگان

رتبه	مدل	شماره	R ²	RMSE	DI	رتبه	مدل	شماره	R ²	RMSE	DI
۱	ANFIS	۱۱	۰/۹۶	۰/۹۶	۰/۹۹	۲۳	ANN	۵	۰/۷۵	۱/۰۵	۰/۹۰
۲	SVM	۱۱	۰/۹۶	۰/۹۶	۰/۹۹	۲۴	M5	۶	۰/۸۱	۱/۰۵	۰/۹۰
۳	ANN	۱۱	۰/۹۵	۰/۹۵	۰/۹۹	۲۵	ANN	۶	۰/۸۴	۱/۰۶	۰/۸۹
۴	M5	۱۱	۰/۹۴	۰/۹۴	۰/۹۹	۲۶	ANFIS	۵	۰/۸۰	۱/۱۰	۰/۸۸
۵	ANFIS	۱۰	۰/۹۴	۰/۹۴	۰/۹۸	۲۷	M5	۵	۰/۷۴	۱/۱۶	۰/۸۷
۶	SVM	۱۰	۰/۹۴	۰/۹۴	۰/۹۸	۲۸	SVM	۵	۰/۸۰	۱/۱۶	۰/۸۶
۷	ANN	۱۰	۰/۹۱	۰/۹۱	۰/۹۷	۲۹	ANFTS	۱	۰/۷۱	۱/۲۲	۰/۸۴
۸	ANFIS	۹	۰/۹۱	۰/۹۱	۰/۹۷	۳۰	M5	۱	۰/۷۱	۱/۲۲	۰/۸۴
۹	M5	۱۰	۰/۹۱	۰/۹۱	۰/۹۷	۳۱	ANN	۳	۰/۶۶	۱/۲۵	۰/۸۲
۱۰	SVM	۹	۰/۹۰	۰/۹۰	۰/۹۷	۳۲	ANFIS	۳	۰/۶۶	۱/۲۷	۰/۸۳
۱۱	ANN	۹	۰/۸۹	۰/۸۹	۰/۹۷	۳۳	SVM	۱	۰/۷۱	۱/۲۷	۰/۸۴
۱۲	M5	۹	۰/۸۷	۰/۸۷	۰/۹۶	۳۴	M5	۳	۰/۶۲	۱/۳۰	۰/۸۳
۱۳	ANFIS	۷	۰/۸۰	۰/۸۰	۰/۹۴	۳۵	SVM	۳	۰/۶۶	۱/۳۰	۰/۸۲
۱۴	SVM	۷	۰/۸۰	۰/۸۰	۰/۹۳	۳۶	ANN	۱	۰/۷۰	۱/۳۱	۰/۸۱
۱۵	ANN	۸	۰/۸۳	۰/۸۳	۰/۹۲	۳۷	SVM	۲	۰/۴۷	۱/۵۸	۰/۶۸
۱۶	ANN	۷	۰/۷۸	۰/۷۸	۰/۹۳	۳۸	M5	۲	۰/۴۸	۱/۵۹	۰/۶۳
۱۷	M5	۷	۰/۷۵	۰/۷۵	۰/۹۳	۳۹	ANFIS	۲	۰/۴۷	۱/۶۱	۰/۶۲
۱۸	SVM	۸	۰/۸۴	۰/۸۴	۰/۹۲	۴۰	SVM	۴	۰/۲۶	۱/۶۵	۰/۵۶
۱۹	ANFIS	۸	۰/۸۵	۰/۸۵	۰/۹۱	۴۱	ANFIS	۴	۰/۲۴	۱/۶۶	۰/۵۰
۲۰	ANFIS	۶	۰/۸۵	۰/۸۵	۰/۹۱	۴۲	ANN	۲	۰/۴۶	۱/۶۶	۰/۵۷
۲۱	SVM	۶	۰/۸۵	۰/۸۵	۰/۹۰	۴۳	M5	۴	۰/۲۱	۱/۶۸	۰/۴۹
۲۲	M5	۸	۰/۸۰	۰/۸۰	۰/۹۰	۴۴	ANN	۴	۰/۲۴	۱/۶۸	۰/۶۹



شکل ۵- مقایسه برآورد ET مدل‌های داده محور برتر با برآورد PMF-56 در اقلیم معتدل مدیترانه‌ای گرگان

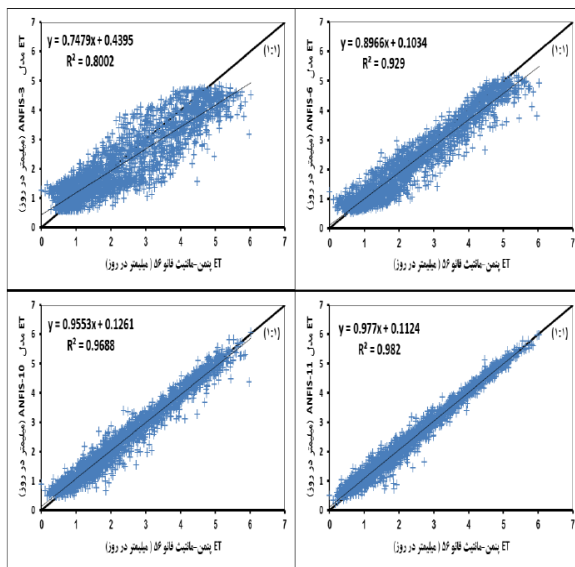
اقلیم معتدل مرطوب بابلسر

با توجه به نتایج ارائه شده در جدول ۴ مدل ANFIS-11 با چهار ورودی به‌عنوان مدل برتر از لحاظ عملکرد و دقت برای اقلیم معتدل مرطوب بابلسر برگزیده شد. میزان RMSE محاسبه‌شده برای این

مدل ۰/۲ میلی‌متر بر روز با ضریب تبیین و شاخص توافق به ترتیب ۰/۹۸ و ۱ است. در سایر ترکیب‌ها، مدل ANFIS-10 با سه متغیر ورودی، مدل ANFIS-6 با دو متغیر ورودی و مدل ANFIS-3 با یک متغیر ورودی بهترین عملکردها را به نسبت تعداد ورودی‌ها

خورشیدی است و پس از آن به ترتیب میانگین دما، میانگین رطوبت نسبی و میانگین سرعت باد هستند که این موضوع از نتایج موجود در جدول ۴ قابل استنباط است. در شکل ۶ نتایج مدل‌های برتر نسبت به نتایج معادله PMF-56 نشان داده شده است.

داشته‌اند. نامناسب‌ترین روش داده محور مدل MSP با ورودی متوسط سرعت باد روزانه است که دارای RMSE و R^2 به ترتیب ۱/۳۷ میلی‌متر بر روز و ۰/۱ است. نتایج به دست آمده نشان می‌دهد در این اقلیم مهم‌ترین عامل در تخمین دقیق‌تر، متغیر تابش خالص



شکل ۶- مقایسه برآورد ET مدل‌های داده محور برتر با برآورد PMF-56 در اقلیم معتدل مرطوب بابل

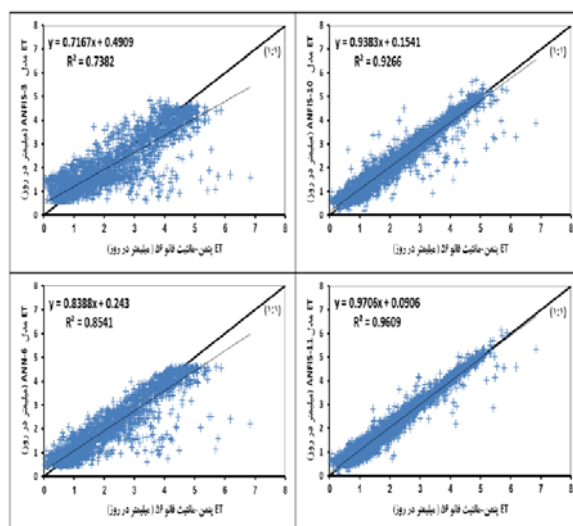
جدول ۴ - عملکرد مدل‌های داده محور در اقلیم معتدل مرطوب بابل

رتبه	مدل	شماره	R^2	RMSE	DI	رتبه	مدل	شماره	R^2	RMSE	DI
۱	ANFIS	۱۱	۰/۹۹	۰/۲۰	۱/۰۰	۲۳	ANN	۵	۰/۹۰	۰/۶۴	۰/۹۴
۲	SVM	۱۱	۰/۹۹	۰/۲۱	۰/۹۹	۲۴	ANFIS	۷	۰/۹۰	۰/۶۴	۰/۹۴
۳	ANN	۱۱	۰/۹۹	۰/۲۳	۰/۹۹	۲۵	SVM	۷	۰/۹۰	۰/۶۵	۰/۹۳
۴	M5	۱۱	۰/۹۹	۰/۲۵	۰/۹۹	۲۶	ANN	۷	۰/۸۹	۰/۶۶	۰/۹۳
۵	ANFIS	۱۰	۰/۹۸	۰/۲۶	۰/۹۹	۲۷	M5	۵	۰/۸۹	۰/۶۷	۰/۹۴
۶	SVM	۱۰	۰/۹۸	۰/۲۷	۰/۹۹	۲۸	ANFIS	۳	۰/۸۹	۰/۶۷	۰/۹۳
۷	ANN	۱۰	۰/۹۸	۰/۲۸	۰/۹۹	۲۹	SVM	۳	۰/۸۹	۰/۶۷	۰/۹۴
۸	M5	۱۰	۰/۹۸	۰/۳۱	۰/۹۹	۳۰	ANN	۳	۰/۸۹	۰/۶۸	۰/۹۳
۹	ANFIS	۸	۰/۹۷	۰/۳۹	۰/۹۸	۳۱	M5	۳	۰/۸۸	۰/۷۰	۰/۹۳
۱۰	SVM	۸	۰/۹۷	۰/۴۰	۰/۹۸	۳۲	M5	۷	۰/۸۷	۰/۷۰	۰/۹۳
۱۱	ANFIS	۶	۰/۹۶	۰/۴۱	۰/۹۸	۳۳	ANFIS	۱	۰/۸۷	۰/۷۵	۰/۹۱
۱۲	SVM	۶	۰/۹۶	۰/۴۱	۰/۹۸	۳۴	M5	۱	۰/۸۶	۰/۷۵	۰/۹۱
۱۳	ANN	۸	۰/۹۶	۰/۴۳	۰/۹۷	۳۵	ANN	۱	۰/۸۷	۰/۷۷	۰/۹۱
۱۴	ANN	۶	۰/۹۶	۰/۴۴	۰/۹۸	۳۶	SVM	۱	۰/۸۶	۰/۷۸	۰/۹۰
۱۵	M5	۶	۰/۹۶	۰/۴۵	۰/۹۷	۳۷	ANFIS	۲	۰/۶۱	۱/۱۵	۰/۶۹
۱۶	M5	۸	۰/۹۵	۰/۴۶	۰/۹۷	۳۸	ANN	۲	۰/۶۲	۱/۱۵	۰/۶۸
۱۷	ANFIS	۹	۰/۹۴	۰/۵۰	۰/۹۷	۳۹	M5	۲	۰/۶۱	۱/۱۵	۰/۶۹
۱۸	ANN	۹	۰/۹۴	۰/۵۱	۰/۹۶	۴۰	SVM	۲	۰/۵۹	۱/۱۷	۰/۶۹
۱۹	SVM	۹	۰/۹۴	۰/۵۲	۰/۹۷	۴۱	SVM	۴	۰/۳۹	۱/۳۵	۰/۳۸
۲۰	M5	۹	۰/۹۱	۰/۶۰	۰/۹۵	۴۲	ANFIS	۴	۰/۳۸	۱/۳۵	۰/۳۱
۲۱	ANFIS	۵	۰/۹۱	۰/۶۱	۰/۹۵	۴۳	ANN	۴	۰/۳۴	۱/۳۶	۰/۳۰
۲۲	SVM	۵	۰/۹۰	۰/۶۲	۰/۹۵	۴۴	M5	۴	۰/۳۲	۱/۳۷	۰/۲۹

اقلیم معتدل فرا مرطوب رشت

بر اساس نتایج به دست آمده در این تحقیق (جدول ۵ و شکل ۷)، مدل ANFIS-11 با RMSE برابر ۰/۲۷ میلی-متر بر روز و R^2 برابر ۰/۹۶ به عنوان بهترین مدل در برآورد تبخیر-تعرق مرجع در اقلیم

معتدل فرامرطوب رشت شناخته شد و پس از آن مدل ANFIS-10 با سه متغیر ورودی، مدل ANN-6 با دو متغیر ورودی و مدل ANFIS-3 با یک متغیر ورودی بهترین نتایج را در ترکیب‌های خود نشان دادند.



شکل ۷- مقایسه برآورد ET مدل‌های داده محور برتر با برآورد PMF-56 در اقلیم معتدل فرامرطوب رشت

جدول ۵ - عملکرد مدل‌های داده محور در اقلیم معتدل فرامرطوب رشت

رتبه	مدل	شماره	R^2	RMSE	DI	رتبه	مدل	شماره	R^2	RMSE	DI
۱	ANFIS	۱۱	۰/۹۶	۰/۲۷	۰/۹۹	۲۳	SVM	۵	۰/۸۱	۰/۶۱	۰/۹۴
۲	SVM	۱۱	۰/۹۵	۰/۳۰	۰/۹۹	۲۴	M5	۵	۰/۷۸	۰/۶۵	۰/۹۳
۳	ANN	۱۱	۰/۹۵	۰/۳۱	۰/۹۹	۲۵	ANFIS	۷	۰/۷۴	۰/۷۱	۰/۹۲
۴	M5	۱۱	۰/۹۵	۰/۳۲	۰/۹۹	۲۶	ANFIS	۳	۰/۷۴	۰/۷۱	۰/۹۱
۵	ANFIS	۱۰	۰/۶۳	۰/۳۷	۰/۹۸	۲۷	ANN	۳	۰/۷۴	۰/۷۱	۰/۹۱
۶	ANN	۱۰	۰/۶۲	۰/۳۹	۰/۹۸	۲۸	ANN	۷	۰/۷۳	۰/۷۱	۰/۹۲
۷	M5	۱۰	۰/۹۱	۰/۴۱	۰/۹۸	۲۹	SVM	۳	۰/۷۴	۰/۷۴	۰/۹۰
۸	SVM	۱۰	۰/۹۲	۰/۴۲	۰/۹۸	۳۰	M5	۳	۰/۷۲	۰/۷۴	۰/۹۱
۹	ANN	۸	۰/۹۱	۰/۴۲	۰/۹۷	۳۱	SVM	۷	۰/۷۱	۰/۷۵	۰/۹۰
۱۰	ANFIS	۸	۰/۹۱	۰/۴۲	۰/۹۷	۳۲	M5	۷	۰/۷۰	۰/۷۶	۰/۹۰
۱۱	SVM	۸	۰/۹۰	۰/۴۵	۰/۹۷	۳۳	ANN	۱	۰/۶۹	۰/۷۸	۰/۸۹
۱۲	M5	۸	۰/۸۹	۰/۴۷	۰/۹۷	۳۴	ANFIS	۱	۰/۶۸	۰/۷۸	۰/۸۹
۱۳	ANFIS	۹	۰/۸۶	۰/۵۱	۰/۹۶	۳۵	M5	۱	۰/۶۸	۰/۷۹	۰/۸۹
۱۴	ANN	۹	۰/۸۶	۰/۵۲	۰/۹۶	۳۶	SVM	۱	۰/۶۷	۰/۸۰	۰/۸۹
۱۵	ANN	۶	۰/۸۵	۰/۵۴	۰/۹۶	۳۷	ANN	۲	۰/۵۳	۰/۹۵	۰/۸۰
۱۶	ANFIS	۶	۰/۸۵	۰/۵۴	۰/۹۶	۳۸	ANFIS	۲	۰/۵۲	۰/۹۶	۰/۸۰
۱۷	M5	۶	۰/۸۴	۰/۵۷	۰/۹۵	۳۹	M5	۲	۰/۵۲	۰/۹۶	۰/۸۰
۱۸	ANFIS	۵	۰/۸۳	۰/۵۷	۰/۹۵	۴۰	SVM	۲	۰/۴۹	۱/۰۰	۰/۷۵
۱۹	SVM	۶	۰/۸۴	۰/۵۸	۰/۹۵	۴۱	ANN	۴	۰/۰۵	۱/۳۵	۰/۳۳
۲۰	ANN	۵	۰/۸۲	۰/۵۹	۰/۹۵	۴۲	ANFIS	۴	۰/۰۴	۱/۳۶	۰/۱۱
۲۱	SVM	۹	۰/۸۳	۰/۵۹	۰/۹۴	۴۳	M5	۴	۰/۰۳	۱/۳۷	۰/۱۸
۲۲	M5	۹	۰/۸۱	۰/۶۰	۰/۹۵	۴۴	SVM	۴	۰/۰۲	۱/۳۷	۰/۰۴

رضایی، ع و میدی، ع. ۱۳۸۴. آمار و احتمالات. انتشارات دانشگاه صنعتی اصفهان، چاپ اول، ۵۹۰ صفحه.

مرادی، ح، تمناء، انصاری، ح و نادریان فرم. ۱۳۹۱. سیستم‌های استنتاج فازی جهت برآورد تبخیر تعرق مرجع ساعتی (مطالعه موردی - منطقه فریمان). مجله پژوهش‌های حفاظت آب و خاک، جلد نوزدهم، شماره اول، صفحات ۱۶۸-۱۵۳.

منهاج، م. ب. ۱۳۷۹. مبانی شبکه‌های عصبی. دانشگاه صنعتی امیر کبیر، معرفی شبکه‌های عصبی مصنوعی، ۷۱۵ صفحه.

Adeloye, A.J., Rustum, R., Kariyama, I.D. 2012. Neural computing modeling of the reference crop evapotranspiration. *Environmental Modelling and software*, 29: 61-73.

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, 356 pages.

Cimen, M., Kisi, O. 2009. Comparison of two different data-driven techniques in modeling lake level fluctuations in Turkey. *Journal of Hydrology*, 378: 253-262.

Cobaner, M. 2011. Evapotranspiration estimation by two different neuro-fuzzy inference systems. *Journal of Hydrology*, 398: 292-302.

Ditthakit, P., Chinnarasri, C. 2012. Estimation of Pan Coefficient using M5 Model Tree. *American Journal of Environmental sciences* 8 (2): 95-103.

Eslamian, S.S., Abedi-Koupai, J., Amiri, M.J., Gohari, S.A. 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.

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 Environments*, 82: 81-90.

Pulido-Calvo, I., Gutierrez-Estrada, J.C. 2009. Improved irrigation water demand forecasting using a soft-computing hybrid model. *Biosystems engineering*, 102: 202-218.

مدل SVM-4 با ورودی متغیر هواشناسی میانگین سرعت باد روزانه و RMSE برابر ۱/۳۷ میلی‌متر در روز کم‌ترین دقت را در برآورد ET_o نشان داد. در میان متغیرهای هواشناسی، متغیر هواشناسی تابش خالص خورشیدی، مؤثرترین ورودی در بالا بردن دقت عملکرد مدل‌ها است و پس از آن به ترتیب تأثیر، میانگین دمای روزانه، رطوبت نسبی و سرعت باد است.

نتیجه‌گیری

مقدار دقیق تبخیر- تعرق پتانسیل در هر منطقه به سهولت قابل اندازه‌گیری نیست از این رو روش‌های متنوعی برای برآورد نزدیک به واقعیت این پارامتر تا به امروز پیشنهاد گردیده است. در این مطالعه به بررسی عملکرد مدل‌های داده محور مختلف در برآورد تبخیر- تعرق پتانسیل در اقلیم‌های موجود در ساحل دریای خزر در ایران پرداخته شد. در این مطالعات بهترین عملکردها مربوط به روش داده محور ANFIS با حداکثر داده‌های هواشناسی می‌باشد که بهترین برآورد آن مربوط به ایستگاه بابلسر با RMSE ۰/۲ میلی‌متر در روز بود. سایر مدل‌ها با توجه به ترکیب متغیرهای هواشناسی ورودی به آن‌ها دقت‌های مختلفی را ارائه نمودند که با توجه به محدودیت متغیرهای هواشناسی در دسترس و دقت مورد نیاز می‌توان از این مدل‌ها بهره جست. نتایج ارائه شده در مقایسه با نتایج به‌دست آمده توسط طبری و همکاران (۲۰۱۲) برای اقلیم نیمه خشک فراسرد همدان (نشان می‌دهد که در هر دو مطالعه، مدل‌هایی داده محور که حداکثر تعداد ورودی داشته‌اند بهترین نتایج را ارائه دادند (RMSE برابر ۰/۹ میلی‌متر در روز) و تفاوت دقت مدل‌ها در دو بررسی را می‌توان به دو عامل مقیاس بررسی ماهانه مطالعه اقلیم همدان و روزانه در این بررسی (Tabari et al., 2012) و تفاوت اقلیمی آن با اقلیم‌های مد نظر در این مطالعه عنوان کرد.

منابع

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

خیرابی، ج، انتصاری، م، توکلی، ع و سلامت، ع. ۱۳۷۶. معرفی جهات نظری و کاربردی روش پنمن-مانیت و ارائه تبخیر-تعرق مرجع استاندارد برای ایران. انتشارات کمیته ملی آبیاری و زهکشی، چاپ اول، ۱۷۹ صفحه.

Tabari,H., Kisi,O., Ezani,A and Talae,P. 2012. SVM, ANFIS, regression and climate based models for reference evapotranspiration modeling using limited climatic data in a semi-arid highland environment. *Journal of Hydrology*, 444-445: 78-89.

Willmott,C.J. 1981. On the validation of models. *Physical Geography* 2: 184-194.

Quinlan,J.R. 1992. Learning with continuous classes. In- Adams, A. sterling, L. (eds) *Proc. AI'92, 5th Australian Joint Conference on Artificial Intelligence*. World scientific, singapore, 343-348.

Shing,J., Jang,R. 1993. ANFIS- Adaptive-Network-Based Fuzzy Inference System. *Man and Cybernetics*, 23(3): 665-685.

Evaluation of Intelligent Data-Driven Models in Daily Reference Evapotranspiration Estimation in Climates of Southern Coasts of the Caspian Sea

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

Recived: Jan. 21, 2014 Accepted: Aug. 27, 2014

Abstract

In many studies accurate estimation of evapotranspiration parameter has special importance. Intelligent data-driven models have great potential in the modeling of complex and nonlinear phenomena. In this study reference evapotranspiration was modeled by using of four data-driven method including Artificial Neural Networks (ANN), Support Vector Machines (SVM), model tree M5 and Adaptive Networks based on Fuzzy Inference System (ANFIS) with Long-term daily meteorological data in three synoptic stations of Climates of southern coast of the Caspian Sea. 11 different combinations of meteorological variables was chosen as input data-driven models and for evaluate the models many statistical were used such as correlation coefficient (R^2), Root Mean Square of Error (RMSE) and index of agreement (DI). The Result that showed ANFIS-11 model Was provided best estimations with RMSE between $0.2 \sim 0.38 \text{ mm day}^{-1}$ and models with single-input wind speed was presented poorest estimations with RMSE between $1.35 \sim 1.68 \text{ mm day}^{-1}$ and R^2 between $0.02 \sim 1.68$ in climates studied.

Key word: ANFIS, ANN, Daily Reference Evapotranspiration, M5, SVM

1 - Assistant Professor, Department of Water Engineering, Faculty of Agriculture, Shahid Bahonar University of Kerman, Kerman Iran

2 -M. Sc. Student of Water Resources Engineering, Department of Water Engineering, College of Agriculture, Shahid Bahonar University of Kerman, Kerman, (member of young researcher of Shahid Bahonar university of Kerman)

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

(*-Corresponding Author Email: Drbakhtiari@uk.ac.ir)