

Estimation of Solar Radiation Using Statistical Methods and Artificial Neural Network (Case Study-Tehran Synoptic Station)

Amir Zarei

* Ph.D Student, Department of GIS & RS, Faculty of Earth sciences, Shahid Beheshti University, Tehran, Iran.
(Corresponding Author)
amir69zarey@yahoo.com

Sirvan Zarei

MSc, Department of Health, Safety and Environment (HSE), Workplace Health Promotion Research Center, School of Public Health and Safety, Shahid Beheshti University of Medical Sciences, Tehran, Iran.

Seyyed Erfan Moomenpoor

Ph. D. Student, Department of Climatology, Faculty of Geography, University of Tehran, Tehran, Iran.

Mohammadreza Golanbaryan

MSc Student, Department of Hydraulic Structures, Faculty of Civil Engineering, University of Maragheh, Iran.

Abaas Solimanpor

MSc, Department of Watershed Engineering, Faculty of Natural Resources, Isfahan University of Technology, Isfahan, Iran.

Abstract

Background and purpose: Solar radiation is one of the important and influential parameters in agricultural hydrology and meteorology studies. According to the purpose of the research, which is to predict the amount of solar radiation, the type of research can be considered user

Materials and Methods: The statistical year in question is for one year (from the first of Farudin 1402 to the end of Farudin 1403) and the data used are the daily data of Tehran synoptic station. It includes the maximum and minimum temperature, maximum and minimum relative humidity and maximum wind speed (5 cases) and PM₁₀ data (PM_{2.5} from Tehran city pollution measurement station). Correlation relationship between solar radiation as a dependent variable and other parameters (independent components) was done through Pearson's correlation coefficient in SPSS.26 software

Results: and the results showed that solar radiation with minimum and maximum temperature, maximum wind speed, and PM₁₀ have a direct correlation with maximum and minimum relative humidity, wind direction and PM_{2.5} have an inverse correlation. The estimation of solar radiation has been done using multiple linear regression and artificial neural network (MLP) methods, with a coefficient of determination (R^2) of nearly 87% that performs better than the other two methods. (Enter and Stepwise)

Conclusion: Also, the comparison of the training and test period of the model shows that the ANN method with the mean square error (MSE) of megajoules per square meter per day, the absolute value of the error is 82.02 and the explanation coefficient is 88.0 in the test phase and 1.87 megajoules respectively. 35.86 and 81.0 Joules per square centimeter per day in the training phase of the models have performed better than other models in estimating solar radiation.

Keywords: Solar Radiation, Meteorological Parameters, Multiple linear Regression, Artificial Neural Network

Open Access Policy: This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited. To view a copy of this licence, visit <https://creativecommons.org/licenses/by/4.0/>

Received: 2024/12/24

Accepted: 2025/04/24

Doi:10.22038/jreh.2025.82332.1683

► **Citation:** Zarei A, Zarei S, Moomenpoor S.E, Golanbaryan M.R, Solimanpor A. Estimation of Solar Radiation Using Statistical Methods and Artificial Neural Network (Case Study-Tehran Synoptic Station). *Iranian Journal of Research in Environmental Health*. 11(2):24-41.

تخمین تابش خورشید با استفاده از روش‌های آماری و شبکه عصبی مصنوعی (مطالعه موردی - ایستگاه سینوپتیک تهران)

امیر زارعی

* دانشجوی دکتری تخصصی، گروه سیستم اطلاعات جغرافیایی و سنجش‌ازدور (GIS & RS)، دانشکده علوم زمین، دانشگاه شهید بهشتی، تهران، ایران. (نویسنده مسئول)
Amir69zarey@yahoo.com

سیروان زارعی

کارشناسی‌ارشد، گروه سلامت، ایمنی و محیط‌زیست (HSE)، مرکز تحقیقات ارتقاء سلامت محیط کار، دانشکده بهداشت و ایمنی، دانشگاه علوم پزشکی شهید بهشتی، تهران، ایران.

سیدعرفان مومن‌پور

دانشجوی دکتری تخصصی گروه آب و هواشناسی، دانشکده جغرافیا، دانشگاه تهران، ایران.

محمدرضا گل‌عنبریان

دانشجوی کارشناسی‌ارشد مهندسی عمران گرایش سازه‌های هیدرولیکی، دانشگاه مراغه، ایران.

عباس سلیمان‌پور

دانش آموخته کارشناسی‌ارشد، گروه مهندسی آبخیزداری، دانشکده منابع طبیعی، دانشگاه صنعتی اصفهان، اصفهان، ایران.

چکیده

زمینه و هدف: تابش خورشیدی یکی از پارامترهای مهم و تأثیرگذار در مطالعات هیدرولوژی و هواشناسی کشاورزی می‌باشد. با توجه به هدف تحقیق که پیش‌بینی میزان تابش خورشیدی می‌باشد، می‌توان نوع تحقیق را کاربری دانست.

مواد و روش‌ها: سال آماری موردنظر به مدت یک سال (اول فروردین ۱۴۰۲ تا آخر فروردین ۱۴۰۳) می‌باشد و داده‌های مورد استفاده، داده‌های روزانه ایستگاه سینوپتیک تهران شامل بیشینه و کمینه دما، بیشینه و کمینه رطوبت نسبی و حداکثر سرعت باد می‌باشد (۵ مورد) و داده‌های PM_{10} ، $PM_{2.5}$ (از ایستگاه سنجش آلودگی شهر تهران) می‌باشد. رابطه‌ی همبستگی بین تابش خورشیدی به عنوان متغیر وابسته و پارامترهای دیگر (مؤلفه‌های مستقل) از طریق ضریب همبستگی پیرسون در نرم‌افزار SPSS.26 انجام گرفت.

یافته‌ها: نتایج نشان داد که تابش خورشیدی با کمینه و بیشینه دما، حداکثر سرعت باد، و PM_{10} دارای همبستگی مستقیم و با بیشینه و کمینه رطوبت نسبی، جهت باد و $PM_{2.5}$ دارای همبستگی معکوس می‌باشد. تخمین تابش خورشیدی با استفاده از روش‌های رگرسیون خطی چندگانه و شبکه عصبی مصنوعی (MLP) انجام شده که روش شبکه عصبی مصنوعی با ضریب تعیین (R^2) نزدیک ۸۷ درصد دارای عملکرد بهتری از دو روش دیگر (Enter و Stepwise) می‌باشد.

نتیجه‌گیری: مقایسه سه روش (Stepwise و Enter، Ann) در مرحله‌ی آموزش و تست نشان از عملکرد بهتر روش شبکه عصبی دارد (مقدار بالاتر ضریب R^2)، مقدار شاخص MSE (میانگین مربعات خطا) در هر دو مرحله‌ی آموزش و آزمون در روش شبکه عصبی بهتر و به ترتیب ۱/۸۷ و ۱/۶۸ می‌باشد در حالی که همین شاخص برای روش Enter، ۳/۶۵ و ۳/۳۷ بوده و هم‌چنین برای روش Stepwise نیز مقادیر ۱۱/۶ و ۱/۷۵ محاسبه گردیده است نتایج نشان می‌دهد که در هر سه روش مدل‌ها در مرحله آزمون بهبود یافته‌اند.

کلیدواژه‌ها: تابش خورشیدی، پارامترهای هواشناسی، رگرسیون خطی چندگانه، شبکه‌ی عصبی مصنوعی

◀ **استناد:** زارعی، ا.، زارعی، س.، مومن‌پور، س.ع.، گل‌عنبریان، م.ر.، سلیمان‌پور، ع. تخمین تابش خورشید با استفاده از روش‌های آماری و شبکه عصبی مصنوعی (مطالعه موردی - ایستگاه سینوپتیک تهران). فصلنامه‌ی پژوهش در بهداشت محیط. تابستان ۱۴۰۴؛ ۱۱(۲): ۴۱-۲۴.

تاریخ دریافت: ۱۴۰۳/۱۰/۰۴

تاریخ پذیرش: ۱۴۰۴/۰۲/۰۴

نوع مقاله: پژوهشی

تابش جهانی خورشیدی^۱ یا به اختصار تابش خورشیدی به مجموع تابش مستقیم، پراکنده شده و بازتابشی خورشیدی اطلاق می‌شود. استفاده از انرژی خورشیدی از جهات مختلف بسیار حائز اهمیت است. انرژی‌های فسیلی مثل نفت و گاز علاوه بر این که تجدیدنپذیرند، در آلودگی محیط‌زیست نیز تأثیرگذارند؛ لذا گرایش به سمت منابع انرژی جدید مثل انرژی خورشیدی امری اجتناب‌ناپذیر است. این پدیده هم‌چنین در هواشناسی و کشاورزی اهمیت زیادی دارد. فتوسنتز و تبخیر و تعرق دو فرایند اساسی رشد محصول هستند که هر دو ارتباط تنگاتنگی با تابش خورشیدی داشته و لذا داشتن اطلاعات از کمیت و کیفیت آن می‌تواند نقش به‌سزایی در این زمینه داشته باشد. امروزه در حوزه‌های مختلفی از این انرژی پاک استفاده می‌شود که آب‌گرم‌کن‌های خورشیدی، خوراک‌پزهای خورشیدی، خشک‌کن‌های خورشیدی، پمپ‌های خورشیدی، حصارهای الکتریکی خورشیدی، آب‌شیرین‌کن‌های خورشیدی و غیره تنها بخشی از آن‌ها هستند. هم‌چنین در زمینه فن‌آوری‌های نوین، استفاده از انرژی خورشیدی در لوله‌های گرمایی، ترموالکتریک، پیل‌های سوختی و نانو، جاذبه‌های فراوانی پیدا کرده‌اند (۱). تمایل به استفاده از انرژی‌های تجدیدپذیر که یکی از مهم‌ترین آن‌ها انرژی خورشیدی می‌باشد. در طول سال‌های اخیر مضرات و محدودیت‌هایی که سوخت‌های فسیلی دارند، افزایش پیدا کرده است به‌طوری‌که در طی سال‌های ۲۰۰۰ تا ۲۰۰۷ سرانه تولید انرژی جهانی ۱/۰۳ افزایش یافته (۲) و پیش‌بینی می‌گردد تا سال ۲۰۳۰ این افزایش به ۶۰ درصد نیز برسد (۳).

منشأ تمام انرژی‌هایی که در سطح کره‌ی زمین وجود دارد، انرژی تابشی از خورشید می‌باشد. این انرژی دریافتی خورشیدی (تابش خورشیدی) در محاسبه‌ی مباحث مهمی از جمله هیدرولوژی، مدیریت منابع آب، مدل‌های بیلان آبی و... به‌کار می‌رود (۲). بسیاری از مردم معتقدند که در میان انواع مختلف منابع انرژی تجدیدپذیر مانند انرژی خورشیدی، انرژی باد، انرژی زیست‌توده، انرژی حرارتی و

غیره. انرژی خورشیدی به‌عنوان بهترین جایگزین برای انرژی‌های فسیلی هستند، به‌دلیل مزایایی مانند دردسترس‌بودن و سازگاری با محیط‌زیست مقدار تابش خورشیدی، یکی از مهم‌ترین پارامترها در طراحی و استفاده‌ی بهینه از سیستم‌های خورشیدی مختلف می‌باشد (۴).

پیچیدگی فرایندهای هواشناسی و هیدرولوژی از نظر انتخاب تمام پارامترهای تأثیرگذار بر آن‌ها و نقص اطلاعات آماری، امکان مدل‌سازی این فرایندها را غیرممکن می‌سازد. در چنین شرایطی استفاده از مدل‌سازی مبنی بر روابط ریاضی، موردتوجه قرار گرفته است (۵). تابش خورشیدی یکی از متغیرهای مهم و مؤثر هواشناختی در برآورد تبخیر و تعرق و نیاز آبی گیاهان است و منشأ انرژی برای همه تحولات جو و سطح زمین است. اندازه‌گیری این متغیر اگرچه در ایران سابقه نسبتاً طولانی دارد؛ ولی به‌دلیل هزینه‌های زیاد وسایل اندازه‌گیری در بسیاری از ایستگاه‌های موجود کشور دستگاه تابش‌سنج یا پیرانومتر وجود ندارد و یا مشکلاتی هم‌چون واسنجی آن، تجمع آب و گردوغبار بر روی سنجنده آن وجود دارد (۶). حتی در ایستگاه‌های هواشناسی هم که تابش را اندازه می‌گیرند، روزهایی وجود دارد که داده‌های تابش ثبت نمی‌شود یا مقادیر غیرواقعی و خارج از بازه‌ی مورد انتظار به‌دلیل نقص دستگاه و یا مشکلات دیگر مشاهده می‌شود (۷). هر چند در اغلب این ایستگاه‌ها ساعات آفتابی به‌طور روزانه اندازه گرفته می‌شود. تحقیقات مختلفی برای تخمین تابش خورشیدی با استفاده از داده‌های هواشناسی انجام شده و روش‌های زیادی توسعه یافته است. محققین به‌دنبال راهی برای تخمین بهتر و دقیق‌تر تابش خورشیدی هستند. در نتیجه روابط تجربی و رگرسیونی، فنون سنجش از دور و میان‌یابی خطی توسعه یافتند (۸)؛ اما در تخمین تابش خورشیدی توسط معادلات تجربی و نیمه‌تجربی، تنها تعداد محدودی از متغیرهای هواشناسی کاربرد دارد. در سال‌های اخیر پژوهشگران زیادی مطالعات خود را بر مبنای استفاده از روش‌های داده‌کاوی و مدل‌سازی

¹ Global Solar Radiation

ریاضی برای تخمین تابش خورشیدی معطوف داشته‌اند (۹، ۱۰، ۱۱، ۱۲، ۱۳، ۱۴، ۱۵، ۱۶، ۱۷، ۱۸، ۱۹، ۲۰، ۲۱، ۲۲، ۲۳ و ۲۴). با توجه به این واقعیت که هزینه‌های پولی زیادی برای نصب و راه‌اندازی و ایجاد تأسیسات و تجهیزاتی که برای اندازه‌گیری میزان تابش در برخی مناطق استفاده می‌شود، مورد نیاز است. اتخاذ معادلات تجربی و هم‌چنین مدل‌سازی روش‌هایی مانند شبکه عصبی (ANN)، آنالیز بردار پشتیبان (SVM) و.... در دهه‌های اخیر در نظر گرفته شده است (۲۵). داده‌های تابش خورشیدی برای بسیاری از پژوهشگران جهت انجام تحقیقات بسیار ضروری می‌باشد از آنجایی که این داده‌ها به‌وسیله‌ی دستگاه‌های پیرانومتر اندازه‌گیری می‌شوند و این دستگاه بسیار پرهزینه و گران‌قیمت هستند؛ لذا فقط در مکان‌های محدودی این داده‌ها اندازه‌گیری شده و این محدودیت باعث می‌شود در بسیاری پروژه‌ها احتیاج به تخمین این داده‌ها باشد؛ لذا با توجه گستردگی و وسعت کشوری ایران نصب دستگاه‌های ثبت مقدار تابش خورشید عملاً غیرممکن می‌باشد. در نتیجه این مشکلات دستگاه‌های اندازه‌گیری، فقط در مکان‌های محدودی قرار داده شده‌اند و در این شرایط، تخمین مقدار این پارامتر اهمیت خاصی پیدا می‌کند. تحقیقات گذشته در ایران نشان می‌دهد بیشتر مطالعات انجام شده مبتنی بر استفاده از مدل‌های تجربی است و ضروری است که کارایی روش‌های هوش مصنوعی در برآورد تابش خورشیدی ارزیابی گردد. از یک طرف، با توجه به این که در بسیاری از ایستگاه‌ها امکان اندازه‌گیری دقیق تابش خورشیدی وجود ندارد و از طرف دیگر لازم است تا متغیرهای هواشناسی کمتری در معادلات مورد استفاده قرار گیرد، لذا بررسی روابط غیرخطی موجود بین متغیرهای هواشناسی مؤثر بر تابش خورشیدی ضروری به نظر می‌رسد. با توجه به اهمیت تابش رسیده به سطح زمین و عوامل متعدد تأثیرگذار بر آن، هدف از این تحقیق تخمین تابش خورشیدی در ایستگاه سینوپتیک تهران با یک اقلیم تقریباً خشک با مجموعه‌ای از داده‌های هواشناسی و آلودگی هوا ضروری به نظر می‌رسد.

با بررسی داده‌های ورودی و خروجی ایستگاه موردنظر تنها قادر به انتخاب یکسال به عنوان پایه آماری مشترک در بین آنها شدیم با بررسی اولیه، مواردی از داده‌های گمشده در

برخی روزها مشاهده شد، به‌ویژه در متغیرهای $PM_{2.5}$ ، PM_{10} و RS برای مقادیر گمشده جزئی (کمتر از ۵٪ از کل داده‌ها)، از روش میانگین‌گیری متحرک با پنجره ۳ روزه استفاده شد. در صورتی که بیش از ۵ روز متوالی داده‌ای ثبت نشده بود، آن بازه از تحلیل حذف گردید.

روش کار

محدوده‌ی مورد مطالعه

ایستگاه هواشناسی مورد مطالعه، ایستگاه سینوپتیک تهران در منطقه‌ای با طول جغرافیایی $51/32$ درجه‌ی شمالی و عرض $35/44$ درجه‌ی شرقی و ارتفاع از سطح دریای آن 1419 متر می‌باشد. این منطقه دارای آب‌وهوای خشک و متوسط ساعات آفتابی منطقه نیز 2851 ساعت می‌باشد داده‌های مورد استفاده در این مطالعه، متغیرهای T_{max} ، T_{min} ، RH_{min} ، RH_{max} ، $PM_{2.5}$ ، PM_{10} و حداکثر سرعت باد به‌عنوان ورودی و RS (مقدار تابش خورشید)، به‌عنوان خروجی مدل‌ها، در مقیاس زمانی روزانه از ایستگاه سینوپتیک مطالعاتی طی دوره‌ی آماری $1402-1403$ از سازمان هواشناسی کشور اخذ گردیدند. برای تحلیل داده‌های موجود از نرم‌افزارهای $Matlab2024$ $SPSS26.0$ استفاده شد.

با بررسی داده‌های ورودی و خروجی ایستگاه موردنظر تنها قادر به انتخاب یکسال به عنوان پایه آماری مشترک در بین آنها شدیم با بررسی اولیه، مواردی از داده‌های گمشده در برخی روزها مشاهده شد، به‌ویژه در متغیرهای $PM_{2.5}$ ، PM_{10} و RS برای مقادیر گمشده جزئی (کمتر از ۵٪ از کل داده‌ها)، از روش میانگین‌گیری متحرک با پنجره ۳ روزه استفاده شد. در صورتی که بیش از ۵ روز متوالی داده‌ای ثبت نشده بود، آن بازه از تحلیل حذف گردید.

رگرسیون خطی چند متغیره

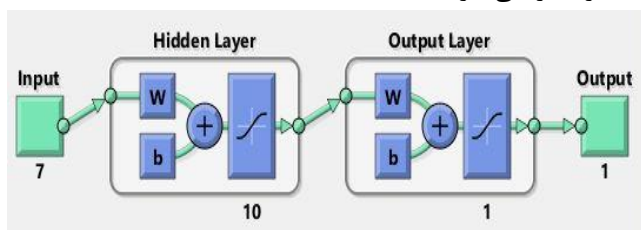
تحلیل رگرسیونی، پرکاربردترین روش در بین تکنیک‌های آماری است که برای بررسی و مدل‌سازی ارتباط بین متغیرها به کار می‌رود. رگرسیون چندمتغیره خطی روشی است برای تحلیل مشارکت جمعی و فردی دو یا چند متغیر مستقل در تغییرات یک متغیر وابسته است؛ بنابراین تحلیل رگرسیونی با به‌دست‌آوردن یک مدل در پیش‌بینی

استفاده می‌گردد. مقدار VIF اگر کمتر از ۵ باشد فرض عدم هم خطی متغیرهای مستقل برقرار است و مقدار آن بیشتر از ۵ برای این شاخص نشان از هم خطی جدی بوده و باعث عدم اعتماد به نتایج می‌شود. روش‌های انتخابی در روش رگرسیون خطی چندگانه دو روش ۱-روش همزمان^۳ و ۲-روش گام به گام^۴ می‌باشد (۲۵).

مزیت روش‌های آماری (مانند رگرسیون خطی) در این است که ساده هستند و رابطه‌ی بین متغیرها رو خطی در نظر می‌گیرد در حالی که در طبیعت کمتر چنین چیزی اتفاق می‌افتد اما روش شبکه‌های عصبی مصنوعی قابلیت مدل سازی روابط غیرخطی و پیچیده بین متغیرها را دارا هستند و از دقت بالایی نیز برخوردار می‌باشند.

روش شبکه عصبی مصنوعی (ANN)

شبکه‌ی عصبی مصنوعی، یکی از روش‌های محاسباتی است که به کمک فرایند یادگیری و با استفاده از پردازشگرهای ساده به نام نرون تلاش می‌کند با شناخت روابط ذاتی میان داده‌ها، نگاشتی میان فضای ورودی (لایه ورودی) و فضای مطلوب (لایه خروجی) ارائه دهد. در شبکه‌های عصبی مصنوعی سعی بر این است که ساختار بیولوژیکی مغز انسان و شبکه بدن طراحی شود تا همانند آن قدرت یادگیری، تعمیم‌دهی و تصمیم‌گیری داشته باشد. متغیرهای زیادی در یک مدل ANN، از جمله ورودی، تعداد لایه‌ها و نرون‌ها، الگوریتم آموزش، عملکرد انتقال وجود دارد. هر گونه تغییر در آن‌ها می‌تواند یک مدل ANN جدید ایجاد کند (۲۲). یک ساختار شبکه‌ای از تعدادی عناصر مرتبط به هم به نام نرون تشکیل شده است که هر نرون دارای ورودی‌ها و خروجی‌هایی است و یک عمل نسبتاً ساده و محلی را انجام می‌دهد ANN عموماً عملکرد خود را طی یک مرحله یادگیری فرا می‌گیرد (۸) (شکل ۱).



متغیرهای وابسته مختلف از جمله پارامترهای اقلیمی و هیدرولوژیکی، نقش بارزی ایفا می‌کند (۲۱). کاربرد مدل‌های رگرسیونی برای برآورد متغیرهای مختلف توسط پژوهشگران بسیاری مورد استفاده قرار گرفته است (۲۰)؛ (۲۶) از آنجایی که این روش‌ها به خوبی می‌توانند تأثیر یک متغیر مستقل را بر متغیر وابسته نشان دهند، سعی گردید با استفاده از معادله رگرسیونی چند متغیره نیز، ارتباط بین تابش خورشیدی با متغیرهای هواشناسی و آلودگی در ایستگاه مورد بررسی قرار گیرد. رگرسیون، یک معادله را که فاصله بین خط برازش شده و همی نقاط داده شده را به حداقل برساند، ارائه می‌کند. معادله‌ی کلی رگرسیون چند متغیره به صورت رابطه‌ی زیر بیان می‌شود.

$$Rs = C_1(T_{min}) + C_2(T_{max}) + C_3(RH_{min}) + C_4(RH_{max}) + C_5(PM_{2.5}) + C_6(PM_{10}) + C_7(WS_{max})$$

که در آن، Rs شدت تابش خورشید (Jcm^{-2})، T_{max} و T_{min} به ترتیب حداکثر و حداقل دما، RH_{min} و RH_{max} حداکثر و حداقل رطوبت نسبی، $PM_{2.5}$ و PM_{10} به ترتیب غلظت مواد معلق با اندازه‌های ۲/۵ و ۱۰ میکرون می‌باشد و در نهایت WS_{max} معرف حداکثر سرعت باد ایستگاه در نظر گرفته شده است.

$C_1, C_2, C_3, C_4, C_5, C_6$ و C_7 پارامترهای مدل هستند مدل رگرسیون خطی چندگانه در فرم ماتریسی را می‌توان به صورت معادله‌ی (۱) نشان داد.

$$Y = X \beta + \varepsilon \quad \text{معادله ۱}$$

که در آن β ماتریس ضرایب رگرسیون، (ماتریس خطای برازش، Y ماتریس پاسخ (متغیر وابسته) و X ماتریس متغیرهای مستقل می‌باشد. هم خطی در مدل رگرسیونی به وجود همبستگی بین متغیرهای مستقل اشاره دارد و نتایج مدل رگرسیون چندگانه را غیر معتبر می‌کند. به منظور شناسایی هم خطی از معیار VIF برای مدل رگرسیونی

³ Enter Method

⁴ Stepwise Method

² Variance Inflation Factor

شکل ۱. معماری یک شبکه عصبی مصنوعی (ANN) جهت

برآورد شدت تابش خورشید با توجه به متغیرهای ورودی

تنظیم پارامترهای شبکه عصبی مورد استفاده: تعداد لایه-های ورودی (بر اساس پارامترهای در نظر گرفته شده برای تخمین تابش خورشیدی) ۷ مورد و تعداد لایه‌های پنهان نیز ۱۰ می‌باشد (شکل ۱) در ضمن چون مدل یک مقدار را پیش‌بینی می‌کند (مثلاً مقدار تابش خورشیدی در یک روز مشخص)، پس فقط یک نورون دارد بنابراین مدل از ورودی‌های ۷ تایی شروع می‌کند، آن‌ها را در یک لایه پنهان با ۱۰ نورون پردازش می‌کند، و در نهایت یک خروجی تولید می‌کند.

معیارهای ارزیابی عملکرد مدل‌ها

تناسب مقادیر حاصل از یک مدل با مقادیر مشاهده شده (که از آن با عنوان ارزیابی یا اعتبارسنجی مدل یاد می‌شود)، معمولاً توسط مقایسه‌های دوگانه بین مقادیر شبیه‌سازی شده با مقادیر مشاهده شده صورت می‌گیرد. شاخص‌های مختلفی برای کارایی مدل‌ها و اعتبارسنجی آن‌ها وجود دارد که می‌توان از معیارهای ضریب تعیین (R^2)، مجذور میانگین مربعات خطا (RMSE)، میانگین مطلق خطا (MAE) نام برد.

R^2 ، نشان‌دهنده‌ی این است که چه مقدار از تغییرات متغیر وابسته تحت‌تأثیر متغیر مستقل مربوطه بوده و مابقی تغییرات متغیر وابسته، مربوط به سایر عوامل می‌باشد. این شاخص اگرچه دارای محدودیت‌هایی در زمینه‌ی بیان عملکرد مدل می‌باشد؛ اما درجه‌ی همبستگی میان مقادیر مشاهده شده و پیش‌بینی شده را نشان می‌دهد (۲۷). ضریب تعیین بین صفر تا یک تغییر می‌کند. مقدار صفر نشان می‌دهد که مدل هیچ یک از تغییرپذیری داده‌های پاسخ در اطراف میانگین آن را تعیین نمی‌کند. اما مقدار یک نشان می‌دهد که تمام مقادیر مشاهده شده با مقادیر برازش شده یکسان خواهند بود و همه‌ی نقاط داده‌ها بر روی خط برازش شده قرار خواهند گرفت.

$$R_{\text{sqr}} = 1 - \frac{\sum_{i=1}^n (Q_i^o - Q_i^m)^2}{\sum_{i=1}^n (Q_i^o - \bar{Q}^o)^2} \quad \text{رابطه‌ی (۲)}$$

RMSE، تفاوت میان مقدار پیش‌بینی شده توسط مدل و مقدار واقعی را نشان می‌دهد و یک ابزار خوب برای مقایسه

خطاهای پیش‌بینی توسط یک مجموعه داده است. این شاخص خطای پیش‌بینی را به‌وسیله‌ی وزن‌دهی به خطاهای بیش از حد بزرگ مورد محاسبه قرار می‌دهد. رابطه RMSE به شرح زیر می‌باشد:

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (\hat{Q}_i - Q_i)^2}{n}} \quad \text{رابطه‌ی (۳)}$$

MAE، شاخصی است که در حالت ایده‌آل باید صفر باشد. مقادیر مثبت و منفی این شاخص به ترتیب نشان‌دهنده بیش‌برآورد و کم‌برآورد مدل هستند این پارامتر معرف دقت روش و مقدار متوسط خطا است (۲۸).

$$MSE = \frac{1}{n} \sum_{i=1}^n |(Q_i^m - Q_i^o)| \quad \text{رابطه‌ی (۴)}$$

که در این روابط، Q_o داده‌های مشاهداتی، $\bar{Q}_o$ میانگین داده‌های مشاهداتی، Q_m داده‌های شبیه‌سازی شده می‌باشد. داده‌ها به دو بخش مجموعه آموزش (Training set) و مجموعه آزمون (Testing set) تقسیم شدند. برای اطمینان از آموزش مؤثر مدل و ارزیابی دقیق عملکرد آن، ۷۰ درصد از داده‌ها برای آموزش و ۳۰ درصد باقی‌مانده برای آزمون مدل‌ها در نظر گرفته شد. با توجه به ماهیت زمانی داده‌ها (سری زمانی روزانه)، این تقسیم‌بندی به‌صورت پیوسته زمانی انجام شد تا از نشت اطلاعات جلوگیری شود. به‌عبارت دیگر، داده‌های ابتدای دوره برای آموزش و داده‌های انتهایی دوره برای آزمون استفاده شدند.

یافته‌ها

آزمون همبستگی

برقراری رابطه‌ی همبستگی بین مقدار تابش خورشید و پارامترهای موردنظر با استفاده از نرم‌افزار SPSS.26 در سال آماری ۱۴۰۲ انجام گرفته است. نتایج همبستگی پیرسون با پارامترهای هواشناسی نشان داد که تابش خورشیدی با کمینه و بیشینه دما، حداکثر سرعت باد و PM_{10} دارای همبستگی مستقیم (با افزایش تابش خورشیدی مقادیر این پارامترها نیز افزایش می‌یابد و بر عکس با کاهش مقدار تابش خورشیدی این مقادیر هم کاهش پیدا می‌کنند) و با بیشینه و کمینه رطوبت نسبی، جهت باد و $PM_{2.5}$ دارای همبستگی معکوس می‌باشد. ستون P-Values (مقدار احتمال)، سطح معنی‌داری بین

همبستگی بالای، ۰/۶ همبستگی معنی دار بین پدیده‌های طبیعی محسوب می‌شوند بر این اساس مقدار تابش خورشیدی با کمینه و بیشینه دما رابطه همبستگی مناسب و مستقیم دارد و با بیشینه و کمینه رطوبت نسبی رابطه همبستگی منفی بالایی دارد.

نتایج کاربرد مدل رگرسیون چند متغیره

روش رگرسیون Enter

به منظور بررسی تأثیر معنادار متغیرهای مستقل (حداقل و حداکثر دما، بیشینه و کمینه رطوبت نسبی، سرعت باد، غلظت روزانه ذرات معلق PM_{10} و $PM_{2.5}$) بر متغیر وابسته (تابش خورشیدی) از رگرسیون استفاده شد. بر اساس نتایج رگرسیون خطی، به پارامترهای جوی وابسته است (نمودار ۱).

بر اساس این روش رگرسیون مقدار متغیر وابسته طبق متغیرهای مستقل شبیه سازی شده و با مقدار مشاهداتی ایستگاه، مقایسه می گردد (نمودار ۲) این مقایسه را به وضوح نشان می دهد.

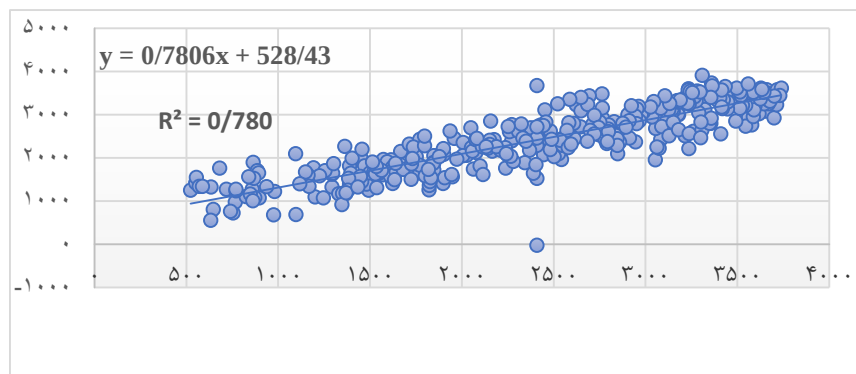
متغیرها رو نمایش می‌دهد، مقدار پارامترهای سرعت حداکثر باد، بیشینه و کمینه رطوبت نسبی، بیشینه و کمینه دما و $PM_{2.5}$ دارای حداکثر مقدار سطح معنی‌داری می‌باشند به این دلیل که مقدار احتمال آن از ۰/۰۱ کمتر می‌باشد (در سطح یک و پنج درصد معنی‌دار می‌باشند). متغیر PM_{10} معنی‌دار نمی‌باشد (در هیچ سطحی). تک‌ستاره به معنی معنی‌دار بودن در سطح ۵ درصد و دو ستاره بودن به معنی معنی‌دار بودن در هر دو سطح یک و پنج درصد می‌باشد (جدول ۱).

جدول ۱. همبستگی پیرسن بین مقدار تابش خورشیدی و پارامترهای دیگر

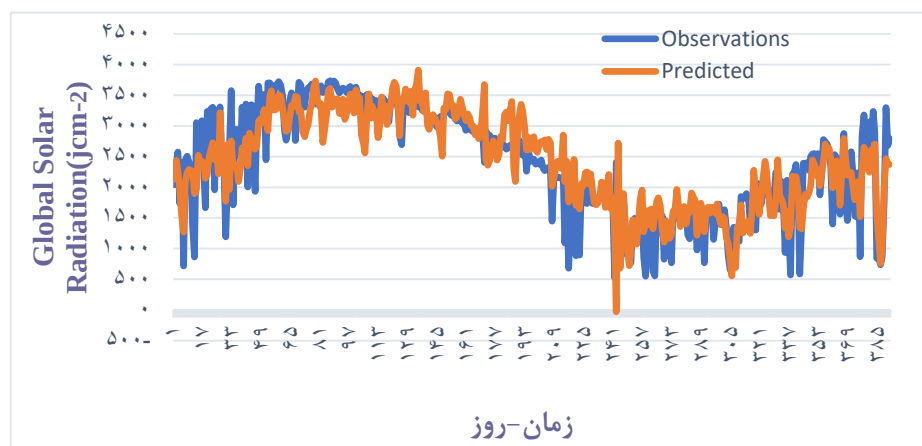
تعداد	P-Values	ضریب همبستگی پیرسن	پارامترهای هواشناسی
۳۹۱	۰/۰۰۰	۰/۱۸۸**	سرعت حداکثر باد
۳۹۱	۰/۰۰۰	۰-/۶۶۱**	بیشینه رطوبت نسبی
۳۹۱	۰/۰۰۰	-۰/۶۸۱**	کمینه رطوبت نسبی
۳۹۱	۰/۰۰۰	۰/۸۱۴**	بیشینه دما
۳۹۱	۰/۰۰۰	۰/۷۳۶**	کمینه دما
۳۹۱	۰/۰۰۰	-۰/۳۸۶**	$PM_{2.5}$
۳۹۱	۰,۰۸۳	۰/۰۸۸	PM_{10}

** : معنی‌داری در سطح اطمینان ۹۹ درصد

بسته به تعداد داده‌های آماری و درجه آزادی معمولاً



نمودار ۱. نمودار پراکنش مقادیر شبیه‌سازی و مشاهداتی میزان تابش خورشید به روش ENTER



نمودار ۲. نمودار مقایسه تغییرات مقادیر تابش مشاهده‌ای و پیش‌بینی شده با مدل ENTER

در (جدول ۳)، برآورد ضرایب و خصوصیات مربوط به آزمون آن‌ها دیده می‌شود. همان‌طور که مشاهده می‌گردد، مقدار ثابت در مدل با مقدار ۱۳۳۶/۳۴ ظاهر شده است. ستون Unstandardize Coefficients که ضرایب واقعی را نشان می‌دهد با توجه به واحد اندازه‌گیری هر یک از متغیرها ایجاد شده‌اند بنابراین نمی‌توان بر اساس بزرگی هر یک از ضرایب اهمیت متغیر مربوطه در مدل رگرسیونی را تشخیص داد. به این منظور از ستون Standardize Coefficients Beta استفاده می‌کنیم. هر ضریبی که دارای Beta بزرگ‌تری باشد، در مدل رگرسیونی از اهمیت بیشتری نیز برخوردار است. به این ترتیب مشخص می‌شود که متغیر بیشینه دما ($Beta = 1/73$) بهترین متغیر برای پیش‌گویی متغیر وابسته است (جدول ۳).

از آنجایی که ضریب همبستگی ($R = 0.88$) و ضریب تعیین ($R \text{ Square} = 0.78$) و هم‌چنین (ضریب تعیین اصلاح‌شده) ($\text{Adjusted } R \text{ Square} = 0.77$) محاسبه شده است، به نظر می‌رسد که مدل رگرسیونی مناسب است. هر چه این مقادارها به ۱ نزدیک‌تر باشند، مدل بیانگر رابطه بیشتری بین متغیر وابسته و مستقل است. به بیان دیگر مدل رگرسیونی توانسته درصد بیشتری از تغییرات متغیر وابسته را تحت پوشش قرار داده یا بیان کند. توجه به نتایج (جدول ۲) و ضریب تشخیص به دست آمده در این مدل مشاهده می‌گردد ۸۸ درصد از تغییرات مقدار تابش خورشید به وسیله پارامترهای دیگر در نظر گرفته شده، در این مدل تشریح می‌گردد.

$$\text{Ener} = (2/01) - RH_{\max} (9/40) - RH_{\min} (140/02) + T_{\max} (92/85) + T_{\min} (1336/34) = \text{معادله خطی به روش Ener} \\ \text{رابطه‌ی (۵)} \quad (10/173) + WS_{\max} (11/14) + PM_{10} (8/768) - PM_{2.5}$$

جدول ۲. خلاصه مقادیر به دست آمده از مدل (Model Summary)

مدل	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
۱	۰/۸۸	۰/۷۸	۰/۷۷	۴۰۷/۸۲	۱/۷۴

جدول ۳. ضرایب همبستگی خطی در مدل Enter

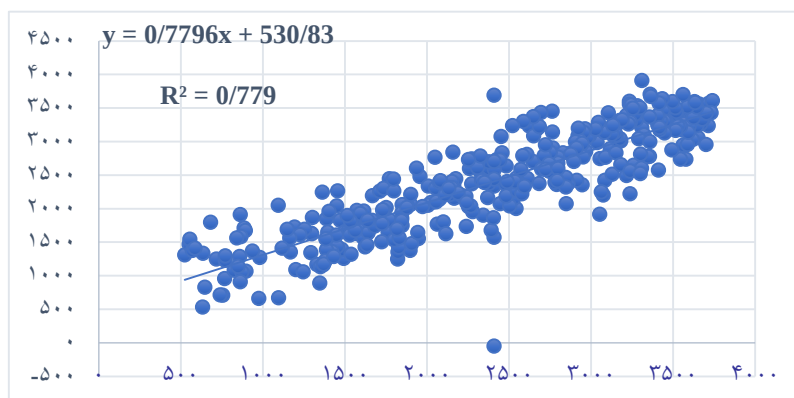
Sig	VIF	آزمون t	ضرایب پیش‌بینی شده در حالت غیراستاندارد			ضریب ثابت (C)
			Beta	Std. Error	ضرایب B	
۰/۰۰۰		۷/۲۳۳		۱۹۲,۶۳	۱۳۳۶,۳۴	

PM 10	-۱۱/۱۴	۱/۴۳	-۰/۳۹۶	-۷/۷۴۴	۱/۰۰۹	۰/۰۰۰
PM 2.5	۸/۷۶	۲/۵۹	۰/۱۷۷	۳/۳۷۸	۱/۵۳۷	۰/۰۰۱
کمینه دما	-۹۲/۸۵	۱۰/۰۷	-۰/۹۳۲	-۹/۲۱۹	۱/۶۹۶	۰/۰۰۰
بیشینه دما	۱۴۰/۰۲	۸/۸۱	۱/۷۳۷	۱۵/۸۸	۱/۸۹	۰/۰۰۰
کمینه رطوبت نسبی	-۹/۴۰	۲/۵۶	-۰/۱۶۳	-۳/۶۶	۲/۱۴	۰/۰۰۰
بیشینه رطوبت نسبی	-۲/۰۱	۱/۸	-۰/۰۵۵	-۱/۱۱	۲/۱۵	۰/۰۰۰
سرعت حداکثر باد	۱۰/۱۷	۹/۹۵	۰/۰۳	۱/۰۲	۱/۱۳	۰/۰۰۰

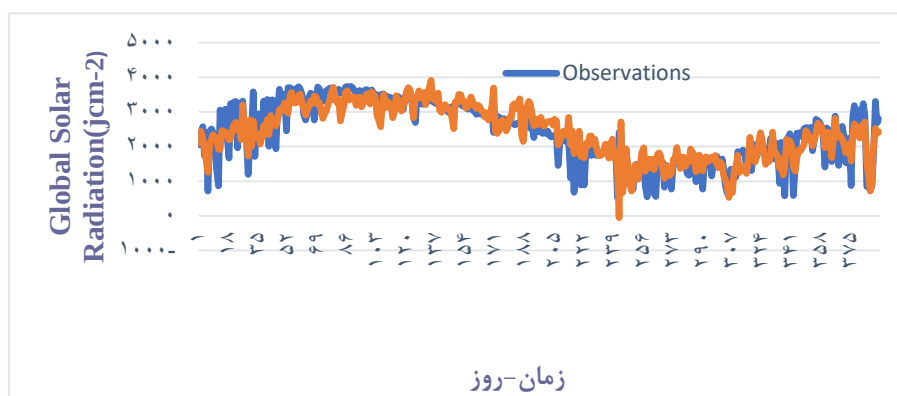
نتایج آزمون رگرسیون گام به گام پیشرونده (Stepwise)

در روش گام به گام از همه متغیرهای مستقل استفاده می‌کند؛ اما متغیرهایی با حذف گام به گام متغیرهای مستقل بهترین ترکیب از متغیرها را انتخاب می‌کند. در این روش صرفاً شبیه سازی متغیر وابسته براساس تمامی متغیرهای مستقل انجام نمی‌گیرد بلکه بهترین ترکیب از

متغیرهای مستقل رو برمی‌گزیند و مقدار تابش مشاهداتی با مقدار پیش بینی حاصل از بهترین ترکیب خطی مقایسه می‌گردد نتایج این روش در (نمودارهای ۳ و ۴) مشخص می‌باشد.



نمودار ۳. نمودار پراکنش مقادیر شبیه سازی و مشاهداتی میزان تابش خورشید به روش Stepwise



نمودار ۴. نمودار مقایسه تغییرات مقادیر تابش مشاهداتی و پیش‌بینی شده با مدل Stepwise

(مقدار تابش خورشید) عبارت است از: بیشینه دما، کمینه دما، PM_{2.5}، PM₁₀ و کمینه رطوبت نسبی؛ که دارای

این تحقیق همان‌طور که از (جدول ۴) مشخص است بهترین ترکیب از متغیرهای مستقل در مقابل متغیر وابسته

بالاترین ضریب همبستگی ($R=0.883$) و ضریب تعیین ($R^2=0.78$) و هم‌چنین "ضریب تعیین اصلاح‌شده" ($Adjusted R^2=0.777$).

نتایج (جدول ۴) و با توجه به ضریب تشخیص در این مدل مشاهده می‌گردد که حدود ۷۸ درصد از تغییرات مقدار تابش خورشید به وسیله پارامترهای مستقل (بیشینه و کمینه دما، $PM_{2.5}$ ، PM_{10} و کمینه رطوبت نسبی) در نظر گرفته‌شده، در این مدل تشریح می‌گردد. ترکیبات در نظر گرفته شده در روش رگرسیون Stepwise به شرح زیر می‌باشد (جدول ۵) و (نمودار ۳):

در (جدول ۶)، برآورد ضرایب و خصوصیات مربوط به آزمون آن‌ها دیده می‌شود. همان‌طور که مشاهده می‌گردد، مقدار

ثابت در مدل با مقدار $1264.009/0.09$ ظاهر شده است. ستون Unstandardize Coefficients که ضرایب واقعی را نشان می‌دهد با توجه به واحد اندازه‌گیری هر یک از متغیرها ایجاد شده‌اند بنابراین نمی‌توان بر اساس بزرگی هر یک از ضرایب اهمیت متغیر مربوطه در مدل رگرسیونی را تشخیص داد. به این منظور از ستون Standardize Coefficients Beta استفاده می‌کنیم. هر ضریبی که دارای Beta بزرگ‌تری باشد، در مدل رگرسیونی از اهمیت بیشتری نیز برخوردار است. به این ترتیب مشخص می‌شود که متغیر مستقل بیشینه دما ($Beta=1/74$) در این مدل هم بهترین متغیر برای پیش‌گویی متغیر وابسته است.

$$\text{Stepwise} \quad RH_{\min} - (10/782) - PM_{10} - T_{\min} (90.838) - T_{\max} (140,238) = 1264.009 + \text{خطی به روش Stepwise} \\ (11/07) + PM_{2.5} (8/768)$$

(۶)

جدول ۴. خلاصه مقادیر به دست آمده از مدل (Model Summary)

ردیف	R	R Square	Adjusted R Square	Std. Error of the Estimate	Dubin-Watson
۱	۰/۰۸۱۴ ^a	۰/۶۶۳	۰/۶۶۲	۵۰۱/۷۲	-
۲	۰/۸۴۴ ^b	۰/۷۱۲	۰/۷۱۰	۴۶۴/۴۱	-
۳	۰/۸۷۳ ^c	۰/۷۶۳	۰/۷۶۱	۴۲۱/۷۱	-
۴	۰/۸۸ ^d	۰/۷۷۴	۰/۷۷۱	۴۱۲/۷۲	-
۵	۰/۸۸۳ ^e	۰/۷۸	۰/۷۷۷	۴۰۷/۶۶	۱/۷۷۶

جدول ۵. ترکیبات مختلف مورد استفاده در روش گام به گام

علامت اختصاری	ساختار ورودی	خروجی
a	$T_{\max}$	R_s
b	$T_{\max}, T_{\min}$	R_s
c	$T_{\max}, T_{\min}, PM_{10}$	R_s
d	$T_{\max}, T_{\min}, PM_{10}, RH_{\min}$	R_s
e	$T_{\max}, T_{\min}, PM_{10}, RH_{\min}, PM_{2.5}$	R_s

جدول ۶. ضرایب همبستگی خطی در مدل Stepwise

کمیت‌های رگرسیونی	ضرایب پیش‌بینی‌شده در حالت غیراستاندارد			ضرایب استاندارد	آزمون آماره t	VIF	Sig
	ضرایب B	Std. Error	Beta				
ضریب ثابت	۱۲۶۴/۰۰۹	۱۲۶/۷۱			۹/۹۷		۰/۰۰۰
بیشینه دما	۱۴۰/۲۳۸	۸/۷۶	۱/۷۴		۱۶,۰۰	۲/۰۱۱	۰/۰۰۰
کمینه دما	-۹۰/۸۳۸	۹/۹۸	-۰/۹۱۲		-۹/۱۰۱	۱/۸۴۱	۰/۰۰۰
PM_{10}	-۱۰/۷۸۲	۱/۴۱	-۰/۳۸۳		-۷/۶۴۸	۱/۷۹۱	۰/۰۰۰
کمینه رطوبت نسبی	-۱۱/۰۷۰	۲/۱۴	۰-/۱۹۲		-۵/۱۵۶	۱/۴۶۶	۰/۰۰۰

PM2.5	۸/۰۶	۲/۴۷	۰/۱۶۳	۳/۲۵۹	۱/۶۷۴	۰/۰۰۱
-------	------	------	-------	-------	-------	-------

اعتبارسنجی مدل و بهینه سازی پارامترهای شبکه عصبی مصنوعی

برای ارزیابی دقت مدل ANN، از داده های آزمون (۳۰ درصد انتهایی سری زمانی) استفاده شد که مدل پیش از آن آموزش ندیده بود. معیارهای آماری زیر برای ارزیابی عملکرد مدل به کار گرفته شدند.

- ضریب تعیین (R^2)
- جذر میانگین مربعات خطا (RMSE)
- میانگین قدرمطلق خطا (MAE)
- میانگین مربعات خطا (MSE)

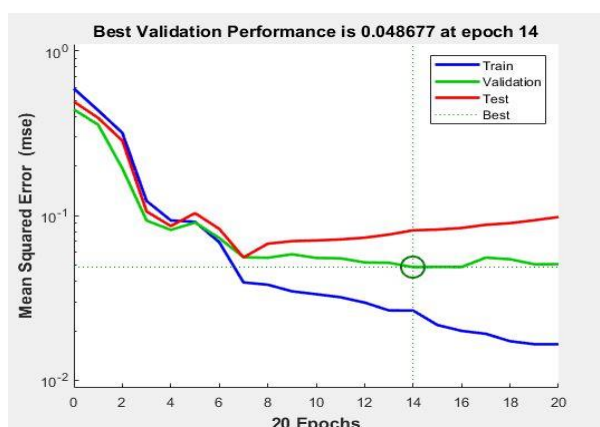
مقدار این شاخص ها برای مجموعه آزمون محاسبه شده و با نتایج روش آماری (مثلاً رگرسیون چندمتغیره) مقایسه گردید تا برتری نسبی مدل شبکه عصبی بررسی شود (جدول ۷).

در این روش، متغیرهای RH_{min} ، T_{min} ، T_{max} ، PM_{10} ، $PM_{2.5}$ و WS_{max} به عنوان ورودی به مدل شبکه های عصبی مصنوعی با استفاده از نرم افزار MATLAB بررسی می شوند. همچنین متغیر تابش کلی خورشید R_s به عنوان خروجی در نظر گرفته شد. مدل

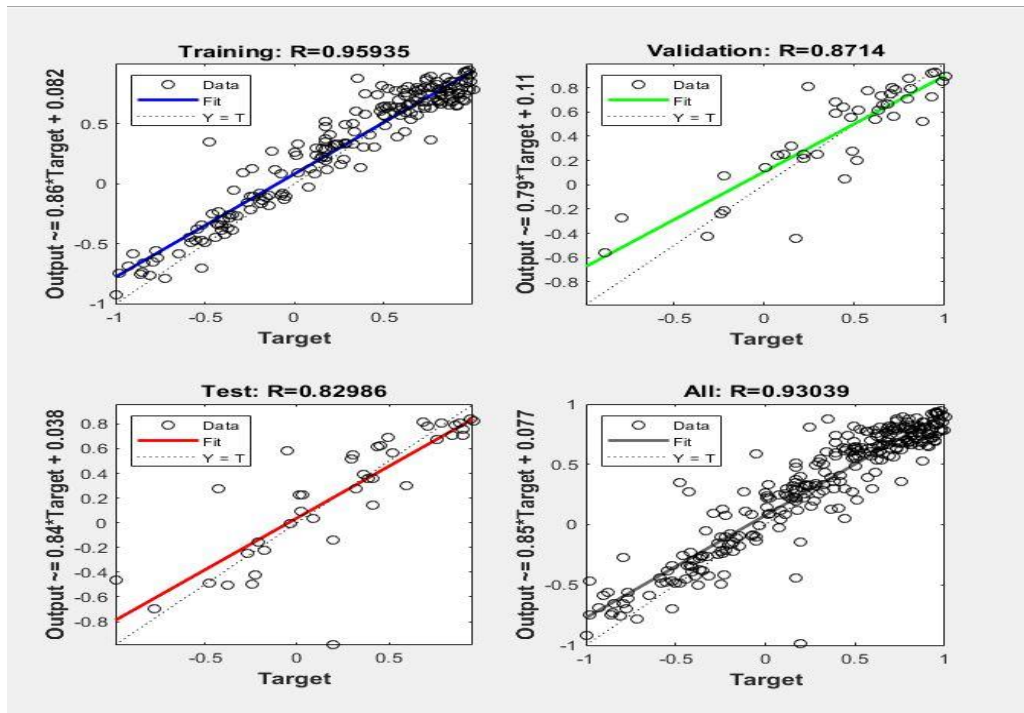
مورد استفاده در شبکه عصبی، مدل پرسپترون چند لایه (MLP) با دولایه میانی (پنهان) و تابع محرک تانژانت سیگموئید می باشد حداکثر تعداد تکرار مدل نیز ۲۰ در نظر گرفته شده که در تکرار ۱۴ طبق (نمودار ۵) به مقدار بهینه رسید.

بعد از شبیه سازی مقدار تابش خورشیدی به روش رگرسیون خطی (Enter و Stepwise) اینکار به با استفاده از شبکه عصبی مصنوعی هم انجام گرفت تا مقایسه جامع تر و همچنین نتیجه گیری مناسب تری حاصل گردد. در روش شبکه عصبی مصنوعی (ANN) ۷ متغیر مستقل روش به عنوان ورودی با ترکیب ۳۰ درصد آموزش و ۷۰ درصد آزمون، به مدل داده شد (نمودارهای ۶ و ۷) نیز مقدار این پیش بینی ها و نتایج را نشان می دهد.

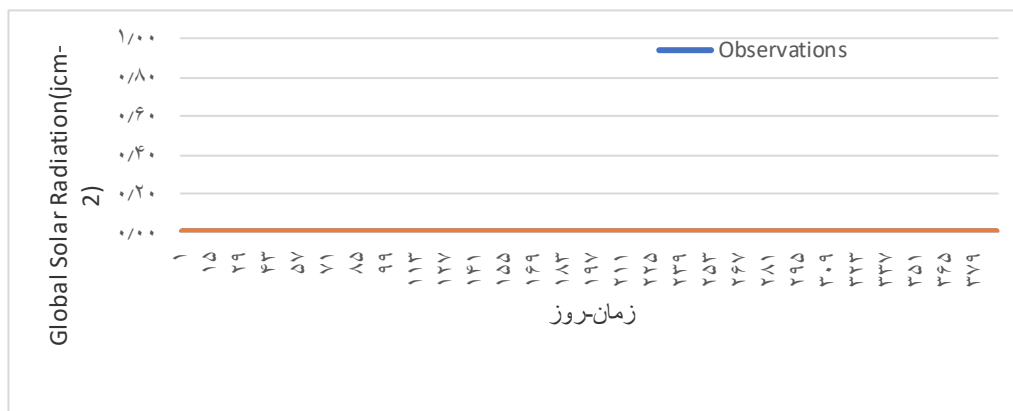
نتایج شبیه سازی با شبکه عصبی مصنوعی نشان از عملکرد بهتر این مدل در پیش بینی مقدار تابش خورشیدی دارد با ضریب R^2 حدود ۸۷ درصد (روش های رگرسیون خطی هر دو حدود ۷۸ درصد) (نمودار ۸) نیز در نقایسه با (نمودارهای ۱ و ۳) این مطلب را اثبات می کند.



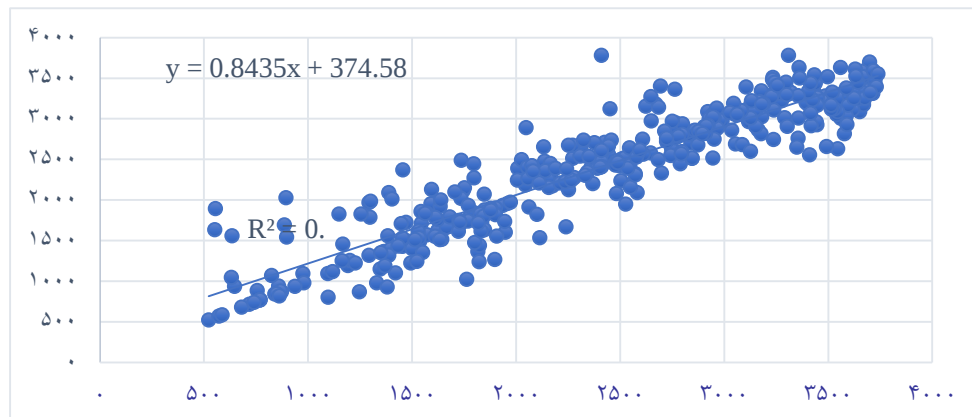
نمودار ۵. نمودار تکرار بهینه در شبکه عصبی مصنوعی



نمودار ۶. نمودار نتایج شبیه سازی در مرحله آزمون و آموزش



نمودار ۷. نتایج مقایسه بین مدل پیش بینی ANN و نتایج واقعی



نمودار ۸. ارتباط بین مقدار شبیه سازی شده با ANN و مقادیر واقعی

جدول ۷. ارزیابی مدل های مختلف استفاده شده در مرحله Train و Test

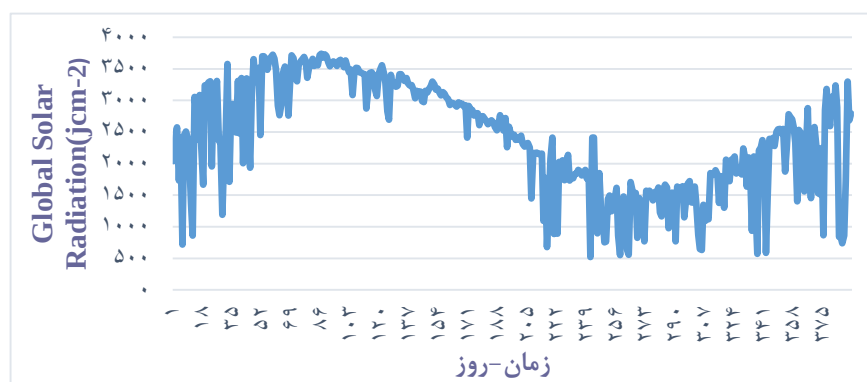
آزمون				آموزش				مدل
MSE	MAE	RMSE	R ²	MSE	MAE	RMSE	R ²	
۳/۳۷	۱۱۸/۲۸	۱۸۳/۸۴	۰/۶۹	۳/۶۵	-۴۶/۹۷	۱۹۱/۰۹۷	۰/۷۱	رگرسیون ENTER
۱/۷۵	۱۱۸/۸۱	۱۸۴/۶۷	۰/۶۷	۱۱/۶	-۴۷/۱۹	۱۹۱/۹۷	۰/۷۲	رگرسیون STEPWISE
۱/۶۸	۸۲/۰۲	۱۲۹/۶۳	۰/۸۸	۱/۸۷	-۳۵/۸۶	۱۳۶/۹۰۸	۰/۸۱	شبکه عصبی ANN

گرفته است. (نمودار ۹) مربوط به مقدار تابش خورشیدی در سال آماری ۱۴۰۲-۱۴۰۳ می باشد که حداکثر آن ۳۷۳۹ (ژول بر سانتی متر مربع) در ۲۴ خرداد ماه ۱۴۰۲ و کمترین مقدار آن ۵۲۳ (ژول بر سانتی متر مربع) در ۲۵ آبان همان سال می باشد. (نمودار ۱۰) بیانگر مقدار غلظت آلاینده های PM_{10} و $PM_{2.5}$ در سال آماری مورد نظر بر حسب $\mu g/m^3$ (میکروگرم بر مترمکعب) می باشد که مقدار حداکثر و حداقل برای PM_{10} به ترتیب ۱۰۰/۳۸ و ۹۹/۸۷ و برای $PM_{2.5}$ ۹۸/۹۳ و ۱۰/۳۶ می باشد.

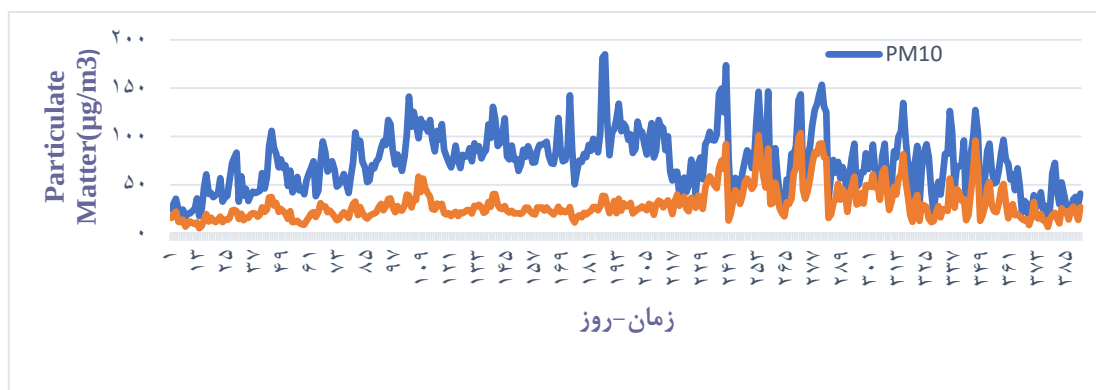
محدودیت های مطالعه: در اختیار نداشتن همزمان داده های تابش خورشیدی با سایر پارامترها به صورت همزمان در سایر ایستگاه های شهر تهران. پیشنهاد می گردد تحقیق موردنظر برای اقلیم های مختلف کشور با سایر روش ها نظیر مدل های ماشین لرنینگ، انجام و نتایج آن در برنامه ریزی های مد نظر قرار گیرد.

بحث

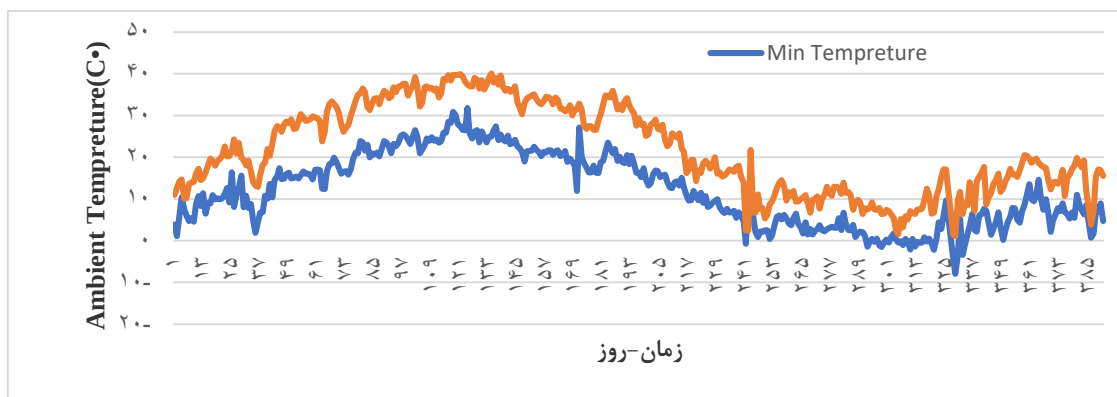
هدف از مطالعه ی حاضر پیش بینی مقدار تابش خورشیدی بر اساس پارامترهای هواشناسی و آلودگی هوا در ایستگاه سینوپتیک تهران می باشد که با استفاده از روش های رگرسیون خطی چندگانه و شبکه عصبی مصنوعی انجام



نمودار ۹. مقدار تابش خورشیدی در سال آماری



نمودار ۱۰. مقدار ذرات معلق هوا PM₁₀ و PM_{2.5} در روزهای مختلف سال

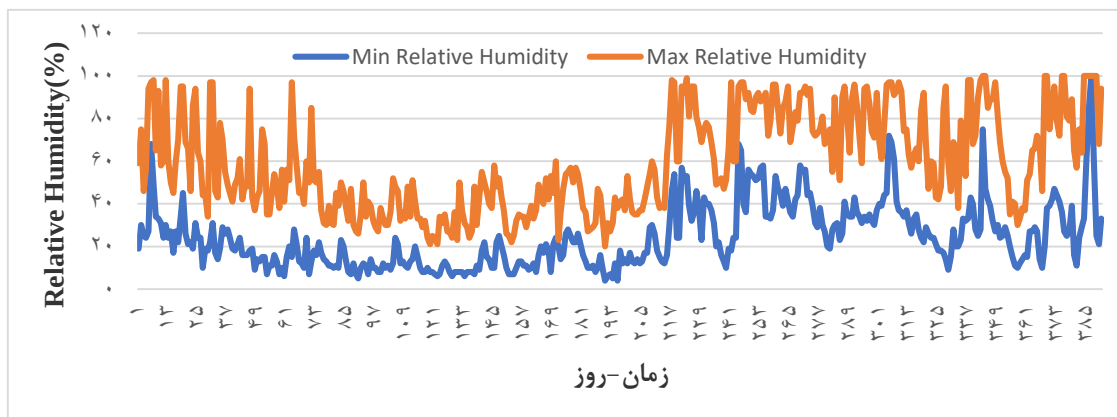


نمودار ۱۱. حداقل و حداکثر دما در روزهای مختلف سال

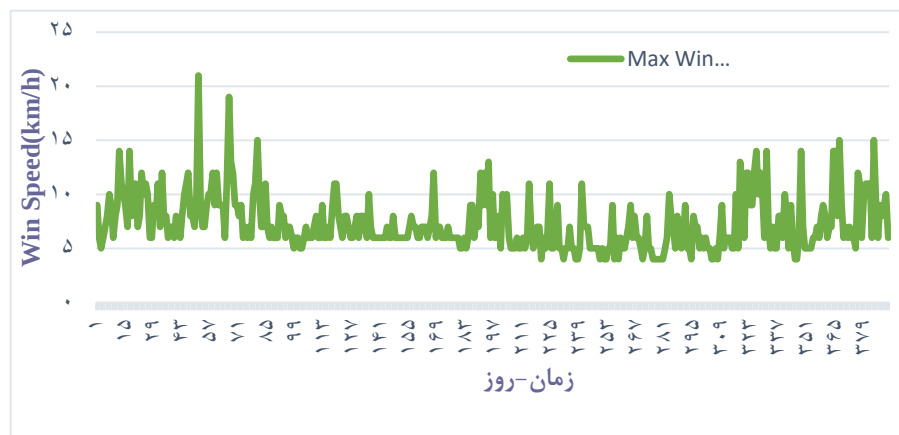
گردید. به منظور ارزیابی میزان همبستگی بین متغیرها از ضریب همبستگی پیرسون استفاده شد. ضریب همبستگی پیرسون رایج ترین همبستگی مورد استفاده است. این ضریب میزان همبستگی بین دو متغیر تصادفی را می سنجد. مقدار این همبستگی بین -۱ تا ۱ تغییر می کند که مقدار ۱ به معنای همبستگی کامل مثبت یا مستقیم، صفر به معنای عدم همبستگی و -۱ به معنای همبستگی کامل منفی یا معکوس است.

حداقل و حداکثر دما در سال موردنظر طبق (نمودار ۱۱) به ترتیب ۹/۹۰، ۰/۱- و ۹/۸۰، ۱/۱۰ سانتی گراد به ثبت رسیده است.

(نمودارهای ۱۲ و ۱۳) نیز به ترتیب تغییرات حداقل و حداکثر رطوبت نسبی و حداکثر سرعت باد در ایستگاه موردنظر را نشان می دهند. برای شناخت مهم ترین متغیرهای هواشناسی و سنجش آلودگی (متغیرهای ورودی) مؤثر بر مقدار شدت تابش خورشید، مقدار ضریب همبستگی بین شدت تابش خورشیدی با هر یک از متغیرهای محاسبه



نمودار ۱۲. حداقل و حداکثر رطوبت نسبی در روزهای مختلف سال



نمودار ۱۳. حداکثر سرعت باد در روزهای مختلف سال

نتیجه‌گیری

تست مدل‌ها نشان می‌دهد بهترین مدل از نظر این شاخص در مرحله آموزش ابتدا شبکه عصبی (با مقدار $MSE=1/87$)، سپس روش Enter (با مقدار $MSE=3/65$) و در نهایت روش Stepwise (با مقدار $MSE=1/6$) می‌باشد. مقایسه مدل‌ها در مرحله تست به ترتیب شبکه عصبی، Stepwise و Enter با مقادیر میانگین مربعات خطای (MSE) به ترتیب $1/68$ ، $1/75$ و $3/37$ مگا ژول بر سانتی مترمربع می‌باشد.

مقایسه‌ی پژوهش حاضر با سایر تحقیقات انجام گرفته در این زمینه (پیش‌بینی تابش خورشیدی) نشان می‌دهد که اولاً معیارهای ارزیابی روش‌های مختلف متفاوت بوده و ثانیاً به جهت این‌که مناطق مختلف از نظر دریافت تابش خورشیدی با هم فرق دارند، در هر تحقیق الگوریتم متفاوتی دارای عملکرد بهینه بوده است. برای مثال، موزنو و همکاران (۲۰۲۱) در اسپانیا روش شبکه عصبی مصنوعی را در مقایسه با یک روش تجربی و رگرسیون کرنل روش برتر دانستند که با نتایج کلی تحقیق حاضر مطابقت دارد. محمدی و همکاران (۲۰۱۹) از بین روش‌های ماشین بردار پشتیبان، شبکه عصبی مصنوعی و الگوریتم ژنتیک در تخمین تابش خورشیدی روزانه و ماهانه در یک شهر ساحلی ایران، روش ماشین بردار پشتیبان را روش مناسب‌تری پیشنهاد دادند.

کوئچ و همکاران (۲۰۲۲) روش ماشین بردار پشتیبان را نسبت به روش‌های انفیس و شبکه عصبی مصنوعی در تخمین تابش خورشیدی در یک منطقه گرم نیمه‌مرطوب

امروزه با توسعه‌ی روش‌های هوش مصنوعی در برآورد پارامترهای گوناگون، مدل‌هایی هم‌چون شبکه‌های عصبی مصنوعی قابلیت خود را در پیش‌بینی این پارامترها به منزله‌ی ابزاری قدرتمند به خوبی نشان داده‌اند. در این تحقیق مدل شبکه عصبی مصنوعی در برآورد تابش کل خورشیدی برای ایستگاه سینوپتیک تهران به کار گرفته شد. بدین منظور از مجموعه‌ای از پارامترهای اقلیمی و کیفیت هوا موثر در برآورد تابش خورشیدی در حکم ورودی و تعیین ساختار شبکه استفاده شد. پارامترهای فوق از جمله پارامترهایی هستند که از سوی سازمان هواشناسی و سازمان کنترل ترافیک هوای شهر تهران به صورت روز آن‌ها اندازه‌گیری و ثبت می‌شوند مقایسه مدل‌های به کار رفته در پژوهش نشان می‌دهد که روش شبکه عصبی مصنوعی MLP نسبت به دو روش دیگر (روش‌های رگرسیون خطی چندگانه) به عملکرد بهتری از خود نشان می‌دهد (ضریب رگرسیون $86/56$ درصد و بالاتر از روش‌های دیگر) و روش‌های رگرسیون خطی تقریباً عملکرد مشابهی دارند (ضریب رگرسیون تقریباً برابر). مقدار شاخص MSE (میانگین مربعات خطا) در هر دو مرحله آموزش و آزمون در روش شبکه عصبی بهتر و به ترتیب $1/87$ و $1/68$ می‌باشد در حالی که همین شاخص برای روش Enter، $3/65$ و $3/37$ بوده و همچنین برای روش Stepwise نیز مقادیر $1/75$ و $1/68$ محاسبه گردیده است همچنین مقایسه مقدار شاخص MSE (میانگین مربعات خطا) در مرحله آموزش و

تشکر و قدردانی: بدین وسیله از مسئولین محترم سازمان هواشناسی تهران، به دلیل در اختیار گذاشتن داده‌های پژوهش، صمیمانه تشکر و قدردانی می‌گردد.

تعارض منافع: پژوهش حاضر دارای تعارض منافع نمی‌باشد.

حمایت مالی: این مقاله منتج از کار تحقیقاتی شخصی می‌باشد و برای این پژوهش هیچ‌گونه حمایت مالی از طرف هیچ نهادی دریافت نشده است.

ملاحظات اخلاقی: نویسندگان تمام نکات اخلاقی شامل عدم سرقت ادبی، انتشار دوگانه، تحریف داده‌ها و داده‌سازی را در این مقاله رعایت کرده‌اند. هم‌چنین هرگونه تضاد منافع حقیقی یا مادی که ممکن است بر نتایج یا تفسیر مقاله تأثیر بگذارد را رد می‌کنند.

سهم نویسندگان: ۱. امیر زارعی ۲. سیروان زارعی ۳. سیدعرفان مومن‌پور ۴. محمدرضا گل‌عنبریان ۵. عباس سلیمان‌پور در اجرای این پژوهش مشارکت داشته‌اند.

References

1. Layeghi, M. 2019. Solar energy, technologies and applications. Jahad daneshgahi publication. (Persian).
2. Zohoori M., (2012): Exploiting Renewable Energy Sources in Iran. Interdisciplinary J. of Contemporary Research in Business ,4, pp:849-862.
3. Solangi K.H., M.R. Islam, R. Saidur, N.A. Rahim and H. Fayaz, (2011): A review on global solar energy policy. Renewable and Sustainable Energy Reviews, 15, pp: 2149-2163. <https://doi.org/10.1016/j.rser.2011.01.007>
4. Shaddel, M., Javan, D.S., Baghermia, P., 2019. Estimation of hourly global solar irradiation on tilted absorbers from horizontal one using Artificial Neural Network for case study of Mashhad. Renew. Sustain. Energy Rev. 53, 59e67. <https://doi.org/10.1016/j.rser.2015.08.023>
5. Abdelhafidi, N., Bachari, N.E.I. and Abdelhafidi, Z., 2021. Estimation of solar radiation using stepwise multiple linear regression with principal component analysis in Algeria. Meteorology and Atmospheric Physics, 133(2), pp.205-216. and hydroclimatic model validation. Water and Resources Research,

ارجح دانستند. اگبولوت و همکاران (۲۰۲۳) در چهار منطقه‌ی آب‌وهوایی مختلف در ترکیه تابش خورشیدی را با روش‌های شبکه عصبی مصنوعی، ماشین بردار پشتیبان، نزدیک‌ترین همسایگی کرنل و یادگیری عمیق برآورد کردند. نتایج آن‌ها حاکی از مناسب بودن روش شبکه عصبی مصنوعی نسبت به سایر روش‌ها بود. تفاوت نتایج تحقیقات (عملکرد مدل‌ها) می‌تواند به دلایل شرایط آب‌وهوایی، طول دوره آماری، داده‌های گمشده و انتخاب داده‌های ورودی در الگوریتم‌های مورد استفاده در پیش‌بینی تابش خورشیدی، باشد هر چند در اکثر تحقیقات روش‌های داده‌کاوی (روش شبکه عصبی، ماشین بردار پشتیبان و...) در همه مناطق روش بهتری نسبت به سایر روش‌ها (روش‌های تجربی) در تخمین تابش خورشیدی خصوصاً در مناطقی که داده‌های تابش خورشیدی موجود نیست، پیشنهاد می‌شود.

- 35(1), 233-241 <https://doi.org/10.1007/s00703-020-00739-0>
6. Rahimikhoob A. 2020. Estimating global solar radiation using artificial neural network and air temperature data in a semi-arid environment. Renew. Energy. 35, 2131-2135. <https://doi.org/10.1016/j.renene.2010.01.029>
7. Hunt, L.A., Kuchar, L., Swanton, C.J., 2018. Estimation of solar radiation for use in crop modeling. Agric. Meteorol. 91, 293-300. [https://doi.org/10.1016/S0168-1923\(98\)00055-0](https://doi.org/10.1016/S0168-1923(98)00055-0)
8. Sabziparvar A.A., and Shetaee H. 2019. Estimation of global solar radiation in arid and semi-arid climates of East and West Iran, Energy 32: 649-655. <https://doi.org/10.1016/j.energy.2006.05.005>
9. Yadav, A. K. and Chandel, S. S. (2019). Solar energy potential assessment of western Himalayan Indian state of Himachal Pradesh using J48 algorithm of WEKA in ANN based prediction model. Renewable Energy, 75, 675-693. <https://doi.org/10.1016/j.renene.2014.10.046>
10. Mehdizadeh, S., Behmanesh, J. and Khalili, K., 2018. Comparison of artificial intelligence methods and empirical equations to estimate

daily solar radiation. *Journal of Atmospheric and Solar-Terrestrial Physics*, 146, pp.215-227.

<https://doi.org/10.1016/j.jastp.2016.06.006>

11. Okundamiya MS, Emagbetere JO, Ogijor EA (2021) Evaluation of various global solar radiation models for Nigeria. *Int J Green Energy* 13(5):505-512.

<https://doi.org/10.1080/15435075.2014.968921>

12. Olalekan, S., Abdullahi, M. I. and Olabisi, A. (2021). Modeling of Solar Radiation Using Artificial Neural Network for Renewable Energy Application. *Journal of Applied Physics*, 10(2), 6-12

13. Meenal, R. and Selvakumar, A.I., 2022. Assessment of SVM, empirical and ANN based solar radiation prediction models with most influencing input parameters. *Renewable Energy*, 121, pp.324-343.

<https://doi.org/10.1016/j.renene.2017.12.005>

14. Radosevic, N., Duckham, M., Liu, G.J. and Sun, Q., 2021. Solar radiation modeling with KNIME and Solar Analyst: Increasing environmental model reproducibility using scientific workflows. *Environmental Modelling & Software*, 132, p.104780

<https://doi.org/10.1016/j.envsoft.2020.104780>

15. Zeng J, Qiao W (2020) Short-term solar power prediction using a support vector machine. *Renewable Energy* 52:118-127.

<https://doi.org/10.1016/j.renene.2012.10.009>

16. Yadav, A. K. and Chandel, S. S. (2019). Solar energy potential assessment of western Himalayan Indian state of Himachal Pradesh using J48 algorithm of WEKA in ANN based prediction model. *Renewable Energy*, 75, 675-693.

<https://doi.org/10.1016/j.renene.2014.10.046>

17. Ball RA, Purcell LC, Carey SK (2018) Evaluation of solar radiation prediction models in North America. *Agronomy Journal* 96:391-397 <https://doi.org/10.2134/agronj2004.3910>

18. Belmahdi, B., Louzazni, M. and El Bouardi, A., 2022. One month-ahead forecasting of mean daily global solar radiation using time series models. *Optik*, 219, p.165207.

<https://doi.org/10.1016/j.ijleo.2020.165207>

19. Taki, M., Rohani, A. and Yildizhan, H., 2023. Application of machine learning for solar radiation modeling. *Theoretical and Applied Climatology*, 143(3-4), pp.1599-1613.

<https://doi.org/10.1007/s00704-020-03484-x>

20. Yang, H., Griffiths, P. R. and Tate, J. D. (2018). Comparison of partial least squares regression and multi-layer neural networks for quantification of nonlinear systems and application to gas phase Fourier transform infrared spectra. *Analytica Chimica Acta*, 489(2), 125-136. [https://doi.org/10.1016/S0003-2670\(03\)00726-8](https://doi.org/10.1016/S0003-2670(03)00726-8)

21. Tabari, H., Grismer, M. E. and Trajkovic, S. (2019). Comparative analysis of 31 reference evapotranspiration methods under humid conditions. *Irrigation Science*, 31(2), 107-117.

<https://doi.org/10.1007/s00271-011-0295-z>

22. Nwokolo, S.C., Obiwulu, A.U., Ogbulezie, J.C. and Amadi, S.O., 2022. Hybridization of statistical machine learning and numerical models for improving beam, diffuse and global solar radiation prediction. *Cleaner Engineering and Technology*, 9, p.100529.

<https://doi.org/10.1016/j.clet.2022.100529>

23. Soltani-Gerdefaramarzi, S. (2023). Prediction of solar radiation intensity in Yazd station by using regression model based on principal components (PCR). *Journal of Agricultural Meteorology*, 11(1), 6-16. doi: 10.22125/agmj.2023.352446.1140.

24. Jahantigh, N., & Piri, J. (2023). Estimation of solar radiation in different climates of Iran using hybrid machine learning methods. *Journal Of Applied and Computational Sciences in Mechanics*, 35(2), 37-54. doi: 10.22067/jacsm.2022.79378.1144.

25. Rehman, S., Mohandes, M., 2018. Artificial neural network estimation of global solar radiation using air temperature and relative humidity. *Energy Policy* 36, 571-576

<https://doi.org/10.1016/j.enpol.2007.09.033>

26. Razi, M. A. and Athappilly, K. (2021). A comparative predictive analysis of neural networks (NNs), nonlinear regression and classification and regression tree (CART) models. *Expert Systems with Applications*, 29(1), 65-74.

<https://doi.org/10.1016/j.eswa.2005.01.006>

27. Legates D. R. and McCabe Jr. G. J. (2017). Evaluating the use of goodness-of-fit measures in hydrologic

28. Mohammadi B, Emamgholizadeh S. Using principal component analysis to inputs the effective rainfall estimates based on entries to help support vector machine and artificial neural

network. Journal of Rainwater Catchment
Systems 2017; 4 (4):67-75.