

مقایسه مدل های تجربی، رگرسیونی و شبکه عصبی مصنوعی در برآورد تابش خالص دریافتی (R_s) در ایستگاه سینوپتیک زاهدان

پریسا کهخا مقدم*

عضو هیئت علمی گروه مهندسی آب دانشگاه زابل، زابل، ایران

محمد مهدی جاری

عضو هیئت علمی گروه مهندسی آب دانشگاه زابل، زابل، ایران

تاریخ پذیرش: ۱۳۹۶/۲/۹

تاریخ دریافت: ۱۳۹۵/۴/۸

چکیده

تابش خورشیدی در بسیاری از مدل های هیدرولوژی، پارامتری مهم در تخمین تبخیر و تعرق است. تهیه و ایجاد وسایل اندازه گیری این پارامتر بسیار پرهزینه می باشد. در این پژوهش از داده های اندازه گیری شده تابش (R_s) در سال های ۱۳۸۵ تا ۱۳۸۹ ایستگاه هواشناسی زاهدان استفاده شده است و همچنین چند مدل غیرخطی نظیر شبکه عصبی با الگوریتم BFGS و شبکه عصبی با کاهش شیب توأم و رگرسیون خطی محلی با استفاده از آزمون گاما توسعه داده شد. سپس این مدل های غیرخطی و دو مدل تجربی شامل آنگستروم- پرسکات و گلور مک کلوت برای پیش بینی تابش مورد ارزیابی قرار گرفت. از پارامترهای دمای ماکزیمم، سرعت متوسط باد و تابش برون زمینی و ساعت آفتابی برای پیش بینی روش های غیرخطی استفاده شد. نتایج مقایسه مقادیر محاسبه شده با مدل ها با مقادیر اندازه گیری شده توسط پیرانومتر نشان می دهد که شبکه عصبی با روند نما BFGS دارای RMSE= ۱/۹۵، MAE= ۱/۴۷ و R²= ۰/۹۳ است که دارای بهترین عملکرد در مدل ها است. بعد از آن مدل شبکه عصبی با کاهش شیب توأم و مدل رگرسیونی خطی محلی است که مقادیر RMSE، MAE و R² برای آن ها به ترتیب برابر ۲/۵۳، ۱/۷۷، ۰/۸۸ و ۲/۸۹، ۱/۸۹، ۰/۸۲ می باشد. روش آنگستروم و گلور-مک کلوت نیز به ترتیب دارای مقادیر RMSE=۴/۳۸، MAE= ۳/۲۱، R²= ۰/۳۳، RMSE=۴/۶۴، MAE= ۳/۰۷ و R²= ۰/۵۰ می باشند.

واژگان کلیدی: تابش خالص دریافتی، شبکه عصبی، گاما تست، زاهدان.

مقدمه

تابش خورشیدی در بسیاری از مدل های هیدرولوژی، پارامتری مهم در تخمین تبخیر و تعرق است. به علاوه داده های تابش آفتابی روزانه در مدل های شبیه سازی رشد گیاه و مدل های بیلان آب مهم می باشد (تان و همکاران، ۲۰۰۷: ۲۶۵-۲۶۶). با وجود اهمیت اندازه گیری این پارامتر به دلیل مشکلات اقتصادی وسایل و ابزار مناسب اندازه گیری تابش در همه

مناطق موجود نبوده و باید به‌گونه‌ای تخمین زده شود. با این وجود اطلاعات زیادی برای توسعه آن صورت نگرفته است. در دهه‌های گذشته مدل‌های فیزیکی و تجربی برای تابش پیشنهاد شده است (نایا و همکاران، ۱۹۹۳: ۴۵۲ و تاور و بالدسانو، ۲۰۰۱: ۱۸۲۴-۱۸۲۵). معادله آنگستروم^۱ که به‌وسیله Angstrom (۱۹۲۴) ارائه شد و سپس توسط پروسکات^۲ ۱۹۴۰ اصلاح شد، یک معادله رایج می‌باشد و در بیشتر روش‌ها از آن برای تخمین میانگین ماهانه تابش خورشیدی استفاده می‌شد (خلیلی و رضایی صدر ۱۳۷۶: ۱۷). قویدل حیدری (۱۳۸۵) توان بالقوه انرژی خورشیدی در استان سیستان و بلوچستان را مورد بررسی قرار داد.

اخیراً روش‌های جدیدی نظیر شبکه‌های عصبی مصنوعی برای پیش‌بینی تابش آفتابی استفاده می‌شود (داینسر و همکاران، ۱۹۹۶: ۱۸۳-۱۹۸؛ مهندس و همکاران، ۱۹۹۸: ۱۷۹-۱۸۴؛ اگولیت و همکاران، ۲۰۰۲: ۳۵۱-۳۵۸ و سایلان و همکاران، ۲۰۰۳: ۱۸۲۹-۱۸۳۷). همچنین در این روش‌ها سعی می‌شود با کمترین اطلاعات و با دقت بالایی تابش را شبیه‌سازی کرد. تامویز و همکاران (۲۰۰۵: ۷۵۲-۷۶۲) از ۴ پارامتر ورودی برای مدل‌سازی تابش خورشیدی با شبکه عصبی بهره بردند. در این کار مدل‌های گوناگون شبکه عصبی مورد ارزیابی قرار گرفته و با یافته‌های مدل آنگستروم سنجیده شده است. بوش و همکاران (۲۰۰۸: ۱۶۲۲-۱۶۲۸) با استفاده از روش شبکه عصبی مصنوعی در برآورد پارامتر تابش کل در مناطق کوهستانی اسپانیا به نتایج قابل قبولی دست یافت. نتایج آن‌ها روشن ساخت که استفاده از شبکه عصبی مصنوعی حتی در برآورد تابش در مناطقی با توپوگرافی پیچیده نیز رویکردی مناسب است. شبکه عصبی مصنوعی را برای برآورد تابش کل برای ۶ منطقه از ایران مورد بررسی قرار دادند و نتایج آن را با مدل آنگستروم مقایسه کردند. سبزی پرور و بیات ورکشی (۱۳۸۹: ۳۴۷-۳۵۷) روش شبکه عصبی و عصبی - فازی را در شبیه‌سازی تابش خورشیدی مورد بررسی قرار دادند و نتیجه گرفتند که به‌طور کلی مدل‌های هوشمند با دقت قابل قبولی توانایی پیش‌بینی را دارند. همچنین ساعات آفتابی دارای بیشترین همبستگی با تابش است. پیری و همکاران (۱۳۹۲) تابش رسیده به سطح زمین را با استفاده از روش (ANFIS) برای ایستگاه‌های زاهدان و بجنورد مدل‌سازی کردند و نشان دادند که روش نرو فازی با دقت مناسبی تابش دریافتی از خورشید را برای این دو ایستگاه مدل‌سازی می‌کند.

آزمون گاما یک مدل‌سازی غیرخطی و ابزاری جهت آنالیز است و اجازه می‌دهد تا رابطه بین ورودی و خروجی‌ها در یک مجموعه داده‌های عددی مورد امتحان قرار گیرد. این آزمون می‌تواند کمترین مقدار قابل محاسبه میانگین مربع‌ها خطا را برای هر مدل ارائه دهد. آزمون گاما مقدار خطای برآورد شده را از روی داده‌ها به‌صورت مستقیم نشان می‌دهد (جونز و همکاران، ۲۰۰۲). این آزمون نخستین بار توسط اگلباجون و همکاران (۱۹۹۷: ۱۳۳-۱۳۱) به‌صورت خلاصه معرفی شد. سپس جزئیات آن در سال‌های بعد توسط چوژانوا و همکاران (۱۹۹۸)، توسی (۱۹۹۹)، دورانت (۲۰۰۱)، توسی و همکاران (۲۰۰۲) و جونز و همکاران (۲۰۰۲) مورد بررسی قرار گرفتند. در طی سالیان اخیر پژوهش‌هایی در ارتباط با استفاده از آزمون گاما و کاربرد آن در زمینه‌های تبخیر و تعرق منتشر شده است. ریمسان و همکاران با استفاده از آزمون

^۱. Angstrom

^۲. Prescott

گاما عوامل مؤثر بر تابش خورشیدی را در انگلستان تعیین کردند. آن‌ها از ترکیب پارامترهای دما، باران، سرعت باد و تابش برون زمینی برای ساخت مدل استفاده کردند و سپس با استفاده از دو مدل رگرسیونی و شبکه عصبی به مدل سازی تابش خورشیدی پرداختند (ریمسان و همکاران، ۲۰۰۸: ۴۳۰۱-۴۳۰۹). مقدم نیا و همکاران (۲۰۰۹: ۹۷۵-۹۸۲) برای مدل سازی تابش در انگلستان از آن استفاده کردند در این پژوهش از پارامترهای دمای حداقل، حداکثر و میانگین، سرعت باد و بارندگی استفاده شد.

پیری و همکاران (۲۰۱۵) با استفاده از روش بردار پشتیبان و رگرسیونی و به کارگیری پارامترهای هواشناسی شامل ساعت‌های آفتابی، دمای حداکثر، دمای حداقل و رطوبت نسبی متوسط مقدار تشعشع رسیده به سطح زمین را برای دو ایستگاه سینوپتیک زاهدان و بجنورد و به ترتیب برای ۵ و ۷ سال آماری مدل سازی کردند. نتایج نشان داد که دقت روش بردار پشتیبان از روش رگرسیونی و همچنین روش‌های تجربی بهتر است. شمشیربند و همکاران (۲۰۱۶) با استفاده از مدل شبکه عصبی ترکیبی خودکار پرسرونده^۱ تشعشع دریافتی برای شهر بندرعباس در جنوب را پیش‌بینی کردند. برای این منظور از دو مجموعه داده شامل ۱- ساعات آفتابی و مجموع ساعت‌های آفتابی ممکن ۲- بیشینه و کمینه دمای محیط استفاده کردند. نتایج نشان داد تشعشع پیش‌بینی شده با استفاده از مجموعه اول داده‌ها شامل ساعت‌های آفتابی و مجموع ساعت‌های آفتابی ممکن دارای دقت بیشتری نسبت به استفاده از روش شماره (۲) است.

موضوع اصلی این مطالعه، ارزیابی کارایی روش‌های غیرخطی مانند رگرسیون خطی و شبکه عصبی در مقایسه با مدل‌های تجربی تخمین تابش خاص دریافتی است. در این مطالعه همچنین قابلیت آزمون گاما در ساختن مدل‌های غیرخطی بدون بعد بر اساس میانگین مربعات خطا مورد بررسی قرار می‌گیرد. همچنین داده‌های مورد نیاز برای ساخت مدل مورد ارزیابی قرار می‌گیرد.

داده‌ها و روش‌ها

منطقه مورد مطالعه

برای انجام مدل سازی تابش خالص دریافتی با استفاده از مدل‌های غیرخطی و تجربی فهرست و اطلاعات مورد نیاز از ایستگاه هواشناسی زاهدان از دی‌ماه ۱۳۸۵ تا خرداد ماه ۱۳۸۹ (از صحت این داده‌ها اطمینان حاصل شد) تهیه گردید. ایستگاه مورد نظر در عرض جغرافیایی ۲۸' ۲۹° شمالی و طول جغرافیایی ۵۳' ۶۰° شرقی قرار گرفته است. ارتفاع آن از سطح دریا ۱۳۶۹ متر می‌باشد. زاهدان بر اساس طبقه‌بندی اقلیمی کوپن جزء مناطق بیابانی محسوب می‌شود.

^۱. Hybrid auto-regressive neural network model

آزمون گاما

پایه اصلی روش آزمون گاما به‌طور کلی تلاش برای زودتر به نتیجه رسیدن آنالیز غیرخطی است. برای مثال مجموعه‌ی زیر را در نظر بگیرید،

$$\{(x_i, y_i), 1 \leq i \leq M\} \quad \text{رابطه ۱}$$

که در آن بردار $x \in R^m$ به‌عنوان ورودی و مطابق آن عدد $y \in R$ ، خروجی و M تعداد سری مجموعه‌ی مشاهده شده از پدیده مورد بررسی می‌باشد. آزمون گاما با فرض عدم قطعیت و پذیرش مقدار خطا در به دست آوردن خروجی از روی ورودی به خاطر پیچیدگی و غیرخطی بودن پدیده‌ها مدل‌سازی این خطا را به‌صورت زیر نشان می‌دهد:

$$y = f(x_1, \dots, x_M) + r \quad \text{رابطه ۲}$$

که در اینجا f یک مدل هموار^۱ است که برای مدل‌سازی داده‌ها استفاده می‌شود و Γ نشان‌دهنده متغیر تصادفی است که برای نمایش خطا به کار می‌رود. با فرض آن که میانگین توزیعی که Γ از آن پیروی می‌کند برابر صفر است (یا هر مقدار ثابت دیگری نیز می‌تواند برای میانگین تابع f در نظر گرفت) و همچنین فرض دیگر آنکه که واریانس خطا $\text{Var}(r)$ کران‌دار است. مدل فوق به مدلی با مشتق جزئی درجه اول محدود می‌شود و آزمون گاما (Γ) نشان‌دهنده آن بخش از واریانس داده‌ها خروجی است که نمی‌تواند توسط مدل هموار محاسبه شود.

آزمون گاما بر اساس $x_{i,k}$ است که p نزدیک‌ترین همسایگی است مقادیر این آزمون برای مجموعه داده‌های ورودی و خروجی بر اساس معادلات ۳ و ۴ به دست می‌آید:

$$\delta_M(k) = \frac{1}{M} \sum_{i=1}^M |x_{N(i,k)} - x_i|^2, \quad (1 \leq k \leq p) \quad \text{رابطه ۳}$$

$\delta_M(k)$ مجذور فاصله بین نقطه مورد نظر ورودی و نزدیک‌ترین همسایگی می‌باشد و $|\dots|$ فاصله اقلیدسی بین نزدیک‌ترین همسایگی و نقطه مورد نظر را مشخص می‌کند و مقدار خروجی برابر است با:

$$\gamma_M(k) = \frac{1}{2M} \sum_{i=1}^M |y_{N(i,k)} - y_i|^2, \quad (1 \leq k \leq p) \quad \text{رابطه ۴}$$

که $y_{N(i,k)}$ مترادف مقدار y برای k امین همسایه نزدیک x_i در معادله (۳) است. با ایجاد رابطه رگرسیون خطی بین p مجموعه‌ی $\delta_M(k), \gamma_M(k)$ مقدار آماره گاما برابر عرض از مبدأ خط رگرسیون ایجاد شده می‌شود که این معادله به‌صورت زیر است:

$$\gamma = A\delta + \Gamma \quad \text{رابطه ۵}$$

اگر محور عمودی قطع شود ($\delta=0$)، مقدار Γ به‌صورت زیر نشان داده می‌شود،

$$\gamma_M(k) \rightarrow \text{Var}(r) \quad \text{as} \quad \delta_M(k) \rightarrow 0 \quad \text{رابطه ۶}$$

^۱. Smooth Function

معادله (۵) نشان می‌دهد که عرض از مبدأ خط مذکور برابر مقدار آزمون گاما است که از آن بخش از واریانس داده‌های خروجی که نمی‌تواند به وسیله مدل محاسبه گردد به دست می‌آید. شیب خط رگرسیون (A) نشان‌دهنده پیچیدگی مدل است هرچه این شیب تندتر باشد نشان‌دهنده پیچیدگی بیشتر مدل است. می‌توان استاندارد بودن نتایج (نزدیک‌تر بودن به واقعیت) را با شاخص V_{ratio} ملاحظه کرد که مابین صفر و یک می‌باشد. V_{ratio} را می‌توان به صورت زیر نشان داد:

$$V = \frac{\Gamma}{\sigma^2(y)} \quad \text{رابطه ۷}$$

که در اینجا $\sigma^2(y)$ ، واریانس داده‌های خروجی y است. این نسبت نشان‌دهنده دقت مدل برای یافتن خروجی مطلوب از روی ورودی‌هاست. هر چه نسبت V به صفر نزدیک‌تر شود پیش‌گویی خروجی y مناسب‌تر می‌شود. هرگاه این متغیر به یک نزدیک شود نمایانگر یک خطای تصادفی بسیار بزرگ است و مدل مورد نظر نامناسب است. در واقع اگر مقدار نسبت V را از یک کم کنیم مقدار ضریب تبیین به دست می‌آید (جونز، ۲۰۰۴: ۱۰۴-۱۴۹؛ مقدم نیا و همکاران، ۲۰۰۹: ۹۷۵-۹۸۲).

مدل‌های غیرخطی

امروزه به دلیل پیشرفت فناوری رایانه‌ای، تعداد مدل‌های غیرخطی زیادی مانند شبکه عصبی، ماشین‌های بردار پشتیبان، سیستم‌های فازی، تابع‌های چند جمله‌ای، رگرسیون خطی محلی و غیره به وجود آمده است. در این مطالعه دو نوع رایج یعنی رگرسیون خطی محلی و شبکه عصبی مورد توجه قرار دادیم.

رگرسیون خطی محلی

این تکنیک، یک روش خطی غیر فراسنجی (پارامتریک) است که در بسیاری از پیش‌بینی‌های کم بعد و مسائل ساده به کار می‌رود. تنها مشکل LLR^۱ تصمیم‌گیری درباره‌ی اندازه‌ی p_{max} است؛ یعنی تعداد نزدیک‌ترین همسایه‌هایی که برای مدل‌سازی به کار می‌روند. روش انتخاب p_{max} برای رگرسیون خطی شاخص تأثیر نامیده می‌شود و در ادامه شرح داده می‌شود. برای به دست آوردن نزدیک‌ترین همسایگی در نقاط p_{max} معادله ماتریس خطی باید حل شود:

$$\mathbf{Xm} = \mathbf{y} \quad \text{رابطه ۸}$$

که X یک ماتریس $p_{max} \times d$ است که p_{max} نقاط ورودی (متغیرهای مستقل) در d بعد است، $\mathbf{x}_i (1 \leq i \leq p_{max})$ نزدیک‌ترین همسایگی نقاط است. y بردار ستونی خروجی متناظر به طول p_{max} است و m یک بردار ستونی است که باید تعیین شود تا رابطه بین x با y برقرار گردد، مانند:

^۱. Local Linear Regression

$$\text{رابطه ۹} \quad \begin{pmatrix} x_{11} & x_{12} & x_{13} & \dots & x_{1d} \\ x_{21} & x_{22} & x_{23} & \dots & x_{2d} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_{x_{p \max} 1} & x_{x_{p \max} 2} & x_{x_{p \max} 3} & \dots & x_{x_{p \max} d} \end{pmatrix} \begin{pmatrix} m_1 \\ m_2 \\ m_3 \\ \vdots \\ m_d \end{pmatrix} = \begin{pmatrix} y_1 \\ y_1 \\ \vdots \\ y_{p \max} \end{pmatrix}$$

اگر X مجذور و یگانه باشد پس تنها راه‌حل معادله (۸)، $\mathbf{m} = \mathbf{X}^{-1}\mathbf{y}$ است. اگر X غیر مجذور و یگانه نباشد معادله

(۸) را اصلاح کرده و با کمینه کردن معادله زیر، m را به دست می‌آوریم:

$$\text{رابطه ۱۰} \quad |\mathbf{Xm} - \mathbf{y}|^2$$

به‌وسیله پینرسون (۱۹۵۵: ۴۰۶-۴۱۳) و پینرسون (۱۹۵۶: ۱۷-۱۹)، نشان داده شده که تنها راه‌حل مسئله

$\mathbf{m} = \mathbf{X}^\# \mathbf{y}$ است که $\mathbf{X}^\#$ ماتریس معکوس ساختگی است.

شبکه عصبی مصنوعی

شبکه‌های عصبی مصنوعی بر اساس ارتباطات داخلی بسیار گسترده، همانند سیستم عصبی و مغز انسان بنا نهاده شده‌اند. شبکه‌های عصبی جزء سیستم‌های دینامیکی می‌باشند که با پردازش داده‌های تجربی قانون نهفته در ورای اطلاعات را به ساختار شبکه منتقل می‌کنند. به همین خاطر آن‌ها را سیستم هوشمند می‌نامند، چرا که بر اساس محاسبات روی داده‌های عددی و یا مثال‌ها، قوانین کلی را فرا می‌گیرند (چین و همکاران، ۱۹۹۹: ۲۶۶). روند نماهای یادگیری متفاوتی در شبکه‌های عصبی وجود دارد یکی از روند نماهای رایج برگشت به عقب بر اساس کاهش شیب است. روند نماهایی مانند شیب توأم، گوسی-نیوتن، لونبرگ-مارکویت و غیره روند نماهای سریع می‌باشند که در بسیاری از روش‌های عددی به‌کاربرده می‌شود. در این مطالعه از روش‌های شبکه عصبی با روند نما یادگیری برودن، فلتشر، گلدفارب، شانو (۱۹۸۷) (BFGS) و روند نما یادگیری کاهش شیب توأم که در نرم‌افزار Win Gamma موجود است (دورانت، ۲۰۰۱) استفاده گردید.

مدل‌های تجربی

در این پژوهش از دو نوع مدل تجربی آنگستروم، پرسکات و گلور - مک کلویت برای انجام شبیه‌سازی تابش خورشیدی استفاده شد.

¹. Broyden-Fletcher-Goldfarb-Shanno

مدل آنگستروم-پرسکات پیشنهادی توسط فائو

این مدل به دلیل سادگی، کاربرد گسترده‌ای در محاسبه تابش رسیده به سطح زمین دارد و در برآورد تبخیر - تعرق به روش پنمن، مانیت و فائو برای مناطقی که روش مناسبی جهت محاسبه تابش آن‌ها ارائه نشده و یا مقادیر تابش آن‌ها اندازه‌گیری نمی‌شود، مورد استفاده قرار می‌گیرد (موسوی بایگی و همکاران، ۱۳۸۹: ۸۳۸). مدل آنگستروم-پرسکات به‌طور کلی به‌صورت زیر است که در آن R_s تابش خورشیدی رسیده به سطح زمین (مگاژول بر متر مربع بر روز)، R_a تابش فرا زمینی (مگاژول بر متر مربع بر روز که از روش ارائه شده توسط آلن و همکاران ۱۹۹۵ می‌توان به دست آورد)، n ساعات آفتابی واقعی، N ساعات آفتابی حداکثر، a و b ضرایب ثابتی هستند که برای هر منطقه قابل محاسبه است (خلیلی و رضایی صدر، ۱۳۷۶: ۱۵-۳۵).

$$R_s = R_a \left(a + b \frac{n}{N} \right) \quad \text{رابطه ۱۱}$$

فائو ضرایب a و b را به ترتیب برابر 0.25 و 0.5 پیشنهاد کرده و بنابراین رابطه بالا را به شکل زیر ارائه نموده است:

$$R_s = R_a \left(0.25 + 0.5 \frac{n}{N} \right) \quad \text{رابطه ۱۲}$$

مدل گلور-مک کلوت

گلور و مک کلوت ضریب a در رابطه آنگستروم-پرسکات را به عرض جغرافیایی (ϕ) ارتباط داده و رابطه زیر را ارائه نمودند (موسوی بایگی و همکاران، ۱۳۸۹: ۸۳۹):

$$\frac{R_s}{R_a} = 0.29 \cos \phi + 0.25 \frac{n}{N} \quad \text{رابطه ۱۳}$$

ارزیابی اعتبار مدل‌ها

کارایی روش‌های رگرسیونی، شبکه عصبی و تجربی با استفاده از سه شاخص آماری جذر میانگین مربعات خطا (RMSE)، میانگین مربعات خطا (MAE) و ضریب تبیین (R^2) مورد ارزیابی قرار گرفت. RMSE اختلاف بین مقادیر مقایسه شده و شبیه سازی شده را نشان می‌دهد. کوچک بودن مقدار MAE درستی تخمین را نشان می‌دهد. R^2 میزان وابستگی خطی دو شبیه را اندازه‌گیری می‌کند.

$$RMSE = \left[\frac{\sum_{i=1}^n (p_i - o_i)^2}{n} \right]^{1/2} \quad \text{رابطه ۱۴}$$

$$MAE = \frac{\sum_{i=1}^n |p_i - o_i|}{n} \quad \text{رابطه ۱۵}$$

رابطه ۱۶

$$R^2 = 1 - \left(\frac{\sum_{i=1}^n (o_i - p_i)^2}{\sum_{i=1}^n (o_i - \bar{o})^2} \right)$$

در این روابط O_i و P_i به ترتیب مقادیر مشاهده شده و پیش‌بینی شده در زمان i ، \bar{O} و \bar{P} میانگین مقادیر مشاهده شده و پیش‌بینی شده و n تعداد نمونه‌ها است.

یافته‌ها

اطلاعات اقلیمی در طی سال‌های ۱۳۸۵ الی ۱۳۸۹ به دو قسمت جهت آموزش و شبیه‌سازی تقسیم ۸۵۰ عدد برای آموزش و ۴۰۰ عدد برای تست به کار رفت. ویژگی‌های آماری داده‌های هواشناسی شامل دما (T)، سرعت باد (U)، ساعات آفتابی (n) و رطوبت نسبی (RH) ایستگاه مورد نظر در جدول (۱) ارائه شده است. در این جدول S_x ، X_{max} ، X_{min} و C_v که به ترتیب میانگین، انحراف معیار، ضریب تغییرات، حداکثر و حداقل داده‌ها می‌باشند.

تجزیه و تحلیل داده‌ها با آزمون گاما

پیش‌بینی تابش خالص دریافتی با استفاده از چهار فاکتور دمای ماکزیمم، سرعت باد متوسط و تابش برون زمینی و ساعات آفتابی انجام گرفت. این ترکیب بر اساس حداقل کردن مقدار آماره گاما به دست آمده است.

جدول ۱: ویژگی‌های آماری پارامترهای آماری روزانه داده‌های هواشناسی ایستگاه هواشناسی زاهدان

موقعیت	داده‌ها	واحد	X_{mean}	S_x	C_v		
					(S_x/X_{mean})	X_{min}	X_{max}
زاهدان	T	°C	۱۹/۶۸	۷/۹۹	۰/۴۵	-۱۱	۴۴
	U	m/s	۲/۲	۱/۴۳	۰/۶۵۳	۰	۹/۲۱
	n	hr	۹/۳	۲/۹۸	۰/۳۲۱	۰	۱۳/۲۳
	RH	%	۲۹/۹۶	۱۵/۳۱	۰/۵۱۱	۰	۱۰۰

(منبع: نگارندگان)

نتایج مربوط به تجزیه و تحلیل با استفاده از آزمون گاما در جدول (۲) موجود است. هرچه مقدار خطای استاندارد به صفر نزدیک‌تر باشد اطمینان بیشتری از آماره گاما حاصل می‌شود. شیب (A) یک شاخص برای پیچیدگی مدل است که از روی داده‌های ورودی و خروجی ساخته می‌شود و مقدار کم آن نشان‌دهنده مناسب بودن مدل است. نسبت V توانایی پیش‌بینی خروجی با استفاده از ورودی‌ها را نشان می‌دهد که هرچه به صفر نزدیک‌تر باشد دقت آن بیشتر می‌شود. در این مدل کوچک بودن مقادیر گاما، خطای استاندارد، شیب و نسبت V می‌تواند توانایی مدل را در شبیه‌سازی اثبات کند.

جدول ۲: نتایج مربوط به آزمون گاما (منبع: نگارندگان)

مقدار	فرا سنج
۰/۰۴۳۳	مقدار گاما
۰/۰۹۲۸	شیب
۰/۰۰۳۲۱	خطای استاندارد
۰/۱۳۲	V-ratio

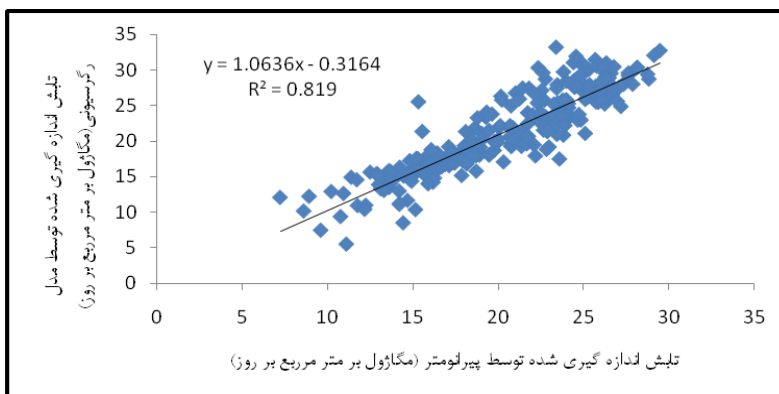
نتایج مدل‌های تجربی و روش‌های غیرخطی

در جدول (۳) مقادیر RMSE، MAE و R^2 که در نتیجه مقایسه مقادیر برآورد شده تابش خورشیدی رسیده به سطح زمین با استفاده از مدل‌های ارائه شده در بالا و نیز مقادیر واقعی اندازه‌گیری شده تابش توسط پیرانومتر نشان داده شده است. این نتایج بر اساس داده‌هایی مورد استفاده در صحت یابی مدل‌ها است. به‌طور کلی هرچه مقدار شاخص‌های خطا سنجی پایین‌تر بوده و میزان R^2 بیشتر باشد مدل مربوطه دقیق‌تر و مناسب‌تر می‌باشد. در شکل (۱) تا (۵) میزان همبستگی این مدل‌ها با مقدار اندازه‌گیری شده تابش توسط پیرانومتر (تابش واقعی) نشان داده شده است.

از جدول (۳) و شکل‌های (۱) تا (۵) مشخص می‌شود که مدل شبکه عصبی با روند نما یادگیری BFGS دارای بالاترین ضریب تعیین (۰/۹۳) و همچنین RMSE و MAE کمتر (به ترتیب ۱/۹۵ و ۱/۷۷) نسبت به سایر مدل‌ها است. بعد از آن نیز دیگر مدل شبکه عصبی یعنی مدل شبکه عصبی با کاهش شیب توأم قرار دارد. بعد از آن دیگر مدل غیرخطی یعنی مدل رگرسیون خطی محلی می‌باشد. از میان مدل‌های تجربی نیز مدل گلور مک-کلوت به علت دارا بودن مقادیر RMSE و MAE کمتر و R^2 بیشتر (به ترتیب برابر ۴/۶۴، ۳/۰۷ و ۰/۵ در مقابل ۳/۸۸، ۳/۲۱ و ۰/۳۳) نسبتاً از مدل آنگستروم-پرسکات بهتر عمل کرده است. به‌طور کلی می‌توان بیان کرد که مدل‌های شبکه عصبی و رگرسیونی عمل شبیه‌سازی را بهتر از مدل‌های تجربی انجام می‌دهند. از مقایسه دو مدل تجربی نیز مشخص می‌شود که این دو مدل تفاوت بسیاری با همدیگر ندارند.

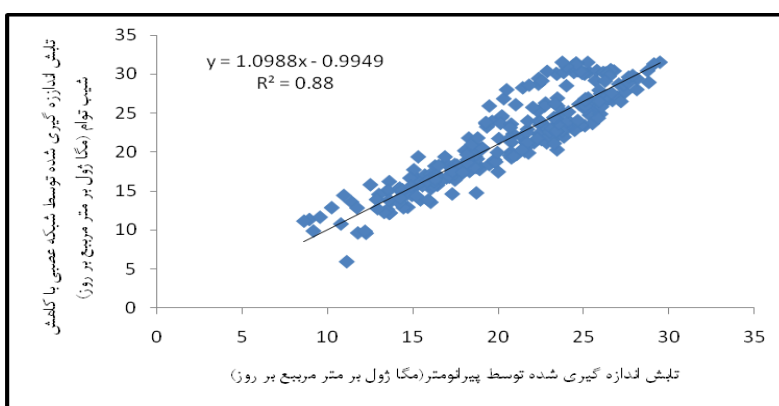
جدول ۳: نتایج تحلیل آماری مدل‌های غیرخطی و تجربی (منبع: نگارندگان)

مدل	RMSE	MAE	R^2
رگرسیون خطی محلی	۲/۸۹	۱/۸۹	۰/۸۲
شبکه عصبی با کاهش شیب توأم	۲/۵۳	۱/۷۷	۰/۸۹
شبکه عصبی با روند نما BFGS	۱/۹۵	۱/۴۷	۰/۹۳
آنگستروم-پرسکات	۴/۳۸	۳/۲۱	۰/۳۳
گلور-مک کلوت	۴/۶۴	۳/۰۷	۰/۵



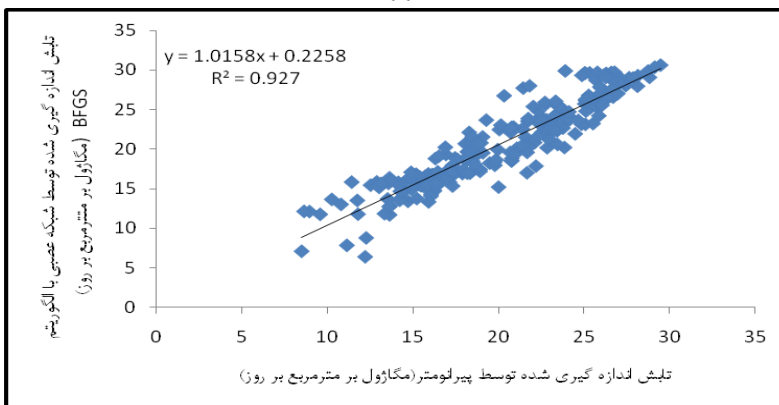
(منبع: نگارندگان)

شکل ۱: میزان همبستگی مدل رگرسیونی با مقادیر اندازه‌گیری شده با پیرانومتر برای داده‌های که برای تست مدل به کار رفته



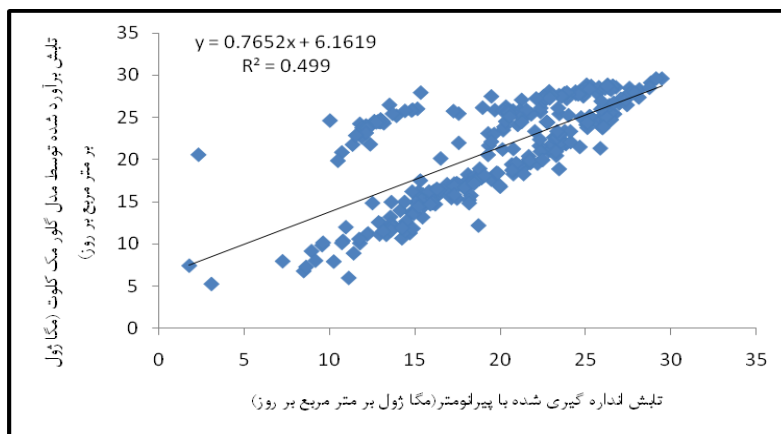
(منبع: نگارندگان)

شکل ۲: میزان همبستگی مدل شبکه عصبی با کاهش شیب توأم با مقادیر اندازه‌گیری شده با پیرانومتر برای داده‌های که برای تست مدل به کار رفته



(منبع: نگارندگان)

شکل ۳: میزان همبستگی مدل شبکه عصبی با الگوریتم BFGS با مقادیر اندازه‌گیری شده با پیرانومتر برای داده‌های که برای تست مدل به کار رفته



(منبع: نگارندگان)

شکل ۴: میزان همبستگی مدل گلور مک-کلوت با مقادیر اندازه گیری شده با پیرانومتر برای داده های که برای تست مدل به کاررفته



(منبع: نگارندگان)

شکل ۵: میزان همبستگی مدل آنکستزوم-پرسکات با مقادیر اندازه گیری شده با پیرانومتر برای داده های که برای تست مدل به کاررفته

نتیجه گیری

این مقاله روش تخمین تابش خالص دریافتی (R_s) با استفاده از آزمون گاما و ترکیب آن با روش مدل های غیرخطی را شرح داد. در این مطالعه توانایی آزمون گاما در ایجاد مدل های غیرخطی برای تخمین تابش با موفقیت نشان داده شد. در این پژوهش با استفاده داده های اندازه گیری شده ایستگاه هواشناسی زاهدان کارایی مدل های رگرسیونی، شبکه عصبی و تجربی مورد ارزیابی قرار گرفت. در پیش بینی تابش به طور تصادفی ۸۵۰ عدد از داده ها برای آموزش مدل ها و ۴۰۰ عدد از داده ها را برای پیش بینی به کار بردیم. بهترین ترکیب با استفاده از چهار پارامتر دمای ماکزیمم، سرعت باد متوسط، تابش برون زمینی و ساعت های آفتابی به دست آمد. این ترکیب بر اساس کمینه کردن خطا است و برای نقاط مختلف متفاوت می باشد. برخلاف پژوهش مقدم نیا و همکاران (۲۰۰۹: ۹۷۵-۹۸۲) و همچنین ریسمان و همکاران (۲۰۰۸: ۴۳۰۱-۴۳۰۹) به دلیل کمی بارش در این منطقه، بارش تأثیر چندانی نداشته است و در ایجاد مدل از آن استفاده نشده است. مشابه با سایر پژوهش های صورت گرفته در زمینه مدل سازی تابش یا استفاده از مدل های غیرخطی و شبکه عصبی ساعت های آفتابی بیشترین تأثیر را داشته و مهم ترین پارامتر بوده است. در مرحله اعتبار سنجی بر اساس ضریب

تعیین بیشترین مقدار مربوط به مدل شبکه عصبی با روند نما BFGS با مقدار $0/93$ و کمترین مقدار مربوط به مدل تجربی آنگستروم - پرسکات با مقدار $0/33$ است. مقدار RMSE و MAE دارای کمترین مقدار $1/95$ و $1/47$ برای مدل BFGS و بیشترین مقدار برای مدل تجربی آنگستروم - پرسکات است. به‌طور مدل‌های غیرخطی خیلی بهتر از مدل‌های تجربی عمل کرده‌اند. دو مدل تجربی نیز تفاوت چندانی با یکدیگر ندارند. نتایج مدل‌های غیرخطی هماهنگی مناسبی را بین نتایج اندازه‌گیری و پیش‌بینی نشان می‌دهد. در این شبیه‌سازی بر عکس پژوهش مقدم نیا و همکاران (۲۰۰۹: ۹۷۵-۹۸۲)، مدل شبکه عصبی با روند نما BFGS بهتر از شبکه عصبی با کاهش شیب توأم و مدل رگرسیونی عمل کرده است. همچنین بر خلاف پژوهش رمان و همکاران (۲۰۰۸: ۴۳۰۹-۴۳۰۱) که بهترین مدل، مدل رگرسیونی بود می‌باشد و بیان کننده این مطلب است که با توجه به شرایط آب و هوایی مختلف نتایج متفاوتی ارائه می‌گردد. بر اساس نتایج به‌دست‌آمده از پژوهش حاضر، می‌توان جمع‌بندی نمود که استفاده صحیح از روش‌های رگرسیونی و مدل‌های شبکه‌های عصبی مصنوعی می‌تواند زمینه‌ساز آنالیز سریع تابش دریافتی به خاک در سیستم‌های آب و خاک و تبخیر و تعرق باشد که البته نیازمند انجام مطالعات بیشتر در این مورد و در مکان‌های دیگر است.

منابع

- ۱- پیری، جمشید. انصاری، حسین. فرید حسینی، علیرضا (۱۳۹۲): مدل‌سازی تابش خورشید رسیده به زمین با استفاده از (ANFIS) و مدل‌های تجربی (مطالعه موردی: ایستگاه‌های زاهدان و بجنورد) نشریه انرژی ایران، دوره ۱۶ شماره ۳ پاییز، صص ۳۷-۸۵.
- ۲- خلیلی، علی. رضایی صدر، حسین (۱۳۷۶)، برآورد تابش کلی خورشید در گستره ایران بر مبنای داده‌های اقلیمی، فصل‌نامه تحقیقات جغرافیایی، شماره ۴۶، صص ۱۵ تا ۳۵.
- ۳- سبزی پرور، علی‌اکبر. بیات ورکشی، مریم (۱۳۸۹)، ارزیابی دقت روش‌های شبکه عصبی مصنوعی و عصبی- فازی در شبیه‌سازی تابش کل خورشیدی، مجله پژوهش فیزیک ایران، جلد ۱۰، شماره ۴، صص ۳۴۷ تا ۳۵۷.
- ۴- موسوی بایگی، محمد. اشرف، بتول. میان‌آبادی، آمنه (۱۳۸۹): بررسی مدل‌های مختلف برآورد تابش خورشیدی به‌منظور معرفی مناسب‌ترین مدل در یک اقلیم نیمه‌خشک، مجله آب‌و‌خاک، جلد ۲۴، شماره ۴، صص ۸۴۴-۸۳۶.
- ۵- قویدل حیدری، عباس (۱۳۸۵): بررسی توان بالقوه انرژی خورشیدی در استان سیستان و بلوچستان، پایان‌نامه کارشناسی ارشد.

- 6- Agalbjörn S, Končar N, Jones AJ. (1997): A Note on The Gamma Test. *Neural Computing and Applications*, 5(3):131-133. ISSN 0-941-0643.
- 7- Allen R. (1995): *Evaluation of Procedures of Estimating Mean Monthly Solar Radiation From Air Temperature*. FAO, Rome.
- 8- Angstrom A. (1924): *Solar and Terrestrial Radiation*. *Quarterly Journal of The Royal Meteorological Society* 50: 121-125.

- 9- Azadeh A, Maghsodi A, Sohrabkhani S, (2009): An Integrated Artificial Neural Networks Approach for Predicting Global Radiation. *Energy Conversion and Management*. 50:1479-1505.
- 10- Bosch JL, Lopes G and Batlles FJ, (2008): Daily Solar Irradiation Estimation Over a Mountainous Area using Artificial Neural Networks. *Renewable Energy*. 33: 1622-1628
- 11- Chuzhanova NA, Jones AJ, Margetts S. (1998): Feature Selection for Genetic Sequence Classification. *Bioinformatics* 14(2): 139–143.
- 12- Dinçer I, Dilmaç S, Türe IE, Edin M. (1996): Simple Technique for Estimating Solar Radiation Parameters and Its Application for Gebze. *Energy Conversion and Management*. 37(2): 183–198.
- 13- Durrant PJ. (2001): winGamma: A Non-Linear Data Analysis and Modelling Tool With Applications to Flood Prediction. PhD Thesis, Department of Computer Science, Cardiff University, Wales, UK.
- 14- Fletcher R. (1987): *Practical Methods of Optimization* (2nd ed.). New York: Wiley.
- 15- Jain, S.K. A. Das, and D. k. Srivastva. (1999): Application of ANN for Reservoir In Flow Prediction and Operation, *Journal of Water Resources Planning and Management*, ASCE, 125(5):263-271
- 16- Jones A J. (2004): *New Tools In Non-Linear Modelling and Prediction*. *Computational Management Science*. 1: 109–149.
- 17- Jones A J, Tsui A, De Oliveira AG. (2002): *Neural Models of Arbitrary Chaotic Systems: Construction and The Role of Time Delayed Feedback in Control and Synchronization*. *Complexity International* Vol. 09
- 18- Moghaddamia, A. Remsan, R. Hassanpour Kashani, M. Mohammadi, M. Han, D Piri, J. (2009): Comparison of LLR, MLP, Elman, NNARX and ANFIS Mmodel- With a Case Study In Solar Radiation Estimation. *Journal of Atmospheric and Solar- Terrestrial Physics*, 71, 975–982.
- 19- Mohandes M, Balchonaim A, Rehman KS, Halawani TO. (1998): Estimation of Global Solar Radiation Using Artificial Neural Networks. *Renewable Energy* 14(1–4): 179–184. Doi:10.1016/S0960-1481(98)00065-2
- 20- Noia M, Ratto CF, Festa R. (1993a): Solar Irradiance Estimation From Geostationary Satellite Data. I. *Statistical Models Solar Energy* 51(6):449-456.
- 21- Ogulata RT, Ogulata, SN. (2002): Solar Radiation on Adana, Turkey. *Applied Energy* 71(4): 351–358.
- 22- Penrose, R. (1955): A Generalized Inverse for Matrices. *Proceedings of The Cambridge Philosophical Society* 51, 406–413.
- 23- Penrose R. (1956): On Best Approximate Solution of Linear Matrix Equations. *Proceedings of the Cambridge Philosophical Society*, 52:17–19.
- 24- Piri J. Shamshirb and S. Petkovic D. Tong CW. Rehman MH. (2015): Prediction of The Solar Radiation on The Earth Using Support Vector Regression Technique. *Infrared Physics & Technology*. 68: 179–185.
- 25- Remesan, R. Shamim, M.A. Han, D. (2008): Model Data Selection Using Gamma Test For Daily Solar Radiation Estimation. *Hydrological Processes* 22 (21), 4301–4309.
- 26- Saylan L, Sen O, Toros H, Arısoy A. (2003): Solar Energy Potential For Heating Cooling Systems in Big Cities of Turkey. *Energy Conversion and Management* 43(14): 1829–1837.
- 27- Shamshirband S. Mohammadi K. Piri J. Petkovic D. Karim A. (2016): Hybrid Auto-Regressive Neural Network Model For Estimating Global Solar Radiation in Bandar Abbas, Iran. *Environ Earth Science*, 75(172): 1-12.
- 28- Tan, S. B. K, Shuy, E.B. and Chua, L.H.C, (2007): Modelling Hourly and Daily Open-Water Evaporation Rates in Areas With An Equatorial Climate, *Hydrological Processes*, 21(486-499), DOI: 10.1002/hyp.6251
- 29- Tovar HF, Baldasano JM. (2001): Solar Radiation Mapping From NOAA AVHRR Data in Catalonia, Spain. *Journal of Applied Meteorology* 40:1821-1834.

- 30- Tsui APM. (1999): Smooth Data Modelling and Stimulus-Response via Stabilization of Neural Chaos. PhD Thesis, Department of Computing, Imperial College of Science, Technology and Medicine, University of London
- 31- Tsui APM, Jones AJ, De Oliveira AG. (2002): The Construction of Smooth Models Using Irregular embedding's Determined By a Gamma Test Analysis. *Neural Computing and Applications* 10(4):318–329.
- 32- Tymvios FS, Jacovides CP, Michaelides SC, Scouteli C, (2005): Comparative Study of Angstrom's and Artificial Neural Networks Methodologies In Estimating Global Solar Radiation: *Sol Energy*. 78: 752-762.