

بازسازی داده‌های مفقوده مجموع ماهیانه ساعات آفتابی با استفاده از شبکه‌های عصبی مصنوعی

مسعود کرباسی¹

تاریخ دریافت: 1394/2/5 تاریخ پذیرش: 1395/8/3

چکیده

تبخیر-تعرق گیاه مرجع یکی از عوامل مهم چرخه هیدرولوژیکی است که باید در طرح سیستم‌های آبیاری، تاسیسات آبی، مطالعات زهکشی و هیدرولوژیکی برآورد شود. یکی از داده‌های موردنیاز برای محاسبه این پارامتر بااهمیت، مقدار تابش خورشیدی می‌باشد که در صورت عدم وجود داده‌های آن از مجموع ماهیانه ساعات آفتابی استفاده می‌شود. با توجه به اینکه در اکثر ایستگاه‌های هواشناسی کشور در سال‌های گذشته داده‌های مربوط به مجموع ماهیانه ساعات آفتابی موجود نمی‌باشد، نیاز به بازسازی داده‌های مربوط به آن احساس می‌شود. در تحقیق حاضر با استفاده از دو نوع شبکه عصبی مصنوعی MLP و RBF و همچنین داده‌های هواشناسی ایستگاه هدف و ایستگاه‌های مجاور اقدام به بازسازی داده‌های مجموع ماهیانه ساعات آفتابی گردید. نتایج این تحقیق نشان داد که می‌توان با استفاده از داده‌های هواشناسی ایستگاه هدف و ایستگاه‌های مجاور، مجموع ماهیانه ساعات آفتابی را با دقت بالایی بازسازی کرد. نتایج سناریوهای مختلف اعمال شده نشان داد که در صورتی که صرفاً از داده‌های هواشناسی ایستگاه هدف استفاده شود می‌توان با پارامترهای هواشناسی حداقل و حداکثر دما، رطوبت نسبی متوسط، تابش فرازمینی و تعداد روزهای صاف، ابری و نیمه‌ابری با RMSE، 16/79 ساعت و درصد خطای متوسط 6/44 درصد مجموع ماهیانه ساعات آفتابی را تخمین زد. همچنین در صورتی که تنها از داده‌های ایستگاه مجاور استفاده شود، استفاده از ایستگاه‌های بیش‌تر منجر به افزایش دقت می‌شود (RMSE، 14/25 ساعت و درصد خطای متوسط 5/71 درصد). بهترین نتیجه زمانی به دست آمد که از هر دو سری داده هواشناسی ایستگاه هدف و ایستگاه‌های مجاور استفاده شود (RMSE، 13/78 و درصد خطای متوسط 4/97 درصد). مقایسه عملکرد دو شبکه عصبی مصنوعی MLP و RBF نشان داد که دقت شبکه عصبی مصنوعی MLP تا حدودی بیش‌تر از شبکه عصبی RBF می‌باشد. در پایان نیز سری زمانی تبخیر-تعرق مرجع برای سال‌هایی که داده مجموع ماهیانه ساعات آفتابی موجود نبود، بازسازی گردید.

واژه‌های کلیدی: بازسازی داده، تبخیر-تعرق، شبکه عصبی، مجموع ماهیانه ساعات آفتابی

مقدمه

دسترسی به داده‌های اندازه‌گیری شده RS توسط پیرانومتر است، اما به دلیل کمبود امکانات اندازه‌گیری آن در تعداد کمی از ایستگاه‌های سینوپتیک ایران امکان‌پذیر است. روش دوم بر مبنای استفاده از یک سری روابط و مدل‌های ریاضی استوار است. در این روش‌ها بین عوامل آب و هوایی از قبیل ساعات آفتابی، رطوبت نسبی، مقدار پوشش ابر و یا دمای حداقل و حداکثر که اندازه‌گیری آن‌ها آسان‌تر است، روابطی رگرسیونی خطی و یا غیرخطی برقرار می‌گردد (بیات و میرلطیفی، 1388).

داده‌های مجموع ماهیانه ساعات آفتابی یکی از داده‌های هواشناسی می‌باشد که در صورت عدم وجود داده‌های تابش خورشیدی در برآورد تبخیر-تعرق گیاه مرجع به‌صورت ماهانه مورد استفاده قرار می‌گیرد. مدل پنمن-مانتیت فائو 56 با مقایسه با داده‌های لایسیمتری به‌عنوان بهترین مدل جهت برآورد تبخیر-تعرق مرجع در مقیاس‌های ماهیانه، روزانه و ساعتی انتخاب شده است. این مدل به داده‌های تابش خورشیدی نیاز دارد که به دلایل ذکرشده در بالا در اکثر ایستگاه‌های هواشناسی در دسترس نمی‌باشد. راه‌حل

برآورد داده‌های مفقوده به‌عنوان اولین مرحله در مطالعات هیدرولوژی و اقلیم‌شناسی شناخته می‌شود. داده‌های مفقوده می‌توانند به علت عدم حضور دیده‌بان، خرابی دستگاه‌های اندازه‌گیری و یا قطعی ارتباط ایجاد شوند (Hassanpour and Dinpashoh., 2012). تابش خورشیدی رسیده به سطح زمین (Rs) از پارامترهای مهم و تاثیرگذار در توازن حرارتی سیستم جو-زمین است و در مهندسی انرژی، هواشناسی، کشاورزی، هیدرولوژی و سایر علوم کاربرد دارد (مجنون و همکاران، 1387). از مهم‌ترین موارد کاربرد تابش خورشیدی (Rs) در کشاورزی و آبیاری، استفاده از آن برای محاسبه تبخیر-تعرق به وسیله معادله‌های متداول، مانند پنمن-مانتیت فائو 56 است. برآورد مقادیر Rs در نقاط مختلف زمین به دو صورت میسر است، روش اول که بهترین و قابل اعتمادترین روش می‌باشد،

1- استادیار گروه مهندسی آب دانشگاه زنجان

خطای مطلق کم‌تری را نسبت به روش نزدیک‌ترین همسایه دارد. در روش نزدیک‌ترین همسایه با افزایش فاصله ایستگاه‌ها خطای برآورد افزایش می‌یابد. همچنین دقت هر دو روش در تخمین دمای حداکثر روزانه بیش‌تر از دمای حداقل روزانه می‌باشد. حسن‌پور و دین‌پژوه 11 روش مبتنی بر هوش مصنوعی و کلاسیک را برای تخمین داده‌های مفقوده در شرایط مختلف اقلیمی مورد بررسی قرار دادند. نتایج آن‌ها نشان داد که روش‌های مبتنی بر هوش مصنوعی قدرت بیش‌تری در تخمین داده‌های مفقوده دارند. در روش‌های مبتنی بر هوش مصنوعی، شبکه‌های عصبی پرسپترون چندلایه و در روش‌های کلاسیک نیز رگرسیون چند متغیره بهترین روش انتخاب شد (Hassanpour and Dinpashoh., 2012). طبری و حسین زاده از دو نوع شبکه عصبی MLP و RBF برای بازسازی داده‌های ماهانه کیفیت رودخانه مارون، استفاده نمودند. نتایج ایشان نشان داد که هر دو نوع شبکه عصبی قادر هستند با دقت بالایی داده‌های مفقوده را شبیه‌سازی نمایند. مقایسه عملکرد دو نوع شبکه عصبی نشان داد که شبکه عصبی MLP دارای دقت بالاتری است (Tabari and Hosseinzadeh., 2015). استفاده از شبکه‌های عصبی در شبیه‌سازی داده‌های مفقوده پارامترهای مختلف هواشناسی نیز گزارش شده است. لوند و همکاران از شبکه‌های عصبی در بازسازی داده‌های مفقوده بارش استفاده کردند (Londhe et al., 2015).

اگرچه مطالعات مختلفی برای بازسازی آماری پارامترهای هواشناسی از قبیل دما و بارش صورت گرفته است، لیکن تحقیقاتی در زمینه بازسازی مجموع ماهیانه ساعات آفتابی صورت نگرفته است. در تحقیق حاضر با استفاده از دو نوع شبکه عصبی مصنوعی MLP و RBF اقدام به بازسازی داده‌ها با توجه به پارامترهای هواشناسی ایستگاه موردنظر و ساعات آفتابی ایستگاه‌های مجاور دارای داده خواهد شد. در پایان نیز سری زمانی مجموع ساعات آفتابی برای سال‌های موجود بازسازی و جهت به دست آوردن سری زمانی طولانی‌مدت تبخیر-تعرق گیاه مرجع مورد استفاده قرار خواهند گرفت. عدم وجود داده‌های مجموع ساعات آفتابی ماهیانه در سال‌های قبل از 1982 برای ایستگاه زنجان، اهمیت بازسازی این داده‌ها را برای محاسبات تعیین روند و همچنین پیش‌بینی تبخیر-تعرق افزایش داده است.

مواد و روش‌ها

موقعیت جغرافیایی

ایستگاه سینوپتیک زنجان در عرض جغرافیایی 36 درجه و 41 دقیقه، طول جغرافیایی 48 درجه و 29 دقیقه قرار گرفته و ارتفاع آن از سطح دریا 1663 متر می‌باشد. متوسط دمای زنجان 10/8 درجه سانتی‌گراد، متوسط بارندگی سالانه 291 میلی‌متر و متوسط تبخیر و

جایگزین استفاده از ساعات آفتابی جهت محاسبه‌ی تابش خورشیدی با استفاده از معادله‌ی آنگستروم- پرسکات (رابطه 1) می‌باشد.

$$R_s = \left[a_s + b_s \frac{n}{N} \right] R_a \quad (1)$$

در رابطه‌ی فوق n ساعات آفتابی واقعی (ساعت)، N حداکثر ساعات آفتابی (ساعت)، R_a تابش فرازمینی (تابع زمان و مکان است) و a_s و b_s ضرایب ثابت معادله هستند که در صورت معلوم نبودن مقدار آن‌ها به ترتیب برابر 0/25 و 0/5 در نظر گرفته می‌شوند (Allen et al., 1998). ضرایب a_s و b_s در معادله‌ی آنگستروم- پرسکات در عین نزدیک بودن نسبی به مقادیر جهانی تفاوت‌هایی را در ایستگاه‌های مختلف ایران نشان می‌دهند. ضریب a_s معادله‌ی آنگستروم در ایران بین 0/21 در یزد تا 0/39 در کرمانشاه متفاوت می‌باشد. ضریب b_s معادله‌ی مذکور بین 0/34 در زنجان تا 0/62 در تبریز متغیر است (خلیلی و رضایی صدر، 1376).

به‌منظور به دست آوردن سری‌های زمانی طولانی مدت از تبخیر-تعرق جهت پیش‌بینی و تعیین روند تبخیر-تعرق نیاز است که سال‌های آماری زیادی برای تعیین سری زمانی استفاده شود. در اکثر ایستگاه‌های هواشناسی ایران در سال‌های گذشته (قبل از 1980) داده مجموع ماهیانه ساعات آفتابی موجود نمی‌باشد. برای برآورد دقیق‌تر تبخیر-تعرق مرجع ماهانه در سال‌های بدون داده ساعات آفتابی، نیاز است که داده‌های مربوط به آن بازسازی شوند.

در سال‌های اخیر استفاده از روش‌های هوش مصنوعی در بازسازی داده مفقوده هیدرولوژیکی و اقلیمی بسیار متداول گشته است. آبد و همکاران برای بازسازی داده‌های بارش از روش‌های شبکه عصبی مصنوعی، فازی و کلاسیک استفاده کردند. نتایج نشان داد که مدل مبتنی بر فازی کم‌ترین مقدار خطا را در برآورد داده‌های مفقوده دارد (Abebe et al., 2000). کولیالی و اورا عملکرد شبکه‌های عصبی مختلف را در برآورد داده‌های مفقوده هواشناسی مورد مطالعه قرار دادند. نتایج نشان داد که مدل شبکه عصبی پرسپترون چندلایه از دقت مناسبی برای برآورد داده‌های مفقوده بارش روزانه و حداقل و حداکثر دمای روزانه دارا است (Coulibaly and Evora., 2007). کیم و پاچپسکی داده‌های مفقوده بارش روزانه را با استفاده از شبکه‌های عصبی مصنوعی و درخت رگرسیونی برآورد نمودند. نتایج نشان‌دهنده دقت قابل‌ملاحظه شبکه‌های عصبی و درخت رگرسیونی در مقایسه با روش‌های کلاسیک است (Kim and Pachepsky., 2010). خورشید دوست و همکاران (1391) برای بازسازی داده‌های دمای حداکثر و حداقل روزانه روش نزدیک‌ترین همسایه و شبکه عصبی مصنوعی را برای 5 ایستگاه هواشناسی واقع در غرب استان تهران مورد بررسی قرار دادند. نتایج آن‌ها نشان داد که برای کلیه‌ی ایستگاه‌های مورد مطالعه روش شبکه عصبی متوسط

تغرق روزانه 2/8 میلی‌متر در روز می‌باشند.

شبکه‌های عصبی RBF¹

یکی از قدرتمندترین شبکه‌های عصبی مورد استفاده در مسایل برازش تابع، شبکه عصبی RBF است. این نوع شبکه نسبت به شبکه‌های عصبی پرسپترون پیش خورده، مزایای استراتژیکی دارد. برخلاف شبکه‌های MLP که دارای لایه‌های متوالی متعددی هستند، شبکه RBF از سه لایه ثابت تشکیل شده است. لایه ورودی که محل تزریق سیگنال‌های ورودی به شبکه است؛ لایه میانی یا طبقه RBF که شامل توابع RBF می‌شود و لایه خروجی که ترکیبی خطی از کلیه خروجی‌های طبقه RBF را می‌سازد. در اکثر موارد از توابع گوسی در لایه RBF استفاده می‌شود، که این توابع با دو پارامتر مرکز گوسی و واریانس یا میزان گستردگی گوسی شناسایی می‌شوند. عملیاتی که در این شبکه انجام می‌شود در فرم ماتریسی به صورت رابطه 2 و 3 است (کریمی زاد و شاهسون، 1392).

$$d_i(P) = \sum_{j=1}^N \varphi_j(P) \cdot w_{ij} \quad (2)$$

$$\varphi_j = \exp\left(\frac{-1}{2\sigma_j^2} \|P - C_j\|^2\right) \quad (3)$$

که در آن $d_i(P)$ خروجی i ام شبکه و φ_j خروجی گره RBF j ام به ازای بردار ورودی P است. w_{ij} نیز اتصال سیناپسی میان آن دو است. تعداد کل نرون‌های لایه پنهان نیز N است. بردار C_j را مرکز گوسی می‌نامیم و σ_j نیز مشخص‌کننده میزان گستردگی گوسی است (Hykin., 1999).

در تحقیق حاضر 70 درصد داده‌ها به صورت تصادفی برای آموزش شبکه و 30 درصد باقی مانده برای صحت سنجی مدل‌های پیشنهادی اختصاص پیدا کردند (Goyal et al., 2014).

معادله پنمن - ماتیت فائو 56

معادله پنمن - ماتیت فائو 56 به عنوان یک روش استاندارد برای برآورد تبخیر و تعرق مرجع گیاهان به صورت رابطه 4 ارایه شده است (Allen et al., 1998).

$$ET_0 = \frac{0.408\Delta(R_n - G) + \frac{900\gamma}{[\Delta + \gamma(1 + 0.34u)]} \cdot \frac{uD}{(T + 273)}}{\Delta + \gamma(1 + 0.34u)} \quad (4)$$

در معادله فوق ET_0 = تبخیر - تعرق مرجع (mm/d)

متوسط دمای هوا G ؛ $^{\circ}C$ ؛ G = شار حرارتی خاک (مگاژول بر مترمربع بر روز)؛ u = سرعت باد در ارتفاع 2 متری (m/s) ؛ R_n = تابش خالص در سطح گیاه (مگاژول بر مترمربع بر روز)؛ D = کمبود فشار بخار (برابر اختلاف فشار بخار اشباع و واقعی (kpa))؛ Δ = شیب منحنی فشار بخار اشباع $(kpa/^{\circ}C)$ ؛ γ = ضریب ثابت سایکرومتری $(kpa/^{\circ}C)$ (Allen et al., 1998). برای محاسبه

آمار و اطلاعات

در این تحقیق اطلاعات ماهانه دوره آماری 1982-2010 (1361-1389 شمسی) مورد استفاده قرار گرفت (در سال‌های فوق داده‌های مجموع ساعات آفتابی ماهانه موجود می‌باشد).

شبکه‌های عصبی مصنوعی

شبکه‌های عصبی سامانه‌ها و روش‌های محاسباتی نوینی هستند برای یادگیری ماشینی، نمایش دانش و در انتها اعمال دانش به دست آمده در جهت پیش‌بینی پاسخ‌های خروجی از سامانه‌های پیچیده، ایده اصلی این گونه شبکه‌ها الهام گرفته از شیوه کارکرد سیستم عصبی زیستی، برای پردازش داده‌ها و اطلاعات به منظور یادگیری و ایجاد دانش قرار دارد. عنصر کلیدی این ایده، ایجاد ساختارهایی جدید برای سامانه پردازش اطلاعات است (Hykin., 1999).

یک شبکه عصبی شامل اجزای سازنده لایه‌ها و وزن‌ها می‌باشد. لایه‌ها به سه دسته تقسیم می‌شوند که شامل لایه ورودی، لایه پنهان و لایه خروجی می‌باشند. در لایه ورودی اطلاعات خام به شبکه وارد می‌شود. لایه پنهان ممکن است از یک تا چند لایه باشد و کار پردازش اطلاعات در این لایه انجام می‌شود. تعداد نرون‌ها نیز در این لایه متغیر است، هر یک از عناصر بردار ورودی به یک لایه، حین اتصال به نرون‌های آن در وزن متناظر خود ضرب می‌شوند. فرآیند آموزش شبکه نیز در واقع تنظیم وزن‌های نرون‌ها در پاسخ به خطای محاسبه شده میان مقادیر محاسباتی و مقادیر هدف می‌باشد. سومین لایه نیز لایه خروجی است که عملکرد این لایه به فعالیت واحد پنهان و وزن نرون‌ها بستگی دارد (کریمی زاد و شاهسون، 1392).

شبکه‌های عصبی پرسپترون چندلایه MLP

شبکه‌های چندلایه پیش‌خور یکی از مهم‌ترین ساختارهای شبکه‌های عصبی مصنوعی می‌باشند. به‌طور معمول این شبکه‌ها شامل مجموعه‌ای از واحدهای حسی (نرون‌های پایه) می‌باشند که تشکیل دهنده لایه ورودی، یک یا چندلایه پنهان و یک لایه خروجی می‌باشند. سیگنال ورودی در خلال شبکه و در مسیری روبه‌جلو به صورت لایه به لایه منتشر می‌شود. این نوع شبکه معمولاً با عنوان پرسپترون چندلایه (MLP) نامیده می‌شود (Hykin., 1999).

از میان روش‌های مختلف آموزش به روش پس انتشار خطا، الگوریتم لونیگ - مارکوارت، به دلیل همگرایی سریع‌تر در آموزش شبکه‌های با اندازه متوسط، برای استفاده در تحقیق حاضر انتخاب شده است (Bilgili and Ozgoren., 2011).

نرخ آموزش در شبکه‌های MLP و مقادیر گستره¹ و تعداد توابع انتقال در شبکه‌های RBF با استفاده از آزمون و خطا تعیین گردید. بررسی شبکه‌های عصبی مصنوعی MLP با آزمون و خطا نشان داد که یک شبکه عصبی پرسپترون چندلایه با نرخ آموزش 0/05 با دولایه مخفی که لایه اول دارای 6 نرون و لایه دوم دارای 5 نرون است، بهترین نتایج را ارائه می‌کند. در شبکه عصبی RBF نیز با حداکثر 20 نرون مقدار خطا به حداقل می‌رسد. همچنین مقدار بهینه گستره نیز با آزمون و خطا 0/1 تعیین گردید.

در این تحقیق ایستگاه زنجان به عنوان ایستگاه هدف برای برآورد مجموع ماهیانه ساعت آفتابی مدنظر قرار گرفت. سه ایستگاه تهران، تبریز و قزوین به علت نزدیکی و موجود بودن داده‌های طولانی مدت مجموع ماهیانه ساعت آفتابی آن‌ها مورد استفاده قرار گرفت. همان‌طور که در بخش مواد و روش اشاره شد، سه نوع داده برای برآورد مجموع ماهیانه ساعات آفتابی مورد استفاده قرار گرفت.

جدول 1 پارامترهای آماری جهت برآورد مجموع ماهیانه ساعات آفتابی را با استفاده از پارامترهای هواشناسی ایستگاه زنجان و شبکه عصبی مصنوعی MLP و RBF را نشان می‌دهد. این روش زمانی کاربرد دارد که در ایستگاه‌های مجاور نیز اطلاعات مربوط به مجموع ماهیانه ساعات آفتابی در دسترس نباشد، این در حالی است که در اکثر ایستگاه‌های هواشناسی ایران در بین سال‌های 1979-1982 اطلاعات مربوط به مجموع ماهیانه ساعات آفتابی موجود نمی‌باشد. با توجه به جدول 1 بهترین نتیجه زمانی حاصل شده است که از تمام داده‌های هواشناسی موجود در ایستگاه هواشناسی استفاده شده است (سناریو 7). در این حالت مقدار RMSE برای شبکه عصبی MLP برای داده‌های آموزش و آزمون به ترتیب 13/65 و 16/79 ساعت به دست آمده است. مقدار MARE نیز برای داده‌های آموزش و آزمون به ترتیب 4/63 و 6/44 درصد به دست آمده است. مقدار ضرایب تبیین (R^2) برای داده‌های آموزش و آزمون به ترتیب 0/973 و 0/966 به دست آمده است. در صورت عدم وجود داده‌های مربوط به ابری بودن هوا و رطوبت نسبی، می‌توان از حداقل داده‌های هواشناسی دمای حداقل و حداکثر و تابش فرازمینی (سناریو 3) استفاده کرد. در این حالت مقدار RMSE برای داده‌های آموزش و آزمون به ترتیب 17/91 و 17/83 ساعت به دست آمده است که نسبت به سناریو شماره 7 مقدار قابل قبولی است. مقدار MARE نیز برای داده‌های آموزش و آزمون به ترتیب 6/66 و 6/59 درصد به دست آمده است که در حالت آزمون تفاوت چندانی با سناریو 7 ندارد. مقایسه دو نوع شبکه عصبی MLP و RBF نشان می‌دهد که دقت شبکه‌های عصبی MLP تا حدودی بیش‌تر از شبکه‌های عصبی RBF می‌باشد.

پارامترهای معادله پنمن-مانتیت نیاز به داده‌های هواشناسی زیر می‌باشد: 1- دمای حداقل (T_{min}) و دمای حداکثر (T_{max}) -2 رطوبت نسبی حداقل (RH_{min}) و رطوبت نسبی حداکثر (RH_{max}) -3 سرعت باد 4- تشعشع خورشیدی.

داده‌های مورد استفاده در بازسازی داده‌های مجموع ساعات آفتابی ماهانه

- 1- استفاده از داده‌های هواشناسی موجود در ایستگاه موردنظر (دمای حداکثر (T_{max})، دمای حداقل (T_{min})، متوسط رطوبت نسبی (RH_{mean})، تعداد روزهای صاف ($Abr02$)، تعداد روزهای قسمتی ابری ($Abr36$) و تعداد روزهای ابری ($Abr78$))
- 2- استفاده از داده‌های ساعات آفتابی و داده‌های هواشناسی ایستگاه‌های مجاور ایستگاه تبریز ($nTab$)، ایستگاه مهرآباد تهران ($nTeh$) و ایستگاه قزوین ($nQaz$)).
- 3- ترکیب موارد 1 و 2 معیارهای آماری برای مقایسه روابط

به منظور مقایسه روابط مختلف از روابط آماری 5-7 استفاده گردیده است (Coulibaly and Evora., 2007).

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (O_i - P_i)^2}{N}} \quad (5)$$

$$MARE = \frac{1}{N} \sum_{i=1}^N \left| \frac{O_i - P_i}{O_i} \right| \quad (6)$$

$$R^2 = \frac{\left(\sum_{i=1}^N (O_i - \bar{O}_i)(P_i - \bar{P}_i) \right)^2}{\sum_{i=1}^N (O_i - \bar{O}_i)^2 \sum_{i=1}^N (P_i - \bar{P}_i)^2} \quad (7)$$

در روابط فوق پارامتر O_i مقادیر مشاهداتی مجموع ساعات آفتابی ماهانه و P_i مقادیر برآوردی از مدل‌های مختلف و \bar{O}_i و \bar{P}_i نیز به ترتیب میانگین مقادیر مشاهداتی و برآوردی می‌باشند.

نتایج و بحث

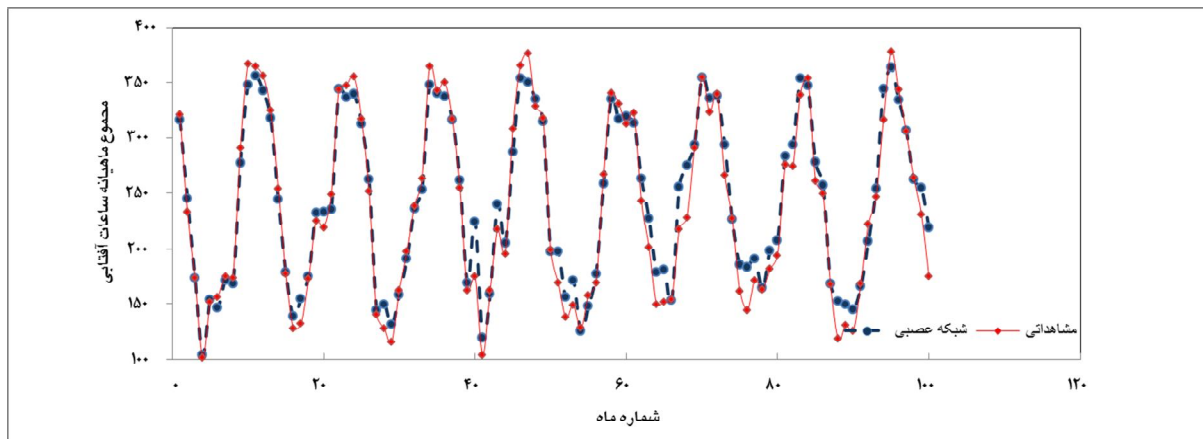
با استفاده از جعبه‌ابزار شبکه عصبی در نرم‌افزار MATLAB 2011، شبکه‌های متعددی با ساختارهای مختلف مورد بررسی و ارزیابی قرار گرفت. تعداد لایه‌های پنهان و تعداد نرون‌های هر لایه و

جدول 1- مقادیر پارامترهای آماری برای برآورد مجموع ماهیانه ساعات آفتابی با استفاده از داده‌های ایستگاه هواشناسی زنجان

شماره سناریو	پارامترهای هواشناسی	مدل	RMSE(hours)	MARE(%)	R ²
1	T_{max}	آموزش MLP	27/97	10/84	0/885
		آزمون MLP	23/24	8/99	0/927
2	T_{max}, T_{min}	آموزش RBF	27/59	10/75	0/886
		آزمون RBF	23/77	8/92	0/924
3	T_{max}, T_{min}, R_a	آموزش MLP	21/29	7/86	0/933
		آزمون MLP	22/61	8/42	0/932
4	$T_{max}, T_{min}, R_a, RH_{mean}$	آموزش RBF	20/79	7/73	0/935
		آزمون RBF	23/27	8/46	0/929
5	$T_{max}, T_{min}, R_a, RH_{mean}, Abr02$	آموزش MLP	17/91	6/66	0/953
		آزمون MLP	17/83	6/59	0/954
6	$T_{max}, T_{min}, R_a, RH_{mean}, Abr02, Abr36$	آموزش RBF	18/36	6/81	0/949
		آزمون RBF	18/61	6/85	0/955
7	$T_{max}, T_{min}, R_a, RH_{mean}, Abr02, Abr36, Abr78$	آموزش MLP	18/31	6/55	0/951
		آزمون MLP	18/15	6/43	0/954
		آموزش RBF	17/95	6/46	0/952
		آزمون RBF	18/77	6/73	0/951
		آموزش MLP	15/1	5/2	0/966
		آزمون MLP	16/87	6/64	0/962
		آموزش RBF	14/79	5/17	0/967
		آزمون RBF	17/33	6/65	0/960
		آموزش MLP	13/64	4/63	0/973
		آزمون MLP	16/93	6/50	0/965
		آموزش RBF	13/33	4/59	0/973
		آزمون RBF	17/93	6/79	0/963
		آموزش MLP	13/65	4/63	0/973
		آزمون MLP	16/79	6/44	0/966
		آموزش RBF	13/42	4/61	0/973
		آزمون RBF	17/45	6/51	0/964

جدول 2 پارامترهای آماری جهت برآورد مجموع ساعات آفتابی ماهیانه ایستگاه زنجان را با توجه به مجموع ساعات آفتابی ماهیانه در ایستگاه‌های مجاور تهران (مهرآباد)، تبریز و قزوین نشان می‌دهد. دلیل انتخاب ایستگاه‌های فوق طولانی بودن دوره آماری و نزدیکی به ایستگاه هواشناسی زنجان می‌باشد. با توجه به جدول 2 بهترین نتیجه در سناریو 14 (ترکیب هر سه ایستگاه هواشناسی) حاصل شده است. در این حالت مقدار RMSE برای شبکه عصبی MLP برای داده‌های آموزش و آزمون به ترتیب 14/63 و 14/25 ساعت به دست آمده است. مقدار MARE نیز برای داده‌های آموزش و آزمون به ترتیب 5/28 و 5/72 درصد به دست آمده است. مقدار ضرایب تبیین (R^2) برای داده‌های آموزش و آزمون به ترتیب 0/968 و 0/971 به دست آمده است.

شکل 1 داده‌های مشاهداتی مجموع ساعات آفتابی ماهیانه واقعی را در مقابل مجموع ساعات آفتابی ماهانه مدل شده توسط شبکه عصبی MLP در داده‌های آزمون سناریو 7 نشان می‌دهد. همان‌طور که از شکل مشاهده می‌شود، شبکه‌های عصبی MLP با دقت بالایی قادر به مدل کردن مجموع ساعات آفتابی ماهیانه ایستگاه هواشناسی زنجان با توجه به داده‌های هواشناسی ایستگاه می‌باشند. نتایج محققین مختلف در نقاط مختلف جهان نشان‌دهنده این موضوع است که با استفاده از داده‌های هواشناسی ایستگاه و کاربرد شبکه‌های عصبی مصنوعی می‌توان مقدار تشعشع رسیده به سطح زمین را که رابطه مستقیم با مجموع ساعات آفتابی دارد، تخمین زد (Atsu et al., 2002; Rehman and Mohandes., 2008; Koca et al., 2011; Yadav and Chandel., 2014).



شکل 1- مقایسه داده‌های مشاهده‌ای و مدل شده توسط شبکه عصبی با استفاده از داده‌های هواشناسی ایستگاه هواشناسی زنجان

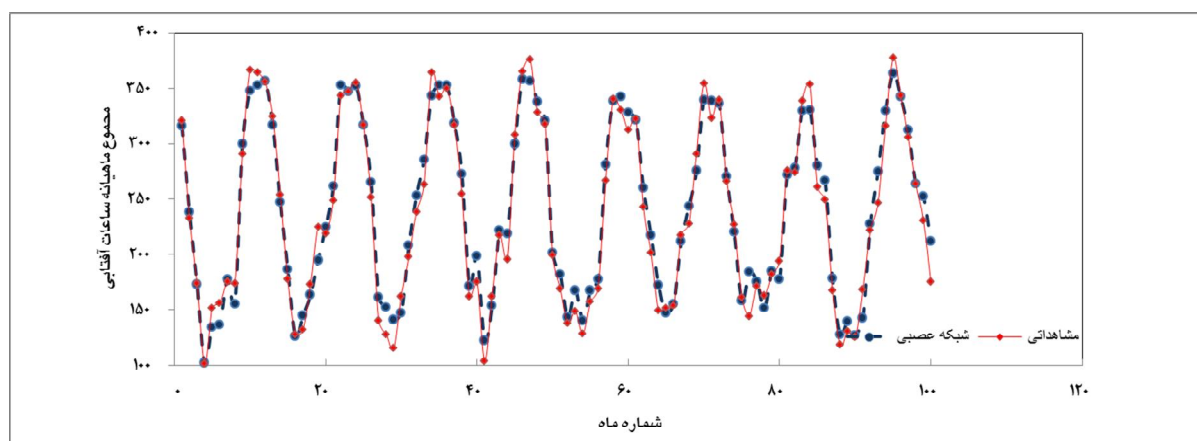
جدول 2- مقادیر پارامترهای آماری برای برآورد مجموع ساعت آفتابی ماهانه با استفاده از داده‌های ایستگاه‌های مجاور هواشناسی

R2	MARE(%)	RMSE(hours)	مدل	پارامترهای هواشناسی	شماره سناریو
0/947	7/21	18/92	آموزش	nTehran	8
0/919	8/11	23/44	آزمون		
0/948	7/11	18/64	آموزش	nTabriz	9
0/909	8/71	25/23	آزمون		
0/928	8/16	22/19	آموزش	nQazvin	10
0/952	7/56	18/64	آزمون		
0/929	7/87	21/7	آموزش	nTabriz, nTehran	11
0/956	7/18	18/56	آزمون		
0/960	6	16/46	آموزش	nTabriz, nQazvin	12
0/967	6/49	15/67	آزمون		
0/961	6/06	16/17	آموزش	nQazvin, nTehran	13
0/950	7/38	18/84	آزمون		
0/960	6/04	16/38	آموزش	nQazvin, nTabriz, nTehran	14
0/956	6/3	17/12	آزمون		
0/964	5/85	15/60	آموزش		
0/962	6/05	16/55	آزمون		
0/968	5/38	14/66	آموزش		
0/972	5/88	14/57	آزمون		
0/971	5/16	13/95	آموزش		
0/961	6/45	17/11	آزمون		
0/962	5/94	16/06	آموزش		
0/964	6/51	15/96	آزمون		
0/966	5/48	15/01	آموزش		
0/952	6/86	18/21	آزمون		
0/968	5/28	14/63	آموزش		
0/971	5/72	14/25	آزمون		
0/971	4/97	13/88	آموزش		
0/953	6/1	18/38	آزمون		

نشان می‌دهد که استفاده از داده‌های ایستگاه‌های مجاور (RMSE=14.25 h) نسبت به داده‌های هواشناسی ایستگاه هدف (RMSE=16.79 h) دقت بالاتری را ارائه می‌کند. مقایسه دو شبکه عصبی MLP و RBF نیز نشان می‌دهد که شبکه عصبی MLP دقت بالاتری را نسبت به شبکه عصبی RBF دارا است.

شکل 2 داده‌های مشاهداتی مجموع ساعات آفتابی ماهیانه واقعی را در مقابل مجموع ساعات آفتابی ماهانه مدل شده توسط شبکه عصبی MLP در داده‌های آزمون سناریو 14 نشان می‌دهد. همان‌طور که از شکل مشاهده می‌شود، شبکه‌های عصبی MLP با دقت بالایی قادر به مدل کردن مجموع ساعات آفتابی ماهیانه ایستگاه هواشناسی زنجان با توجه به داده‌های ساعت آفتابی ایستگاه‌های مجاور می‌باشند.

در صورتی که بخواهیم از داده‌های یک ایستگاه هواشناسی استفاده کنیم، ایستگاه قزوین با RMSE 15/67 ساعت برای داده‌های آزمون بهترین نتیجه را ارائه کرده است. بعد از ایستگاه قزوین، ایستگاه‌های تبریز و تهران به ترتیب با RMSE 18/64 و 23/44 ساعت قرار می‌گیرند. با توجه به نزدیکی ایستگاه قزوین و همچنین تشابه آب و هوایی آن با ایستگاه زنجان، نتیجه فوق منطقی به نظر می‌رسد. نتایج خورشید دوست و همکاران (1391) نیز نشان‌دهنده این موضوع است که با افزایش فاصله ایستگاه‌ها از ایستگاه موردنظر موجب افزایش خطا می‌گردد. اگر بخواهیم از دو ایستگاه استفاده شود، ترکیب دو ایستگاه قزوین و تبریز نتیجه بهتری را با RMSE 14/57 ساعت برای داده آزمون ارائه می‌کند که دلیل نتیجه بهتر مشابهت آب و هوایی و نزدیک بودن ایستگاه‌ها می‌باشد. مقایسه نتایج جداول 1 و 2



شکل 2- مقایسه داده‌های مشاهده‌ای و مدل شده توسط شبکه عصبی با استفاده از داده‌های هواشناسی ایستگاه‌های مجاور (تهران، تبریز و قزوین)

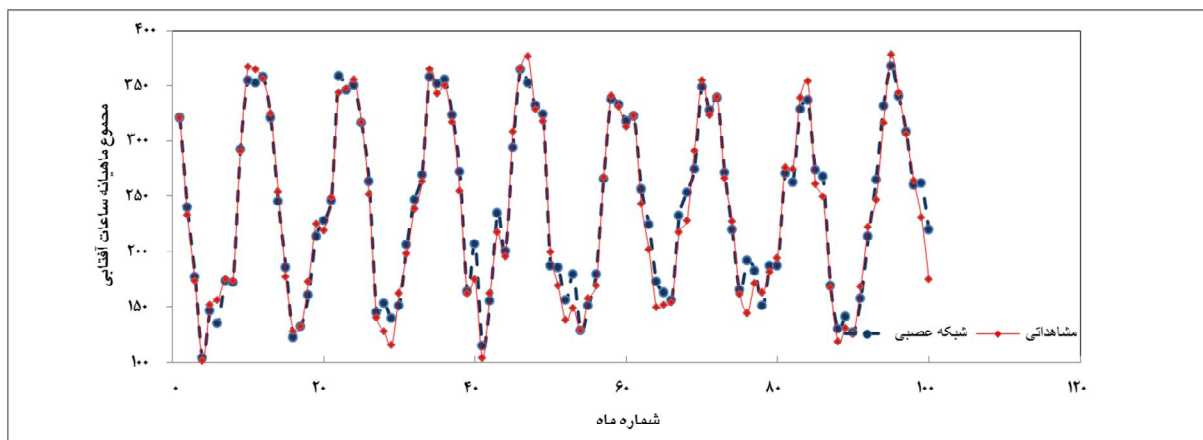
است. شکل 3 داده‌های مشاهداتی مجموع ساعات آفتابی ماهیانه واقعی را در مقابل مجموع ساعات آفتابی ماهانه مدل شده توسط شبکه عصبی MLP در داده‌های آزمون سناریو 19 نشان می‌دهد. همان‌طور که از شکل مشاهده می‌شود، شبکه‌های عصبی MLP با دقت بالایی قادر به مدل کردن مجموع ساعات آفتابی ماهیانه ایستگاه هواشناسی زنجان با توجه به داده‌های ایستگاه هواشناسی زنجان و ساعت آفتابی ایستگاه‌های مجاور می‌باشند.

بازسازی داده‌های مفقوده ساعات آفتابی در ایستگاه سینوپتیک زنجان در دو دوره زمانی (1955 الی 1972 و 1977 الی 1981) داده‌های مجموع ساعات آفتابی ماهانه در ایستگاه سینوپتیک زنجان موجود نمی‌باشد.

جدول 3 پارامترهای آماری جهت برآورد مجموع ساعات آفتابی ماهیانه ایستگاه زنجان را با توجه به مجموع ساعات آفتابی ماهیانه در ایستگاه‌های مجاور و داده‌های هواشناسی ایستگاه زنجان نشان می‌دهد. نتایج نشان داد که استفاده از ترکیب داده‌های هواشناسی در ایستگاه هدف و ایستگاه‌های مجاور نتایج بهتری را ارائه کرده است. در سناریو 19 استفاده از داده‌های هواشناسی ایستگاه زنجان و داده‌های مجموع ساعات آفتابی ماهانه سه شهر تهران، تبریز و قزوین با RMSE 13/52 ساعت برای داده‌های آزمون بهترین نتیجه را ارائه کرده است. مقایسه دو نوع شبکه عصبی MLP و RBF نشان داد که شبکه‌های عصبی MLP با اختلاف اندکی نتایج بهتری را ارائه می‌کنند. نتیجه فوق توسط طبری و حسین زاده (2015) و صیادی و همکاران (1388) نیز گزارش شده است. ولی نتایج کریمی‌زاد و شاهسون (1392) حاکی از برتری شبکه عصبی RBF بر MLP

جدول 3- مقادیر پارامترهای آماری برای برآورد مجموع ساعات آفتابی ماهانه با استفاده از داده‌های ایستگاه هواشناسی زنجان و ایستگاه‌های مجاور

0/972	4/97	13/78	آموزش	MLP	$T_{\max}, T_{\min}, R_a, nTab, nTeh$	15
0/953	6/54	18/02	آزمون			
0/971	5/34	14/02	آموزش	RBF	$T_{\max}, T_{\min}, R_a, RH_{mean}, Abr02, nTeh$	16
0/942	7/28	20/14	آزمون			
0/973	4/72	13/64	آموزش	MLP	$T_{\max}, T_{\min}, R_a, RH_{mean}, nTab, nTeh$	17
0/960	6/32	16/57	آزمون			
0/961	5/89	14/78	آموزش	RBF	$T_{\max}, T_{\min}, R_a, RH_{mean}, Abr02, nTab, nTeh$	18
0/932	7/33	20/75	آزمون			
0/974	4/74	13/39	آموزش	MLP	$T_{\max}, T_{\min}, R_a, RH_{mean}, Abr02, nTab, nTeh, nQaz$	19
0/956	6/21	17/21	آزمون			
0/970	5/14	14/16	آموزش	RBF		
0/956	5/96	17/86	آزمون			
0/976	4/41	12/67	آموزش	MLP		
0/963	6/24	15/93	آزمون			
0/975	4/55	12/99	آموزش	RBF		
0/958	6/12	17/48	آزمون			
0/979	3/98	11/85	آموزش	MLP		
0/974	5/14	13/52	آزمون			
0/978	4/25	12/14	آموزش	RBF		
0/974	4/97	13/78	آزمون			



شکل 3- مقایسه داده‌های مشاهده‌ای و مدل شده توسط شبکه عصبی با استفاده از داده‌های هواشناسی ایستگاه زنجان و ایستگاه‌های مجاور (تهران، تبریز و قزوین)

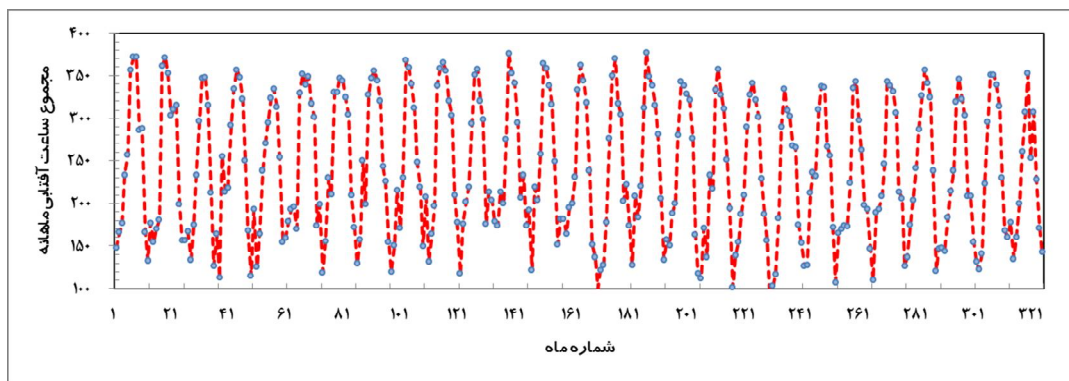
مفقوده استفاده شد. در سال‌های 1962 الی 1972 از ترکیب داده‌های هواشناسی ایستگاه زنجان و داده‌های مجموع ساعات آفتابی ماهانه ایستگاه‌های مهرآباد و تبریز برای بازسازی داده‌های مفقوده استفاده شد.

در دوره‌ی دوم که بین سال‌های 1977 الی 1981 می‌باشد، در سال 1977 داده‌های ایستگاه‌های مهرآباد و تبریز موجود است ولی در سال‌های 1978 الی 1981 داده‌ها موجود نمی‌باشد، بنابراین برای

در دوره‌ی اول در ایستگاه‌های مجاور در سال‌های 1955 تا 1962 تنها در ایستگاه مهرآباد تهران داده‌های مجموع ساعات آفتابی ماهیانه موجود می‌باشد. در ایستگاه هواشناسی تبریز از سال 1962 تا سال 1972 داده‌ها موجود می‌باشد و در ایستگاه قزوین به‌مانند ایستگاه زنجان داده‌ها موجود نمی‌باشند؛ بنابراین در طی سال‌های 1955 الی 1962 از ترکیب داده‌های هواشناسی ایستگاه زنجان و داده‌های مجموع ساعات آفتابی ماهانه ایستگاه مهرآباد برای بازسازی داده‌های

ماهنامه ساعت آفتابی را نشان می‌دهد. با استفاده از این سری زمانی می‌توان سری زمانی بلندمدت-تخیر-تعرق مرجع را ایجاد نمود.

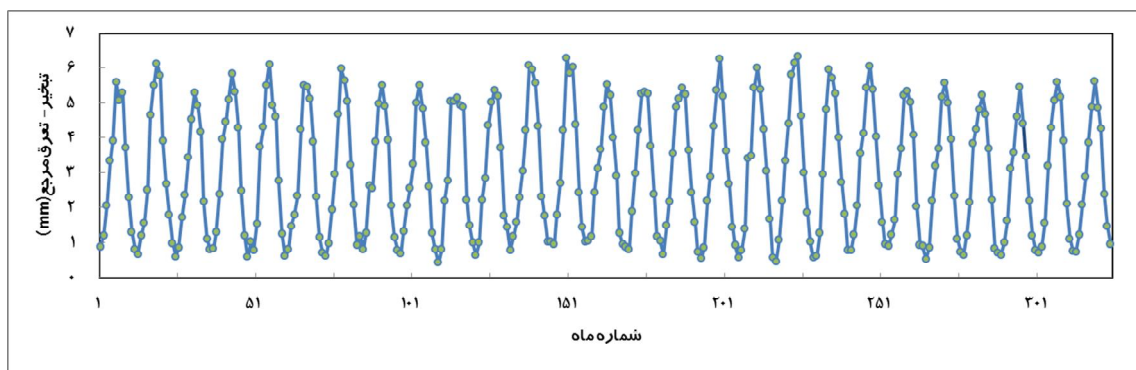
بازسازی داده‌های مفقوده صرفاً از داده‌های هواشناسی ایستگاه زنجان استفاده شد. شکل 4 سری زمانی بازسازی‌شده داده‌های مجموع



شکل 4- سری زمانی داده‌های بازسازی‌شده مجموع ساعات آفتابی

توجه به افزایش حافظه‌ی سری زمانی) می‌توان جهت بررسی دقیق‌تر پدیده‌های تغییر اقلیم و برنامه‌ریزی‌های بلندمدت استفاده کرد.

شکل 5 سری زمانی-تخیر-تعرق مرجع را در سال‌های 1955 تا 1981 را نشان می‌دهد. این سری زمانی از داده‌های بازسازی‌شده مجموع ماهنامه ساعت آفتابی ایجاد شده است. از سری زمانی فوق (یا



شکل 5- سری زمانی-تخیر-تعرق مرجع با استفاده از داده‌های بازسازی‌شده مجموع ساعات آفتابی ماهنامه

در صورت عدم وجود داده‌های ایستگاه هدف می‌توان از داده‌های ساعات آفتابی ایستگاه‌های مجاور استفاده کرد. نتایج این تحقیق نشان داد که با افزایش تعداد ایستگاه‌های مجاور می‌توان دقت را افزایش داد.

بیشترین دقت زمانی حاصل می‌شود که از ترکیب داده‌های هواشناسی ایستگاه هدف و داده‌های ساعات آفتابی ایستگاه‌های مجاور استفاده شود.

مقایسه دو نوع شبکه عصبی MLP و RBF نشان داد که شبکه‌های عصبی MLP با اختلاف اندکی بهتر عمل می‌کنند و دقت بالاتری را ارائه می‌کنند.

از بازسازی داده‌های ساعات آفتابی می‌توان سری زمانی طولانی-تری را استخراج نمود و در مطالعات پیش‌بینی آینده میزان-تخیر-

نتیجه‌گیری

در تحقیق حاضر به منظور بازسازی داده‌های مجموع ساعات آفتابی ماهنامه در ایستگاه سینوپتیک زنجان از شبکه‌های عصبی مصنوعی پرسپترون چندلایه و بر مبنای شعاعی استفاده گردید. خلاصه نتایج این تحقیق به شرح زیر می‌باشد:

شبکه‌های عصبی مصنوعی قادر هستند با دقت مناسبی مقدار مجموع ماهنامه ساعات آفتابی را با استفاده از داده‌های ایستگاه هدف و ایستگاه‌های مجاور تخمین بزنند.

در صورت استفاده از داده‌های ایستگاه هدف، علاوه بر داده‌های دما و رطوبت می‌توان از داده‌های تعداد روزهای هوای صاف، نیمه‌ابری و ابری در ماه استفاده نمود. نتایج این تحقیق نشان داد که استفاده از داده‌های ذکر شده دقت را افزایش داده است.

Goyal, M.K., Bharti, B., Quilty, J., Adamowski, J and Pandey, A. 2014. Modeling of daily pan evaporation in sub-tropical climates using ANN, LS-SVR, Fuzzy Logic, and ANFIS. Expert systems with applications. 41.11:5267-5276.

Hasanpour, M., Dinpashoh, Y. 2012. Evaluation of efficiency of different estimation methods for missing climatological data. Stochastic Environment Research and Risk Assessment 26:59-71

Hykin, S. 1999. Neural Networks: A Comprehensive Foundation. Printice-Hall, New Jersey.

Kim, J.W and Pachepsky, Y.A. 2010. Reconstructing missing daily precipitation data using regression trees and artificial neural networks for SWAT streamflow simulation. Journal of hydrology. 394.3:305-314.

Kadirgama, K., Amirruddin, A.K., Bakar, R.A. 2014. Estimation of Solar Radiation by Artificial Neural Networks: East Coast Malaysia. Energy Procedia. 52.3: 383-388

Koca, A., Hakan, F., Varol, Y., Koca, G.O. 2011. Estimation of solar radiation using artificial neural networks with different input parameters for Mediterranean region of Anatolia in Turkey. Expert Systems with Applications. 38.7: 8756-8762.

Londhe, S., Dixit, P., Shah, S and Narkhede, S., 2015. Infilling of missing daily rainfall records using artificial neural network. ISH Journal of Hydraulic Engineering. 21.3:255-264.

Rehman, S., Mohandes, M. 2008. Artificial neural network estimation of global solar radiation using air temperature and relative humidity. Energy Policy. 36.2: 571-576.

Tabari, H and Talaei, P.H. 2015. Reconstruction of river water quality missing data using artificial neural networks. Water Quality Research Journal of Canada. 50.4:326-335.

Yadav, A.K., Chandel, S.S. 2014. Solar radiation prediction using Artificial Neural Network techniques: A review. Renewable and Sustainable Energy Reviews. 33.5: 772-781

منابع

بیات، ک و میرلطیفی، م. 1388. تخمین تابش کل خورشیدی روزانه با استفاده از مدل‌های رگرسیونی و شبکه‌های عصبی مصنوعی. مجله علوم کشاورزی و منابع طبیعی. 16,3: 1-11

خلیلی، ع، رضایی صدر، ح. 1376. برآورد تابش کلی در گستره ایران بر مبنای داده‌های اقلیمی. تحقیقات جغرافیایی. 46,2: 15-35

خورشید دوست، ع، نساجی زواره، م و قرمز چشمه، ب. 1391. بازسازی سری‌های زمانی دمای حداکثر و حداقل روزانه با استفاده از روش نزدیک‌ترین همسایه و شبکه عصبی مصنوعی (مطالعه موردی غرب استان تهران). فصلنامه علمی-پژوهشی فضای جغرافیایی. 12,38: 197-214

صیادی، ح، غفاری، ا، فعالیان، ا و صدرالدینی، ع. 1388. مقایسه عملکرد شبکه‌های عصبی RBF و MLP در برآورد تبخیر و تغرق گیاه مرجع. مجله دانش آب و خاک، 19,1: 1-12

کریمی‌زاد، ف و شاهسوندا، ا. 1392. مقایسه نتایج حاصل از شبکه‌های عصبی MLP و RBF در پیش‌بینی نتایج حاصل از هم‌زمانی پدیده‌های انتقال جرم و انتقال حرارت. مجله مدل‌سازی در مهندسی. 11,33: 27-43

مجنونی، ا، زند پارسا، ش، سیاستخواه، ع و ناظم السادات، م. ج. 1387. توسعه و ارزیابی مدل‌های تخمین تابش خورشیدی بر اساس ساعات آفتابی و اطلاعات هواشناسی. علوم و فنون کشاورزی و منابع طبیعی. 12(46): 491-499

Abebe, A.J., Solomatine, D.P., Venneker, R.G.W. 2000. Application of adaptive fuzzy rule-based models for reconstruction of missing precipitation events. Hydrological Science Journal 45.3:425-436

Atsu, S.S., Dorvlo Joseph, A., Jervase, Ali Al-Lawati. 2002. Solar radiation estimation using artificial neural networks. Applied Energy. 71.4:307-319.

Bilgili, M and Ozgoren, M. 2011. Daily total global solar radiation modeling from several meteorological data. Meteorology and Atmospheric Physics. 112.3-4:125-138.

Coulibaly, P.N.D., Evora, B. 2007. Comparison of neural network methods for infilling missing daily weather records. Journal of hydrology.

Reconstruction of Missing Data of Monthly Total Sunshine Hours Using Artificial Neural Networks

M. Karbasi¹

Received: Oct.24, 2015

Accepted: Apr.25, 2016

Abstract

Reference crop evapotranspiration is one of the important factors of hydrological cycle. This parameter is used to design irrigation systems, hydraulic structures and drainage systems. One of data that required to calculate the amount of reference crop evapotranspiration is solar radiation which in the absence of this data, monthly sunshine duration data will be used. At the most of the weather stations of Iran the data of monthly total sunshine hours is not available at previous years, so the need to rebuild the data is felt. In the present study two kind of artificial neural network model (MLP and RBF) and meteorological data of target station and monthly total sunshine hours of neighbor stations are used to rebuild the missing data. The results showed that using data from meteorological data of target station and neighbor station, the total monthly sunshine reconstructed with high precision. The results of the different scenarios showed that if only the meteorological of target station such as minimum and maximum temperature, average relative humidity, solar radiation, extraterrestrial radiation and straight, dark and cloudy days number is used, with a precision of RMSE=16.79 hour and MAR=6.44% the monthly total sunshine hours is estimated. Also if only the data from nearby stations is used, the estimates would be more conducive to accuracy (RMSE=14.25 hour and MARE=5.71%). The best results were obtained when both weather data set of target station and adjacent stations are used (RMSE=13.78 hour and MARE=4.97%). Comparison of the performance of the ANN-MLP and RBF ANN-MLP showed that the accuracy of MLP neural network is somewhat greater. Finally the time series of monthly total sunshine hours and reference evapotranspiration were renovated.

Key words: Data Reconstruction, Evapotranspiration, Monthly Total Sunshine Hours, Neural Networks

1- Assistant Professor of Water Engineering Department, University of Zanjan
m.karbasi@znu.ac.ir