

بهبود عملکرد نرم افزارهای هوش مصنوعی در شبیه سازی بارش- رواناب با استفاده از روش حذف - تزریق نوفه

افشین پرتویان^۱، وحید نورانی^{۱*}، محمد تقی اعلمی^۲

چکیده

شبیه سازی دقیق فرایندهای آب شناسی از قبیل بارش- رواناب می تواند اطلاعات مهمی را از یک حوضه ی آبریز برای مدیریت منابع آب فراهم کند. اخیراً در این راستا، نرم افزارهای مختلف جعبه ی سیاه برای شبیه سازی چنین پدیده ی پیچیده ای به کاررفته اند. نرم افزارهای داده محور وابستگی زیادی به کمیت و کیفیت داده ها دارند و داده های دارای نوفه کارایی نرم افزارها را تحت تاثیر قرار داده و رفع نوفه از داده ها با استفاده از یک روش مناسب می تواند منجر به کارایی بهتر نرم افزارهای داده محور شود. لذا در این مقاله ابتدا با استفاده از روش رفع نوفه موجکی اقدام به رفع نوفه بد از گروه های زمانی روزانه کرده و سپس با افزودن نوفه خوب به این داده های رفع نوفه شده و تشکیل مجموعه های آموزشی مختلف، شبیه سازی بارش - رواناب برای ایستگاه پل آنیان واقع در حوضه آبخیز زرینه رود در سراب سد بوکان، با استفاده از نرم افزارهای جعبه سیاه ANN و ANFIS انجام شده است. همچنین برای بررسی دقت نرم افزار پیشنهادی، نتایج این شبیه سازی ها با مدل های سنتی جعبه سیاه ARIMA و MLR مورد مقایسه قرار گرفته اند. نتایج نشان می دهند که اعمال هم زمان رفع نوفه موجکی و افزودن نوفه خوب در شبیه سازی با ANN بهبود ۲۳ درصدی و در شبیه سازی با ANFIS باعث بهبود ۱۴ درصدی، در مرحله ی آزمایش نرم افزارها شده است.

واژه های کلیدی: شبیه سازی بارش- رواناب، نرم افزارهای جعبه سیاه، رفع نوفه موجکی، تزریق نوفه، حوضه ی آبخیز زرینه رود.

^۱ گروه عمران آب، دانشکده عمران، واحد نجف آباد، دانشگاه آزاد اسلامی، نجف آباد، ایران

Email: nourani@tabrizu.ac.ir

^۲ دانشکده عمران، دانشگاه تبریز، تبریز، ایران

۱- مقدمه

از مهمترین و تاثیرگذارترین فرایندهای آب شناسی، فرآیند تبدیل بارش به رواناب است. برحسب استفاده از روش های گوناگون برای ساده سازی فرآیند، انواع نرم افزارها برای شبیه سازی بارش - رواناب توسعه یافته و می یابند. در این راستا نرم افزارهای جعبه سیاه در شبیه سازی پدیده های پیچیده از جمله بارش - رواناب به طور گسترده ای مورد استفاده قرار می گیرند. از معروف ترین این مدل های جعبه سیاه می توان به نرم افزار $ARIMA^1$ اشاره کرد که دارای کاربردهای متعددی در پیش بینی فرایندهای آب شناسی است (سالاس و همکاران، ۱۹۸۰). این نرم افزار که بدون توجه به فیزیک پدیده، یک فرایاز خطی میان داده های ورودی و خروجی برقرار می کند، توانایی های محدودی را در شبیه سازی گروه های زمانی غیر خطی و نایستا دارد (نورانی و همکاران، ۲۰۱۱). لذا در پی این محدودیت، روش های مبتنی بر هوش مصنوعی پا به عرصه ظهور گذاشتند. یکی از پر کاربردترین و موفق ترین نرم افزارهای هوش مصنوعی، نرم افزار شبکه عصبی مصنوعی می باشد که به طور گسترده ای در شبیه سازی و پیش بینی گروه های زمانی غیر خطی فرایندهای آب شناسی مورد استفاده قرار گرفته است (آبراهارت و همکاران، ۲۰۱۲). به طور کلی مزایای مدل های ANN را می توان به صورت زیر بیان نمود:

با توجه به خاصیت جعبه سیاه بودن، در استفاده از این نرم افزارها نیازی به دانش قبلی در مورد ماهیت فیزیکی فرآیند نمی باشد و به دلیل اعمال یک صافی غیرخطی به نام تابع محرک بر روی نرون ها، دارای ماهیت غیر خطی بوده و قابلیت اعمال ورودی های چند متغیره را با ویژگی های مختلف دارا می باشند (نورانی و صالحی، ۱۳۸۷). مقالات متعددی به کاربردهای موفق ANN در شبیه سازی فرایندهای آب شناسی، به ویژه بارش - رواناب اشاره کرده اند: (انجمن مهندسان عمران آمریکا، ۲۰۰۰؛ انتروهمکاران، ۲۰۰۶؛ قربان زاده و همکاران، ۱۳۹۴).

از سوی دیگر، نرم افزارها و داده های استفاده شده برای شبیه سازی فرایندهای آب شناسی معمولاً شامل عدم قطعیت می باشند لذا، ترکیب دو نرم افزار ANN و سامانه فازی تحت عنوان نرم افزار $ANFIS^2$ می تواند به

عنوان گزینه ای جهت برخوردار شدن از مزایای هر دو نرم افزار مطرح شود (نورانی و همکاران، ۲۰۱۱). نرم افزار ANFIS کاربردهای متنوعی را در شبیه سازی های آب شناسی، از جمله شبیه سازی بارش - رواناب داشته است (حقی زاده و همکاران، ۱۳۹۴؛ نورانی و همکاران، ۱۳۸۸؛ نورانی و صالحی، ۱۳۸۷؛ فرات و گنجور، ۲۰۰۶؛ اقیل و همکاران، ۲۰۰۷).

اما قابلیت روشهای هوش مصنوعی، مانند ANN و ANFIS برای داده های گروه های زمانی، که شامل نوفه^۴ هستند محدود می باشد. این کاستی به دلیل وابستگی زیاد مدل های داده محور به کیفیت داده های مورد استفاده می باشد. برای غلبه بر این مشکل می توان از روش های مختلف برای حذف نوفه علامت های آب شناسی استفاده کرد (الشورباگی و همکاران، ۲۰۰۲).

زمانی که روش های سنتی و مرسوم در شبیه سازی گروه های زمانی آب شناسی نمی توانند نیازهای عملی را بنابر محدودیتشان در مواجهه با مشخصات نایستایی و مقیاس های چند زمانی برآورده کنند، روش رفع نوفه ای موجکی آستانه ای، که به وسیله دونوهو (۱۹۹۵) ابداع شده است به عنوان یک روش قدرتمند در مقایسه با روش های مرسوم و سنتی می تواند مورد استفاده قرار گیرد (سنگ، ۲۰۱۲). به عنوان نمونه می توان به کاربرد موفق رفع نوفه موجکی در شبیه سازی شبکه ی عصبی در شبیه سازی جریان Murder Creek واقع در آمریکا اشاره نمود (نژاد و نورانی، ۲۰۱۲).

در مدل های داده محور هوش مصنوعی، علاوه بر کیفیت داده ها، کمیت داده ها نیز نقش بسزایی را در کارایی نهایی شبیه سازی دارند. در فرایند آموزش، هر مجموعه داده آموزشی یک نمونه ی محدود از کل داده های موجود خواهد بود لذا، یک مجموعه داده انتخاب شده به تنهایی نمی تواند بیانگر تمام الگوهای موجود باشد. از طرفی، به دلیل محدودیت امکانات و عدم دسترسی، امکان نمونه برداری داده های فراوان، که بتواند الگوهای مختلف از پدیده را شامل گردد، وجود ندارد. به عنوان مثال بده اندازه گیری شده در خروجی در یک زمان خاص از روز به عنوان نماینده دی روزانه در گروه زمانی بده مد نظر قرار می گیرد، حال آنکه در نظر گرفتن چنین عدد ثابتی می

³ - Adaptive Neuro Fuzzy Inference System

⁴ - Noise

¹ - Auto Regressive Integrated Moving Average

² - Artificial Neural Network

۲- مواد و روشها

۲-۱- حوضه ی مورد مطالعه و داده های مورد

استفاده

در این مطالعه از داده های بارش و رواناب ایستگاه پل آنیان (عرض ۱۲-۳۶ و طول ۲۶-۴۶) واقع بر زرينه رود استفاده گردید. حوضه ی آبخیز زرينه رود در شمال غرب ایران، و در جنوب دریاچه ارومیه گسترده شده است. حوضه ی آبخیز زرينه رود در سراب سد زرينه رود (سد بوکان) از چهار زیر حوضه ی اصلی به نامهای سقر چای، جیغاتوچای، خرخره چای و ساروق چای تشکیل می گردد. بر روی رودخانه ی جیغاتوچای دو ایستگاه پل قشلاق و پل آنیان فعال می باشند. حوضه ی آبخیز پل آنیان دارای مساحت ۱۲۹۸ کیلومتر مربع و محیط ۲۱۸ کیلومتر بوده، و شیب متوسط حوضه ۱۱.۹٪ می باشد. طول رودخانه اصلی ۹۹.۴ کیلومتر، و شیب متوسط آن ۰.۳٪ است. متوسط بارندگی سالانه ی این حوضه ۷۰۳ میلی متر بوده و حجم بارندگی سالانه ی آن ۹۱۲.۴ میلیون متر مکعب می باشد. در شکل (۱) حوضه ی جیغاتوچای، و موقعیت ایستگاه پل آنیان در آن نشان داده شده است.

در توسعه و ایجاد نرم افزار از داده های روزانه ی ۲۱ سال آماری از سالهای ۱۳۷۰ الی ۱۳۹۰ استفاده شد. در تقسیم بندی داده ها، جهت آموزش از ۷۰ درصد داده ها، یعنی حدود ۱۵ سال و برای آزمون مدلها از ۳۰ درصد داده ها، یعنی حدود ۶ سال استفاده گردید. گروه های نوفه تولید شده باید دارای ویژگی میانگین صفر و انحراف معیار ثابت باشند تا خواص گروه اصلی را حفظ نمایند. از طرفی، این گروه های نوفه تولید شده دارای مقادیر منفی نیز می باشند. لذا، جهت جلوگیری از ورود داده های منفی در این تحقیق قبل از آموزش و آزمون نرم افزارها مجموعه ی داده های موجود با استفاده از رابطه (۱) بین مقادیر ۰.۹ و ۰.۱ نرمال سازی گردید (راجور کارو همکاران ۲۰۰۲).

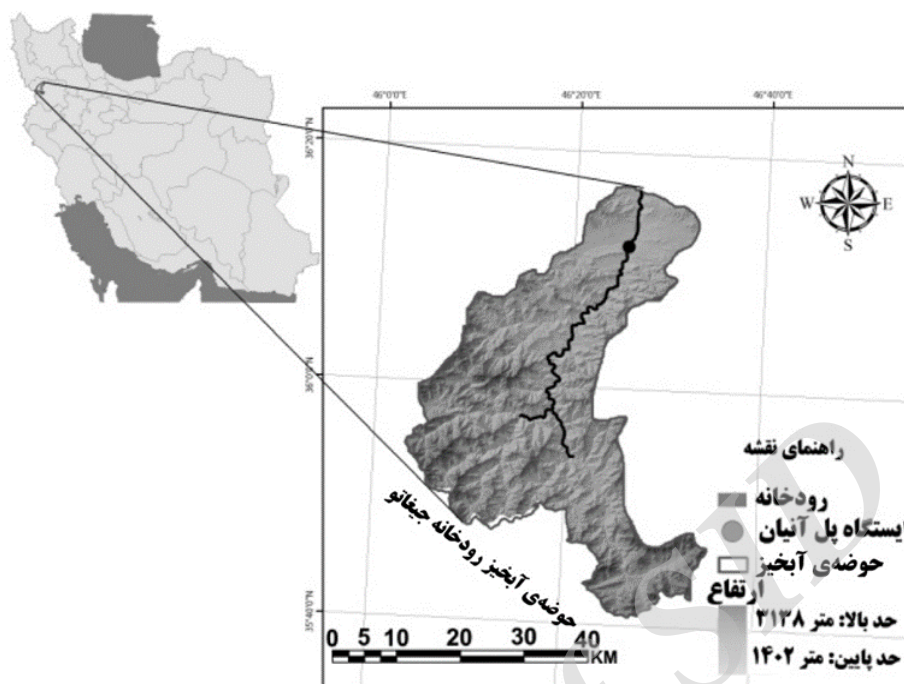
$$N_i = 0.8 \left[\frac{(X_i - X_{\min})}{X_{\max} - X_{\min}} \right] + 0.1 \quad (1)$$

در این رابطه، N_i مقادیر بهنجار شده، X_i مقادیر واقعی مشاهداتی، X_{\max} حداکثر مقادیر واقعی و X_{\min} حداقل مقادیر واقعی می باشند. برای انتخاب بهترین ورودی برای نرم افزارهای مختلف ۵ ترکیب انتخاب شدند، و این امر برای پیدا کردن میزان وابستگی رواناب یک روز آینده به

تواند معروض به عدم قطعیت بوده و نتواند به طور کامل الگوی داده های آن روز را شامل گردد. بدیهی است هر اندازه تعداد داده های نمونه برداری در یک روز بیشتر باشد، در نظر گرفتن متوسط این مقادیر می تواند عدم قطعیت را پایین آورد، لکن به دلیل محدودیت های مالی و انسانی، در عمل امکان نمونه برداری های زیاد امکان پذیر و اقتصادی نمی باشد؛ بنابراین، افزایش تعداد داده ها با استفاده از روش های داده سازی مصنوعی، که داده ها دارای ماهیتی شبیه به داده های اصلی باشند، می تواند به عنوان یک پیش پردازش جهت آمادگی بهتر داده ها برای آموزش نرم افزار در نظر گرفته شود. این فرایند باعث عملکرد بهتر نرم افزار در یادگیری قوانین و الگوی بین داده های واقعی می شود. فرایند افزودن نوفه به داده های اصلی برای تولید داده های آموزشی اضافی را Jitter می نامند (ژنگ ۲۰۰۷). آموزش با این داده ها باعث ایجاد قیده های بیشتر، پیشرفت قابلیت تعمیم مدل و جلوگیری از فرایند از بر کردن شبکه در طی فرایند یادگیری خواهد شد.

(ژنگ، ۲۰۰۷؛ آن، ۱۹۹۶؛ رید و همکاران، ۱۹۹۵)

از طرفی، تاکنون هیچ روش خاصی برای نحوه ی انتخاب میزان نوفه اعمال شده به داده ها، به طور کامل در مقالات و تحقیقات ارائه نشده است، لذا برای رسیدن به یک سطح معقول و بهینه از نوفه باید سطوح مختلف نوفه را ایجاد و با اعمال آنها بر داده های اصلی، اثر هر یک بر روی عملکرد نرم افزار مورد نظر بررسی کرد تا سطح بهینه نوفه برای یک گروه خاص در مسئله ی مورد نظر به دست آید، که برای رسیدن به این مهم از مفهوم تولید داده بر اساس توزیع های مختلف احتمالاتی، استفاده می شود. با این حال، بیشتر تحقیقات انجام گرفته در این زمینه مربوط به مسائل طبقه بندی بوده و در زمینه گروه های زمانی مالی صورت گرفته اند، و بر اساس پیشینه ی تحقیق صورت گرفته، به نظر می رسد هیچ تحقیقی در زمینه شبیه سازی های آب شناسی صورت نگرفته است. علاوه بر این، تا کنون تاثیر همزمان رفع نوفه بدون افزودن نوفه ی خوب نه تنها در آب شناسی بلکه، در هیچ زمینه تحقیقاتی دیگری مورد توجه نبوده است. لذا، لزوم تحقیقات بیشتر در این حوزه، و ارائه ی یک راهکار مناسب جهت شبیه سازی این پدیده ها امری ضروری به نظر می رسد، که در این مقاله صورت گرفته است.



شکل (۱) نقشه DEM حوضه آبخیز جیغاتو چای و موقعیت ایستگاه پل آنیان.

نرم افزار ها شامل شناسایی، تخمین و پیش بینی می باشند. تابع این نرم افزار به صورت (p,d,q) نمایش داده می شود. p بیانگر مرتبه‌ی عبارت خود همبستگی، d مرتبه‌ی تفاضل غیر فصلی و q مرتبه‌ی تأخیر زمانی برای نوفه پیش بینی است. نرم افزار (p,d,q) ARIMA به صورت رابطه‌ی (۲) تعریف می شود (باکس و جنکینس، ۱۹۷۶):

$$W_t = \mu + \frac{\theta_q(B)}{\psi_p(B)} a_t \quad (2)$$

عبارت $W_t = (1-B)^d Y_t$ بیانگر تفاضلی گروه زمانی تک متغیره Y_t می باشد که در آن عملگر تأخیر زمانی است، (به عبارتی $BX_t = X_{t-1}$) μ مقداری ثابت،

$\psi_p(B)$ جمله خودهمبسته B از مرتبه p می باشد:

$$\psi(B) = 1 - \psi_1(B) - \dots - \psi_q(B)^p \quad (3)$$

$\theta_q(B)$ چند جمله میانگین متحرک B از مرتبه q

می باشد که:

$$\theta_q(B) = 1 - \theta_1(B) - \dots - \theta_q(B)^q \quad (4)$$

d مرتبه تفاضلی برای قسمت منظم و غیرفصلی گروه زمانی بوده و a_t اغتشاش تصادفی با میانگین صفر و واریانس مشترک صفر می باشد. زمانی که نرم افزار ARIMA بیش

بارش و رواناب چند روز قبل انجام گردید. ترکیب ورودی‌ها به صورت زیر می باشد:

- Comb. 1: R_t, Q_t
- Comb. 2: R_t, Q_{t-1}, Q_t
- Comb. 3: $R_t, R_{t-1}, Q_{t-1}, Q_t$
- Comb. 4: $R_t, Q_{t-2}, Q_{t-1}, Q_t$
- Comb. 5: $R_{t-1}, R_t, Q_{t-2}, Q_{t-1}, Q_t$

که در ترکیب‌های فوق R_t, Q_t به ترتیب رواناب و بارش مشاهداتی در زمان t می باشد. خصوصیات آماری داده‌های مورد استفاده در جدول (۱) ارائه شده اند.

۲-۲- مدل میانگین متحرک تفاضلی خود همبسته (ARIMA)

نرم افزارهای سنتی جعبه‌ی سیاه همچون ARIMA به صورت گسترده برای پیش بینی گروه‌های زمانی آب شناسی استفاده می شوند (سالاس و همکاران، ۱۹۸۰). ذات این نرم افزار خطی بوده، و بر فرض ایستا بودن داده‌های گروه زمانی استوار است. نرم افزار ARIMA توانایی شناسایی الگوهای پیچیده را در داده‌ها دارا بوده، و امکان پیش بینی را برای آینده مطابق با داده‌های ورودی در زمان گذشته فراهم می آورد. نرم افزارهای ARIMA توانایی تحلیل و پیش بینی داده‌های گروه زمانی تک متغیره را دارا می باشند. سه گام لازم برای استفاده از این

$$f_1 = p_1 x + q_1 y + r_1$$

قانون دوم: اگر x برابر A_2 و y برابر B_2 باشد آن گاه

$$f_2 = p_2 x + q_2 y + r_2$$

که r_i, p_i, q_i ($i=1,2$) فراسنج‌های خطی در

بخش نتیجه نرم افزار فازی تاکاگی - سوگنو درجه اول هستند. مطابق شکل (۳) ساختار ANFIS شامل پنج لایه می‌شود: لایه‌ی اول، گره‌های ورودی: هر گره این لایه یک مجموعه ی فازی است و خروجی هر گره در این لایه درجه ی عضویت متغیر ورودی در این مجموعه فازی است. پارامترهای هر گره تعیین کننده ی شکل تابع عضویت مجموعه ی فازی آن گره است. تابع عضویت مجموعه های فازی معمولاً به وسیله‌ی توابع زنگوله‌ای شکل به صورت زیر بیان می‌شود:

$$O_i^1 = \frac{1}{1 + [(x - c_i) / a_i]^{2b_i}} \quad (5)$$

x مقدار ورودی به گره i و a_i, b_i, c_i فراسنج‌های تابع عضویت این مجموعه هستند. این پارامترها به نام فراسنج‌های شرط نامیده می‌شوند.

لایه‌ی دوم، گره‌های قاعده: هر گره در این لایه درجه‌ی فعالیت یک قانون را محاسبه می‌کند:

$$O_i^2 = w_i = \mu_{A_i}(x) \times \mu_{B_i}(y), i=1,2 \quad (6)$$

$\mu_{A_i}(x)$: درجه عضویت x در مجموعه‌ی A_i و $\mu_{B_i}(y)$

: درجه‌ی عضویت y در مجموعه‌ی B_i است.

لایه‌ی سوم، گره‌های متوسط: در این لایه i امین گره نسبت درجه‌ی فعالیت قانون i ام را به مجموع درجه‌ی فعالیت کلیه‌ی قوانین بصورت زیر محاسبه می‌کند:

$$O_i^3 = w_i^n = \frac{w_i}{w_1 + w_2}, i=1,2 \quad (7)$$

در این رابطه w_i^n : درجه‌ی فعالیت بهنجار شده‌ی قانون i ام است.

لایه‌ی چهارم، گره‌های نتیجه: در این لایه، خروجی هر گره برابر است با:

$$O_i^4 = w_i^n f_i = w_i^n (p_{ix} + q_{iy} + r_i), i=1,2 \quad (8)$$

در این رابطه r_i, q_i, p_i فراسنج‌های تطبیقی لایه‌اند

و به نام فراسنج‌های نتیجه خوانده می‌شوند.

از یک گروه زمانی (متغیر) به عنوان ورودی داشته باشد، موسوم به نرم‌افزار^۱ ARIMAX خواهد شد. نرم‌افزار ARIMAX، به دلیل استفاده از اطلاعات بیشتر (تعداد ورودی بیشتر) در پیش بینی، معمولاً نتایج بهتری را نسبت به نرم‌افزارهای ARIMA ارائه می‌دهد.

۲-۳- شبکه‌ی عصبی مصنوعی (ANN)

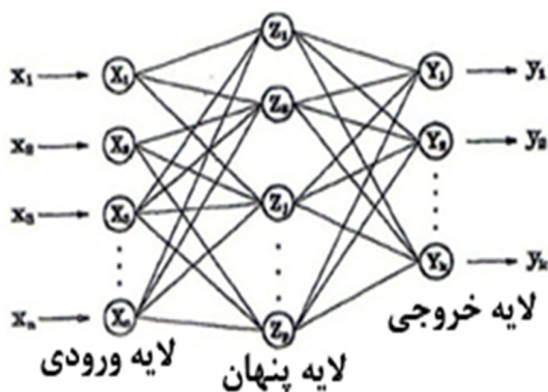
در شکل گرفتن علم شبکه های عصبی مصنوعی از شبکه های عصبی زیستی الهام گرفته شده است. شبکه های عصبی مصنوعی با پردازش داده های تجربی، دانش یا قانون نهفته را ورای داده‌ها به ساختار شبکه منتقل می‌کنند، و در واقع، بر اساس محاسبات روی داده های عددی یا مثالها، قوانین کلی را فرا می‌گیرند، به همین دلیل به این سامانه‌ها هوشمند گفته می‌شود. این سامانه‌ها در مدل سازی ساختار اتصال عصبی مغز بشر می‌کوشند. در شبکه‌های مصنوعی پردازش اطلاعات در اجزایی به نام نرون صورت می‌گیرد، و اطلاعات بین نرون‌ها از طریق ارتباطات موجود بین آنها رد و بدل می‌گردد. هر یک از این ارتباطات، مشخصه یا وزنی را مختص به خود دارد که در اطلاعات انتقال یافته از یک نرون به نرون دیگر ضرب می‌شود. هر یک از نرون‌ها برای محاسبه‌ی خروجی خود، یک تابع محرک را، که معمولاً غیر خطی است، به ورودی‌هایش (جمع اطلاعات وزن دار شده) اعمال می‌نماید. در شکل (۲) شکل کلی یک شبکه عصبی پیشرو نشان داده شده است.

۲-۴- سامانه استنتاج فازی-عصبی تطبیقی (ANFIS)

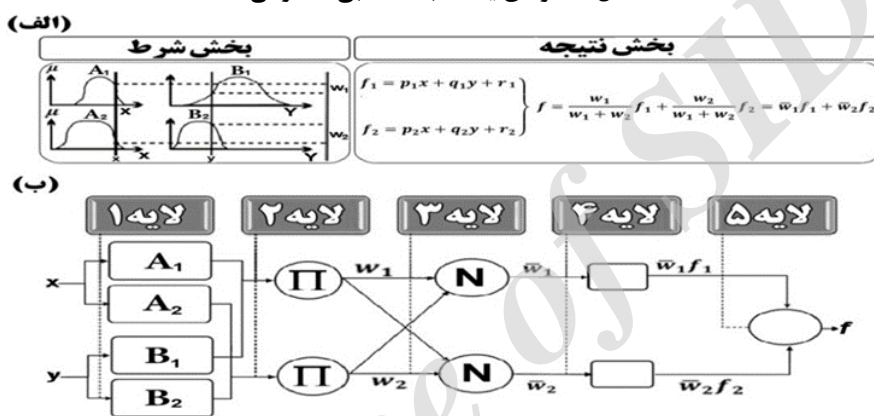
اصطلاح ANFIS مخفف عبارت سامانه استنتاج فازی تطبیق پذیر مبتنی بر شبکه های عصبی مصنوعی است. این نظام پیاده سازی یک سامانه فازی سوگنو به صورت ساختار شبکه ای پیش رونده است. به منظور ساده سازی در تعریف سامانه استنتاج عصبی فازی تطبیقی، فرض می‌شود که سامانه استنتاجی مورد نظر دو ورودی x, y و یک خروجی f دارد. برای یک نرم افزار فازی سوگنو درجه‌ی اول، می‌توان یک صورت زیر را بیان کرد (نورانی وهمکاران، ۲۰۱۱):

قانون اول: اگر x برابر A_1 و y برابر B_1 باشد آن گاه

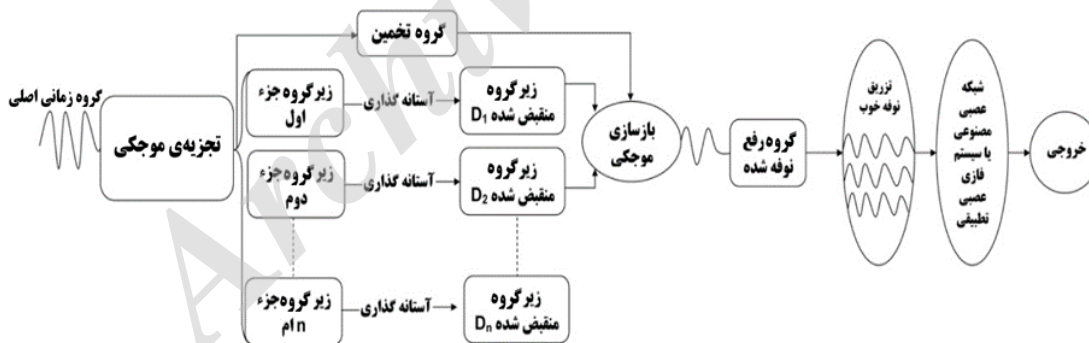
¹- Auto Regressive Integrated Moving Average with exogenous input



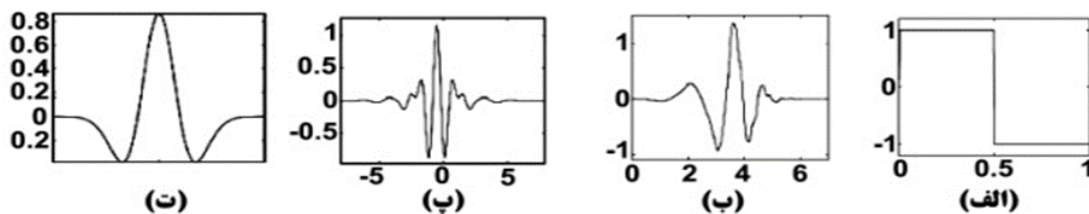
شکل ۲- نمونه‌ی یک شبکه عصبی مصنوعی.



شکل ۳- الف) ساختار نرم افزار ANFIS (ب) نحوه عملکرد و ارتباط بین لایه ها.



شکل ۴- نمودار طرح واردهی نرم افزار پیشنهادی.



شکل ۵- نمودار (الف) تابع موجک Haar (ب) تابع موجک db4 (پ) تابع موجک Meyer (ت) تابع موجک Mexican hat

می‌گردد. دونوهو (۱۹۹۵) روش رفع نوفه موجکی را براساس آستانه‌ها برای به دست آوردن علامت‌های رفع نوفه شده صحیح ارائه داد. این روش، که متداول‌ترین راه رفع نوفه موجکی است، تحت مراحل زیر انجام می‌گیرد:

الف- پس از انتخاب موجک مادر مناسب و سطح تفکیک پذیری M ، سری زمانی اصلی به زیر گروه تخمین در سطح تفکیک پذیری M ، و زیر گروه‌های جزء در سطوح تفکیک پذیری گوناگون با استفاده از تبدیل موجک تجزیه می‌شوند.

ب- مقادیر مطلق زیر گروه‌های جزء $d_j(t)$ ($j=1, 2, \dots, M$) در صورت کمتر بودن از آستانه مشخص T' در هر سطح تفکیک پذیری به صفر تبدیل می‌شوند. زیروند نشانگر سطوح تفکیک پذیری زام می‌باشد. اگر مقادیر مطلق زیر گروه‌های جزء بزرگتر از مقدار آستانه T' در آن زیر سری باشند آن‌گاه اختلافشان با T' به وسیله‌ی رابطه (۱۱) جایگزین مقادیر گروه‌های جزء می‌گردد (دونوهو ۱۹۹۵):

$$d_j(t) = \begin{cases} \operatorname{sgn}(d_j(t))(|d_j(t)| - T') & |d_j(t)| > T' \\ 0 & |d_j(t)| \leq T' \end{cases} \quad (11)$$

رابطه‌ی (۱۱) مقادیر آستانه را برای رفع نوفه‌ی موجکی در هر سطح تفکیک در زیر گروه‌های جزء بکار می‌گیرد. لازم بذکر است که در این روش زیر گروه تخمین آستانه گذاری نمی‌شود.

ج- بازسازی موجکی گروه زمانی رفع نوفه شده از زیر گروه‌های رفع نوفه شده که شامل زیر گروه تخمین در مرتبه تفکیک M ، و زیر گروه‌های جزئی پردازش شده $d_j(t)$ می‌باشد، در همه سطوح تفکیک حاصل می‌شود.

۲-۵-۲- تولید نوفه و گروه‌های مصنوعی

امروزه به جای مشاهده‌ی مستقیم یک سامانه فیزیکی که دارای ورودیهای تصادفی متعدد و رفتار تصادفی است برای بررسی اثرات فراسنج‌های مختلف می‌توان از شبیه‌سازی رایانه‌ای استفاده کرد. خصوصاً، در مواردی که محاسبه‌ی مستقیم مشکل است یا فراسنج‌های متعددی در مسئله دخیل هستند، جهت شبیه‌سازی به مولد اعداد تصادفی نیاز می‌باشد همچنین، برای نمونه‌برداری تصادفی از یک جامعه، از اعداد تصادفی

لایه‌ی پنجم، گروه‌های خروجی: در این لایه هرگره مقدار خروجی نهائی را به صورت زیر محاسبه می‌کند (تعداد گر-ها برابر تعداد خروجی‌هاست):

$$O_i^5 = \sum w_i^n f_i = \frac{\sum w_i f_i}{\sum w_i} \quad (9)$$

به این صورت یک نرم افزار فازی، که قابلیت یادگیری را دارد پیاده سازی می‌شود. روش آموزش اصلی در این نرم افزار، روش پس انتشار خطا است. در این روش، با استفاده از الگوریتم شیب نزولی خطا، مقدار خطا به سمت ورودی‌ها پخش می‌گردد، و فراسنج‌ها تصحیح می‌شوند. می‌توان خروجی کل را بصورت ترکیب خطی فراسنج‌های نتیجه نوشت:

$$\begin{aligned} f &= w_1^n f_1 + w_2^n f_2 \\ &= (w_1^n x) p_1 + (w_1^n y) q_1 + (w_1^n) r_1 \\ &\quad + (w_2^n x) p_2 + (w_2^n y) q_2 + (w_2^n) r_2 \end{aligned} \quad (10)$$

۲-۵-۲- نرم افزار پیشنهادی پیش پردازش داده‌ها

به طور خلاصه می‌توان نرم افزار پیشنهادی را در این مقاله چنین بیان کرد که در این روش ابتدا با اعمال روش رفع نوفه موجکی^۲ برای داده‌های خام پیش پردازش انجام شده و رفع نوفه انجام می‌گیرد. سپس بر اساس توزیع بهنجار اقدام به ایجاد گروه‌های نوفه با انحراف معیارهای مشخص شده و بعداز این مرحله گروه‌های نوفه تولید شده به گروه زمانی رفع نوفه شده موجکی اضافه، و گروه‌های جدید با الگوهای یکسان با گروه‌زمانی اصلی تولید شده و به عنوان ورودی نرم افزار های ANN، ANFIS، ARIMA و MLR به کار می‌روند. در شکل (۴) نمودار طرح واژه نرم افزار پیشنهادی نشان داده شده است.

۲-۵-۱- رفع نوفه با آستانه گذاری نرم موجکی

تبدیل موجک یکی از تبدیل‌های ریاضی کارآمد در زمینه‌ی پردازش علامت می‌باشد. توابع موجک‌ها دارای انواع بسیار متنوعی هستند که در شکل (۵) نمودار تابع بعضی از آنها نشان داده شده است.

در تحلیل علامت‌های گسسته از تبدیل موجک گسسته، که به اختصار DWT^3 گفته می‌شود، استفاده

²- Wavelet- denoising

³- Discrete wavelet - transform

$$y = G^{-1}(\widehat{F(x)})$$

توزیع یکنواخت
دارای توزیع تجمعی G

در نرم افزار متلب می توان بر اساس توزیع های مختلف اعداد تصادفی را ایجاد کرد. در این تحقیق از toolbox متلب، قسمت تولید داده های تصادفی این نرم افزار با استفاده از دستور NORMRND، جهت ایجاد گروه های اعداد تصادفی با توزیع نرمال بهره وری شده است.

۲-۶- معیار ارزیابی نرم افزار

یک نرم افزار مناسب لازم است نتایج قابل قبولی را از نظر ضریب تبیین (DC^4) و جذر میانگین مربعات خطا ($RMSE^5$) در هر دو مرحله آموزش و صحت سنجی نشان دهد، بنابراین برای ارزیابی کارایی نرم افزار در این مطالعه، از روابط (۱۲) و (۱۳) جهت مقایسه ی عملکرد نرم افزارهای مختلف استفاده شد (نورانی و صالحی ۱۳۸۷):

$$DC = 1 - \frac{\sum_{i=1}^n (O_{obs_i} - O_{com_i})^2}{\sum_{i=1}^n (O_{obs_i} - \bar{O}_{obs_i})^2} \quad (12)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (O_{obs_i} - O_{com_i})^2}{N}} \quad (13)$$

در این روابط N ، O_{obs_i} ، O_{com_i} ، \bar{O}_{obs_i} به ترتیب تعداد مشاهدات، داده های مشاهداتی، داده های محاسباتی و میانگین داده های مشاهداتی می باشند. مقدار بیشتر (تا یک) برای DC و مقدار کمتر برای RMSE در هر دو مرحله آموزش و صحت سنجی نشان دهنده ی عملکرد بهتر مدل می باشد.

۳- نتایج و بحث

جهت شبیه سازی فرایند بارش - رواناب در حوضه ی آبخیز سد بوکان با استفاده از مقادیر بارش و بده ثبت شده در ایستگاه پل آنیان از شناسه نویسی و قابلیت های نرم افزار متلب، استفاده شد. شبیه سازی ها با نرم افزار های ANN، ANFIS، ARIMA و MLR برای سه حالت مختلف با داده های پیش پردازش نشده ی خام، داده های رفع نوفه شده موجکی، و بالاخره داده های رفع نوفه شده

استفاده می شود. دنباله ای از نمونه های متغیرهای تصادفی را دنباله اعداد تصادفی گویند. طبیعتاً، تولید اعداد تصادفی فقط به وسیله ی یک پدیده ی فیزیکی تصادفی امکان پذیر است. آنچه را به وسیله ی رایانه می توان تولید کرد، دنباله ی اعداد شبه تصادفی (pseudo random) است، یعنی به وسیله ی الگوریتمی به طور deterministic تولید می شوند، ولی خواص دنباله ی اعداد تصادفی را دارا هستند و لذا تصادفی به نظر می رسند. آزمون هایی وجود دارند که تصادفی بودن دنباله را بررسی می کنند. تولید داده های تصادفی معمولاً شامل دو مرحله است. ابتدا داده ی تصادفی با توزیع یکنواخت ایجاد و سپس از این اعداد تصادفی با توزیع یکنواخت، جهت ایجاد اعداد تصادفی با توزیع دلخواه استفاده می شود. برای تولید دنباله ی اعداد تصادفی با توزیع یکنواخت، یک روش که در عین ساده بودن، دنباله ی حاصل از آن از نظر تصادفی بودن خیلی مطلوب است، الگوریتم Lehmer می باشد (لیبرمن و باوکر، ۱۳۸۸). در این الگوریتم فرض می شود m یک عدد اول بسیار بزرگ و z_0 عددی دلخواه و کوچکتر از m باشد. عدد $a < m$ طوری انتخاب می شود که:

$$\forall n < m - 1 : a^n \neq 1 \pmod{m}$$

در این صورت

$$a^{m-1} = 1 \pmod{m}$$

به این ترتیب دوره تکرار دنباله بیشینه می شود و کلیه ی اعداد ۱ تا $m-1$ ، یک و فقط یک بار ظاهر می شوند. بعد از ایجاد اعداد تصادفی با توزیع یکنواخت می توان از روش های متنوعی، از جمله روش تبدیل معکوس برای ایجاد اعداد تصادفی با توزیع دلخواه، استفاده کرد. در صورتی که متغیر تصادفی x دارای توزیع تجمعی F باشد، در این صورت $u=F(x)$ دارای توزیع یکنواخت $u(0,1)$ خواهد بود، و بر عکس، اگر $u \sim u(0,1)$ باشد، در این صورت $x=F^{-1}(u)$ دارای تابع توزیع انباشته ی F خواهد بود، و نیز اگر X دارای توزیع F باشد، برای ساختن متغیر تصادفی y با تابع توزیع داده شده G خواهیم داشت (لیبرمن و باوکر ۱۳۸۸):

4 - Determination Coefficient

5 - Root Mean Square Error

جدول‌های (۳) و (۴)، و نتایج بهترین نرم افزار رفع نوفه شده موجکی در MLR و ARIMA در جدول (۵) ارائه شده و از ارائه سایر نتایج در این جدول خودداری شده است. در این جدول‌ها، منظور از ساختار شبکه عصبی به صورت a-b-c، تعداد نرون‌های لایه ورودی، میانی و خروجی می‌باشد و Epoch تعداد تکرارهای مرحله آموزش است.

نتایج گزارش شده حاکی از آنند که بهترین نرم افزار در مرحله رفع نوفه موجکی در شبکه‌ی عصبی، نرم افزاری است که با موجک مادر Db4، و سطح تفکیک پذیری ۸، و آستانه‌ی ۱۵ رفع نوفه شده است، که در این حالت، DC آموزش به ۰.۹۲ و DC آزمایش به ۰.۸۷ افزایش می‌یابد. نتایج شبیه سازی ANN با داده‌های رفع نوفه شده در مقایسه با مدل سازی با داده‌های پردازش نشده بهبود ۱۶ درصدی را در مرحله‌ی آزمایش نشان می‌دهد. این نتیجه نشانگر انطباق بیشتر موجک مادر Db4 با گروه زمانی مورد مطالعه در مقایسه با سایر موجک های مادر می باشد. نتایج رفع نوفه موجکی در شبیه سازی‌های ANFIS نیز حاکی از افزایش کارایی نرم افزارها داشته، به طوری که مقادیر ضرایب تبیین آموزش و آزمایش را به ترتیب به ۰.۹۳ و ۰.۸ می‌رساند. این نتایج بهبود ۷ درصدی را در مقایسه با مرحله‌ی آموزش با داده‌های پردازش نشده در مرحله‌ی آزمایش مدل‌ها نشان می‌دهند. بهترین فراسنج‌های رفع نوفه‌ی موجکی در تمام نرم افزار ها موجک مادر Db4 با سطح تفکیک پذیری ۸ و آستانه‌ی ۱۵ بوده، و این نتیجه تاییدکننده‌ی انطباق شکل گروه زمانی مورد مطالعه با شکل موجک مادر Db4 است.

پس از تعیین نرم افزار بهینه‌ی رفع نوفه شده ANN و ANFIS در ایستگاه مورد مطالعه، اقدام به ساختن داده‌های مصنوعی با استفاده از تولید نوفه براساس خواص گروه زمانی اصلی گردید. برای این منظور گروه‌های نوفه با میانگین صفر و انحراف معیارهای ۰.۰۰۰۱، ۰.۰۰۰۱، ۰.۰۰۲۵، ۰.۰۰۱ و ۰.۰۵ تولید، سپس این نوفه‌ها به نرم افزار رفع نوفه شده بهینه اضافه گردید، و برای هر گروه جدید ساخته شده (نوفه دار شده) مجدداً نرم افزار های ANN، ANFIS، ARIMA و MLR آموزش داده شده و نتایج این مدل‌ها مورد ارزیابی قرار گرفتند. نتایج نرم

و پیش پردازش شده با نوفه‌ی خوب، صورت گرفته و نتایج در جداول (۲) تا (۸) ارائه شده اند. قابل توجه است که در این جدول‌ها فقط نتایج بهترین شبیه سازی‌ها ارائه شده و از ارائه‌ی نتایج سایر شبیه سازی‌ها خودداری شده است. قابل ذکر است که در توسعه‌ی نرم افزارهای مبتنی بر ANFIS، از روش گسسته سازی شبکه‌ای و الگوریتم آموزش ترکیبی استفاده شده است، و تابع عضویت بردار ورودی تابع گوسی، تابع عضویت بردار خروجی خطی، و برای هر بردار ورودی ۳ تابع عضویت انتخاب شدند. نتایج بهینه‌ی نرم افزارهای مختلف برای یافتن بهترین ترکیب ورودی در جدول (۲) ارائه شده اند.

بررسی نتایج نرم افزار شبکه عصبی مصنوعی بیانگر آن است که بهترین ترکیب، آن است که رواناب فردا را با استفاده از بارش یک و دو روز قبل، و رواناب یک و دو روز قبل، مدل سازی می‌کند. این نرم افزار دارای مقدار ضریب تبیین ۰.۸۹ در مرحله‌ی آموزش، و ۰.۷۵ در مرحله‌ی آزمایش نرم افزار می‌باشد. همچنین بهترین ترکیب ورودی در نرم افزار ANFIS آن است، که ورودی‌های آن رواناب یک و دو روز قبل و بارش یک روز قبل می‌باشند، که نتایج ضرایب تبیین آموزش و آزمایش با این ورودی‌ها به ترتیب برابر با ۰.۹۰ و ۰.۷۴ می‌باشند. نرم افزار های ARIMA و MLR^۶ (فراپاز خطی چند متغییره) نیز نتایج قابل قبولی را در شبیه سازی بارش - رواناب ارائه کرده‌اند، که بهترین نتایج این شبیه سازی‌ها در نرم افزار MLR، (Comb.5) و در نرم افزار ARIMA، (Comb.3) با ساختار (۳ و ۱ و ۳) می‌باشند، که نرم افزار ARIMA نتایج نسبتاً بهتری را از خود نشان داده است.

پس از این مرحله، اقدام به رفع نوفه‌ی موجکی با استفاده از موجکهای مادر خانواده‌ی دوبجی‌ها، یعنی Haar و Db2 و Db3 و Db4 برای رواناب روزانه‌ی ایستگاه پل آنیان با سطوح تفکیک پذیری ۸ و ۹ (با توجه به آنکه داده‌های مورد تحقیق روزانه می‌باشند)، و آستانه‌های متفاوت گردید. علت انتخاب این دو سطح تفکیک پذیری آن است که یک سال ۳۶۵ روز بوده و بین دو وجه ۲^۸ و ۲^۹ قرار می‌گیرد و لذا این دو امکان وقوع شدت دوره سالانه مورد توجه قرار می‌گیرد. نتایج نرم افزارهای رفع نوفه شده موجکی ANN و ANFIS به ترتیب در

⁶ - Multi Linear Regression

از مقایسه‌ی نتایج نرم افزارهای پردازش شده و پردازش نشده، اثرات مفید گروه‌های پیش پردازش شده کاملاً مشهود بوده، و باعث بهبود نتایج شبیه سازی‌ها گردیده اند. این امر به علت وابستگی زیاد نرم افزارهای داده محور به کیفیت و کمیت داده‌های مورد استفاده می‌باشد. نتایج شبیه سازی با داده‌های رفع نوفه شده در مقایسه با شبیه سازی با داده‌های پردازش نشده، به ترتیب بهبود ۱۶ و ۷ درصدی را در مرحله‌ی آزمایش در مدل‌های ANN و ANFIS نشان می‌دهند. نتایج نرم افزارهای اعمال همزمان رفع نوفه‌ی موجکی و افزودن نوفه‌ی خوب، تاثیر بیشتری را در کارایی مدل‌ها داشته، و باعث بهبود ۲۳ و ۱۴ درصدی به ترتیب در نرم افزارهای ANN و ANFIS در مرحله‌ی آزمایش نرم افزارها شده است. قابل ذکر است که در نرم افزارهای MLR و ARIMA نیز، گروه نوفه تولید شده با انحراف معیار ۰.۰۰۱ بیشترین تأثیر را در افزایش کارایی مدل‌ها داشته، و روش پیشنهادی موجب بهبود ۱۱ و ۱۰ درصدی در مرحله آزمایش در نرم افزارهای ARIMA و MLR، در مقایسه با حالتی که شبیه سازی با داده‌های پیش پردازش نشده انجام گرفته، گردیده است.

افزارهای ANN و ANFIS در جدول‌های (۶) و (۷)، و نتایج نرم افزارهای MLR و ARIMA در جدول (۸) ارائه شده‌اند که در این جدول‌ها نتایج مدل‌های بهینه ارائه شده و از ارائه سایر نتایج خودداری شده است.

لازم به ذکر است که برای هر نوفه با انحراف معیار مشخص، ده گروه زمانی تولید شد که در این جدول‌ها فقط بهترین نتایج ارائه شده اند.

بهترین نتیجه‌ی شبیه سازی ANN، بعد از افزودن نوفه‌ی خوب به گروه زمانی رفع نوفه شده، مربوط به گروه نوفه تولید شده با انحراف معیار ۰.۰۰۱ بوده، که DC آموزش و آزمایش را به ترتیب تا مقادیر ۰.۹۵ و ۰.۹۲ افزایش می‌دهد، که این نتیجه‌ی خوب، انطباق هر چه بیشتر داده‌های پردازش شده را با داده‌های مشاهداتی نشان می‌دهد.

بهترین نتیجه‌ی شبیه‌سازی ANFIS، بعد از افزودن نوفه‌ی خوب به گروه زمانی رفع نوفه شده، گروه نوفه‌ی تولید شده با انحراف معیار ۰.۰۰۱ بوده، که DC آموزش را به ۰.۹۵، و DC مرحله آزمون مدل را به ۰.۸۵ می‌رساند. نمودار بهترین نرم افزارهای ANN و ANFIS در مرحله‌ی آزمایش در شکل‌های (۶) و (۷) نشان داده شده‌اند.

جدول ۱- مشخصات آماری داده‌های بارش و رواناب روزانه پل آنیان.				
فراسنج های آب شناسی	مشخصات آماری	کل	آموزش	آزمایش
بارش (mm)	حداقل	۰	۰	۰
	حداکثر	۷۶.۵	۷۶.۵	۴۰
	میانگین	۱.۰۵	۱.۱۱	۰.۹۱
	انحراف از معیار	۳.۹۶	۴.۲۲	۳.۲۴
رواناب (m ³ /s)	حداقل	۰	۰	۰
	حداکثر	۴۰.۷	۴۰.۷	۲۴۷
	میانگین	۱۸.۵۲	۱۷.۰۶	۲۱.۹۴
	انحراف از معیار	۳۶.۹۶	۴۶.۱	۳۲.۱۶

جدول ۲- نتایج بهینه‌ی نرم افزارهای مختلف با داده های خام.					
نرم افزار	ترکیب ورودی	DC		RMSE (بهنجار شده)	
		آموزش	آزمایش	آموزش	آزمایش
ANN	Comb3	۰.۸۹	۰.۷۵	۰.۰۲۵۴	۰.۰۳۱۴
ANFIS	Comb2	۰.۹۰	۰.۷۴	۰.۰۲۱۵	۰.۰۳۱۷
ARIMA	Comb3	۰.۸۲۰	۰.۷۳	۰.۰۳۱۸	۰.۰۳۳۱
MLR	Comb5	۰.۷۵	۰.۶۷	۰.۰۳۳	۰.۰۳۶۷

جدول ۳- نتایج شبیه سازی ANN با داده های رفع نوفه شده.							
موجک مادر	بهترین ساختار	سطح تفکیک	آستانه	Epoch	DC		RMSE (بهنجار شده)
					آموزش	آزمایش	آموزش
Haar	۴-۸-۱	۸	۴	۶۰	۰.۹۱	۰.۷۸	۰.۰۲۴۴
Db2	۴-۳-۱	۸	۱۰	۵۰	۰.۹۲	۰.۸۵	۰.۰۲۴
Db3	۴-۵-۱	۸	۲۰	۱۰۰	۰.۹۲	۰.۸۴	۰.۰۲۴۹
Db4	۴-۳-۱	۸	۱۵	۵۰	۰.۹۲	۰.۸۷	۰.۰۲۲۷

جدول ۴- نتایج شبیه سازی ANFIS با داده های رفع نوفه شده.							
موجک مادر	سطح تفکیک	آستانه	آموزش	آزمایش	DC		RMSE (بهنجار شده)
					آموزش	آزمایش	آموزش
Haar	۸	۱۰	۰.۹۲	۰.۷۸	۰.۲۰۴	۰.۰۲۹۲	
Db2	۹	۴	۰.۹۲	۰.۸	۰.۲۰۱	۰.۰۲۷۹	
Db3	۹	۱۰	۰.۹۲	۰.۷۹	۰.۰۲	۰.۰۲۸۶	
Db4	۸	۱۵	۰.۹۳	۰.۸	۰.۱۹۵	۰.۰۲۷۸	

جدول ۵- نتایج شبیه سازی MLR و ARIMA با داده های رفع نوفه شده.								
نرم افزار	موجک مادر	سطح تفکیک	آستانه	آموزش	آزمایش	DC		RMSE (بهنجار شده)
						آموزش	آزمایش	آموزش
ARIMA	Db4	۸	۱۵	۰.۸۴	۰.۷۶	۰.۳۰۹	۰.۰۳۱۱	
MLR	Db4	۸	۱۵	۰.۷۸	۰.۷۲	۰.۳۲۹	۰.۰۳۳۹	

جدول ۶- نتایج افزودن نوفه ی خوب به داده های رفع نوفه شده در نرم افزار ANN.							
انحراف معیارنوفه	بهترین ساختار	Epoch	آموزش	آزمایش	DC		RMSE (بهنجار شده)
					آموزش	آزمایش	آموزش
۰.۰۰۰۱	۴-۱۰-۱	۴۰	۰.۹۵	۰.۹۱	۰.۱۷۹	۰.۰۲۰۱	
۰.۰۰۱	۴-۱۰-۱	۸۰	۰.۹۵	۰.۹۲	۰.۱۷۲	۰.۰۲۰۲	
۰.۰۲۵	۴-۹-۱	۸۰	۰.۹۴	۰.۸۹	۰.۲۰۴	۰.۰۲۱۶	
۰.۰۱	۴-۱۰-۱	۶۰	۰.۹۴	۰.۹۱	۰.۱۸۸	۰.۰۲۲	
۰.۰۵	۴-۹-۱	۹۰	۰.۸۷	۰.۸۰	۰.۲۸۳	۰.۰۳۲	

جدول ۷- نتایج افزودن نوفه ی خوب به داده های رفع نوفه شده در نرم افزار ANFIS.							
انحراف معیارنوفه	آموزش	آزمایش	DC		RMSE (بهنجار شده)		
			آموزش	آزمایش	آموزش	آزمایش	
۰.۰۰۰۱	۰.۹۳	۰.۷۳	۰.۱۹۳	۰.۳۲۵			
۰.۰۰۱	۰.۹۵	۰.۸۵	۰.۱۷۳	۰.۲۴۶			
۰.۰۲۵	۰.۹۴	۰.۸۲	۰.۱۹۲	۰.۲۶۵			
۰.۰۱	۰.۹۴	۰.۸۳	۰.۱۷۹	۰.۲۶۱			
۰.۰۵	۰.۹۰	۰.۶۵	۰.۲۴۱	۰.۳۷۲			

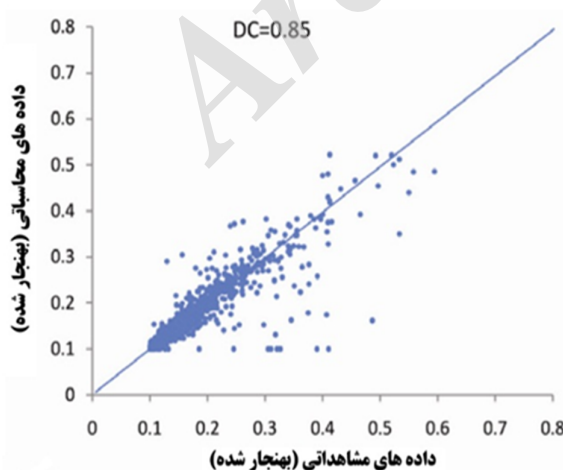
جدول ۸- نتایج افزودن نوفه‌ی خوب به داده های رفع نوفه شده در نرم افزار های MLR و ARIMA.				
نرم افزار	انحراف معیارنوفه	DC		RMSE (پهنجار شده)
		آموزش	آزمایش	آموزش
ARIMA	۰.۰۰۱	۰.۸۸	۰.۸۱	۰.۰۲۶۶
MLR	۰.۰۰۱	۰.۸۰	۰.۷۴	۰.۰۳۲۴

۴- نتیجه گیری و پیشنهادات

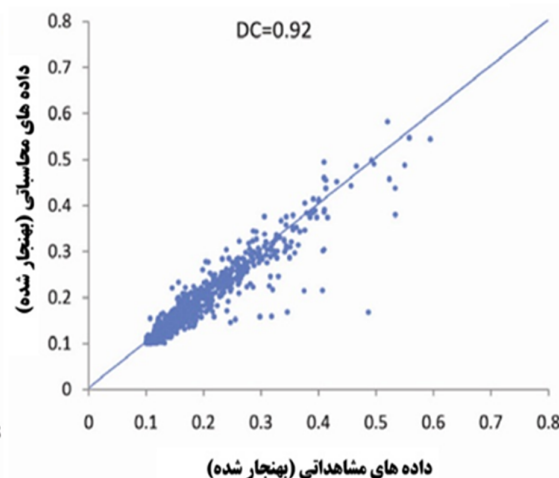
در این تحقیق، در راستای استفاده از روشهای متنوع پیش پردازش داده ها، اثر رفع نوفه‌ی موجکی و اثر افزودن نوفه‌ی خوب به داده های رفع نوفه شده موجکی، برای شبیه سازی بهینه بارش - رواناب نرم افزارهای هوش مصنوعی ANN و ANFIS و نرم افزارهای سنتی جعبه‌ی سیاه MLR و ARIMA مورد تحقیق قرار گرفتند.

اگر چه نرم افزارهای ARIMA و MLR نیز نتایج قابل قبولی را در شبیه سازی بارش - رواناب ارائه کرده اند، ولی در مقایسه با نرم افزار پیشنهادی این مقاله از دقت کمتری برخوردارند. علت این امر توانایی‌های محدود نرم افزارهای خطی در شبیه سازی گروه‌های زمانی غیر خطی و ناپایستا، و از طرفی وابستگی زیاد نرم افزارهای داده محور به کمیت و کیفیت داده‌های مورد استفاده می باشد. بررسی نتایج افزودن نوفه‌ی خوب بیا نگر آن است که گروه زمانی نوفه‌های تولید شده با انحراف معیار بزرگ عملاً باعث ایجاد الگوهای غیر واقعی در داده‌های اصلی شده، به طوری که ضرایب تبیین را در هر دو نرم افزار

ANN و ANFIS در گروه‌های زمانی رفع نوفه شده موجکی پائین آورده، و عملاً باعث پائین آمدن کارایی نرم افزارهای فوق گردیده است. از مقایسه‌ی نتایج مدل‌های پردازش شده و پردازش نشده، اثرات مفید گروه‌های پیش پردازش شده کاملاً مشهود بوده، که این امر ناشی از شناسایی دقیق تر الگوهای نهفته بین داده‌ها، به وسیله‌ی شبکه بوده، و همین امر باعث بالا رفتن کارایی مدل‌های پیش پردازش شده گردیده، به طوری که نرم افزارهای توسعه یافته توانسته‌اند مقادیر رواناب را با درصد خطای پائین تر، و سطح اطمینان بیشتر، شبیه سازی و پیش بینی نمایند. در مقایسه کلی نرم افزارهای پردازش شده با یکدیگر، نکته‌ی قابل توجه آن است که پیش پردازش داده‌ها از قبیل، رفع نوفه بداز گروه زمانی و اعمال دو فرایند رفع نوفه‌ی بدو افزودن نوفه‌ی خوب به صورت همزمان بر گروه زمانی، کارایی نرم افزارها را بالا برده، ولی بهترین نتایج مربوط به اعمال همزمان رفع نوفه‌ی موجکی و افزودن نوفه‌ی خوب می‌باشد، به طوری که نرم افزار پیشنهادی در مقایسه با حالتی که شبیه سازی با داده‌های



شکل ۷- نتایج نرم افزار بهینه‌ی ANFIS در مرحله‌ی آزمایش



شکل ۶- نتایج نرم افزار بهینه‌ی ANN در مرحله‌ی آزمایش

رواناب. نشریه مهندسی عمران و محیط زیست دانشگاه تبریز. شماره ۴. ۶ صفحه.

- 6) Abrahart, R.J., F. Anctil, P. Coulibaly, C.W. Dawson, N.J. Mount, L.M. See, A.Y. Shamseldin, D.P. Solomatine, E. Toth, and R.L. Wilby. 2012. Two decades of anarchy? Emerging themes and outstanding challenges for neural network river forecasting. *Progress in Physical Geography* 36: 480-513.
- 7) An, G. 1996. The effect of adding noise during back propagation training on a generalization performance. *Neural Computation* 8: 643-674.
- 8) Antar, M.A., I. Elassiouti and M.N. Alam. 2006. Rainfall-runoff modeling using artificial neural networks technique: a Blue Nile catchment case study. *Hydrological Process* 20 (5): 1201-1216.
- 9) Aqil, M., I. Kita, A. Yano, and S. Nishiyama. 2007. A comparative study of artificial neural networks and neuro-fuzzy in continuous modeling of the daily and hourly behaviour of runoff. *Journal of Hydrology* 337:22-34.
- 10) ASCE Task Committee on Application of Artificial Neural Networks in Hydrology. 2000. *Artificial Neural Networks in hydrology II: Hydrologic Applications*. Journal of Hydrologic Engineering. 386:27-37.
- 11) Box, G.E.P., and G. Jenkins. 1976. *Time Series Analysis: Forecasting and control*, seconded. Holden-Day, San Francisco.
- 12) Donoho, D.H. 1995. De-noising by soft-thresholding. *IEEE Transactions on Information Theory* 41(3):613-617.
- 13) Elshorbagy, A., S.P. Simonovic, and U.S. Panu. 2002. Noise reduction in chaotic hydrologic time series: facts and doubts. *Journal of Hydrology* 256:147-165.
- 14) Firat, M., and M. Gungor. 2006. River flow estimation using Adaptive Neuro Fuzzy Inference System. *Mathematics and Computers in Simulation* 11:52-62.
- 15) Nejad, F., and V. Nourani. 2012. Elevation of wavelet denoising performance via an ANN-Based streamflow forecasting model. *International Journal of Computer Science and Management Research* 1:764-770.
- 16) Nourani, V., O. Kisi and M. Komasi. 2011. Two hybrid artificial intelligence approaches for modeling rainfall-runoff process. *Journal of Hydrology* 402:41-59.
- 17) Rajurkar, M.P., U.C. Kothiyari, and U.C. Chaube. 2002. Artificial neural networks for daily rainfall-runoff modeling. *Hydrological Sciences Journal* 47(6): 865-877.
- 18) Reed, R., R.J. Marks II, and S. Oh. 1995. Similarities of error regularization, sigmoid gain scaling, target smoothing, and training with jitter.

پیش پردازش نشده انجام گرفته است، بهبود ۲۳، ۱۴، ۱۱ و ۱۰ درصدی در مرحله ی آزمایش مدل ها را به ترتیب در نرم افزارهای ANN، ANFIS، ARIMA و MLR نشان می دهد.

برای تکمیل این مقاله پیشنهاد می شود که نرم افزار ارائه شده برای گروه زمانی روزانه حوضه های دیگر، خصوصاً حوضه های خارجی، صورت گرفته تا از این طریق هم مقایسه ای بین میزان نوفه ی بد موجود در گروه های زمانی برای داده های حاصله از نهادهای مختلف انجام گردد و از طرفی کارایی نرم افزار پیشنهادی در افزایش بازده مدل های شبیه سازی در حوضه های دیگر نیز بررسی شود. توصیه می گردد که رفع نوفه ی موجکی برای گروه زمانی بارش نیز انجام شود، و سپس اثر افزودن نوفه ی خوب به گروه بارش رفع نوفه گردیده، و مدل ترکیبی بارش و بده رفع نوفه شده در مدل سازی ها نیز بررسی گردد. همچنین، پیشنهاد می شود که اثر و کارایی نرم افزار پیشنهادی این مقاله در گروه های زمانی ماهانه، و در سایر فرایندهای آب شناسی از قبیل رواناب - رسوب نیز بررسی گردد.

منابع

- ۱) حقی زاده، ع؛ محمدلو، م؛ نوری، ف. ۱۳۹۴. شبیه سازی فرایند بارش رواناب با استفاده از شبکه عصبی مصنوعی و سیستم فازی عصبی تطبیقی و رگرسیون چندمتغیره (مطالعه موردی: حوضه آبخیز خرم آباد). *اكوهیدرولوژی* ۲: ۲۴۳-۲۳۳.
- ۲) قربان زاده، م؛ ملای نیا، م.ر؛ قره سوفلو، ج. ۱۳۹۴. ارزیابی شبکه عصبی مصنوعی در برآورد بارش رواناب (مطالعه موردی رودخانه كرج). *کنفرانس ملی مهندسی معماری. عمران و توسعه شهری*. ۱۰ صفحه.
- ۳) لیبرمن، ج و باوکر، آ.ه. ۱۳۸۸. *آمار مهندسی*. مرکز نشر دانشگاهی. ۶۰۸ صفحه.
- ۴) نورانی، و؛ و صالحی، ک. ۱۳۸۷. مدل سازی بارش- رواناب با استفاده از روش شبکه عصبی فازی تطبیقی و مقایسه آن با روش های شبکه عصبی و استنتاج فازی. *چهارمین کنگره ملی مهندسی عمران*. دانشگاه تهران. ۸ صفحه.
- ۵) نورا نی، و؛ کی نژاد، م.ع؛ ملکا نی، ل. ۱۳۸۸. استفاده از سیستم فازی- عصبی تطبیقی در مدل سازی بارش -

IEEE Transaction on Neural Networks 6 (3): 529–538.

19) Salas, J.D., J.W. Delleur, V. Yevjevich, and W.L. Lane. 1980. Applied modeling of hydrological time series. Water Resources Publications. Denever.

20) Sang, Y.F. 2012. A practical guide to discrete wavelet decomposition of hydrologic time series. Water Resources Management 26: 3345-3365.

21) Zhang, G.P. 2007. A neural network ensemble method with jittered training data for time series forecasting. Information Sciences 177:5329–5346.

Archive of SID