

# تأثیر نوع الگوریتم آموزش شبکه عصبی پرسپترون چندلایه در دقت پیش‌بینی بارش ماهانه ایران، مطالعه موردی: مدل ECMWF

مرتضی پاکدامن\*

استادیار، پژوهشکده اقلیم‌شناسی و تغییر اقلیم، پژوهشگاه هوشناسی و علوم جو، مشهد، ایران

(دریافت: ۱۴۰۰/۵/۱۰، پذیرش نهایی: ۱۴۰۰/۱۰/۲۰)

## چکیده

امروزه انواع شبکه‌های عصبی مصنوعی در حوزه‌های مختلف علوم جو و اقلیم‌شناسی برای اهدافی نظیر طبقه‌بندی، رگرسیون و پیش‌بینی استفاده می‌شوند. اما سوال اساسی در استفاده از شبکه‌های عصبی مصنوعی، نحوه طراحی و معماری آنهاست. یکی از نکات مهم در استفاده از شبکه‌های عصبی مصنوعی که باید مد نظر طراحان قرار بگیرد، انتخاب الگوریتم مناسب برای آموزش شبکه است. در این مقاله، شش روش مختلف آموزش شبکه عصبی پرسپترون چند لایه شامل روش‌های منظم‌سازی بیزی، لوبنرگ-مارکوات و گرادیان مزدوج پاول-بل، الگوریتم شبه‌نیوتنی BFGS، گرادیان مزدوج فلچر-پاول و گرادیان مزدوج مقیاس‌شده برای پیش‌بینی ماهانه بارش کشور مورد بررسی و مقایسه قرار می‌گیرند. برای این منظور، یک شبکه عصبی پرسپترون برای پس پردازش خروجی بارش ماهانه مدل ECMWF طراحی می‌شود که برای آموزش آن از داده‌های ERA5 و روش‌های آموزش مختلف استفاده می‌شود. برای بررسی عملکرد شش روش آموزش مختلف، مقدار سه شاخص ضریب همبستگی، میانگین مربعات خطا و نیز شاخص نش-ساتکلیف برای هر مدل محاسبه شد. همچنین، عملکرد روش‌های مذکور در مناطق جغرافیایی مختلف کشور برای ماه ژانویه، به‌عنوان نمونه، مورد ارزیابی قرار گرفت. نتایج حاصل از مقایسه شاخص‌ها نشان داد که عملکرد دو روش منظم‌سازی بیزی و لوبنرگ-مارکوات، در مقایسه با چهار روش دیگر برای آموزش شبکه عصبی بهتر است. همچنین این دو روش توانستند، در مقایسه با داده‌های مدل ECMWF، قبل از پس پردازش، نتایج با دقت بیشتری را به‌دست آورند.

واژه‌های کلیدی: شبکه عصبی پرسپترون، مدل ECMWF، الگوریتم منظم‌سازی بیزی، الگوریتم لوبنرگ-مارکوات، الگوریتم گرادیان مزدوج پاول-بل.

## ۱. مقدمه

(۲۰۲۰a) از شبکه عصبی برای پیش‌بینی ماهانه بارش کشور با پس‌پردازش مدل‌های همادی آمریکای شمالی استفاده کردند. استفاده از شبکه‌های عصبی مصنوعی برای طبقه‌بندی نیز مرسوم است. به‌عنوان مثال می‌توان از آنها برای پیش‌بینی وقایع دودویی نظیر رعد و برق نیز استفاده کرد (پاکدامن و همکاران، ۲۰۲۰b).

اما در یک شبکه عصبی مصنوعی پرسپترون، طراحی معماری شبکه بسیار حائز اهمیت است و این طراحی می‌تواند در توانایی شبکه عصبی برای حل مسئله به‌صورت مستقیم اثر بگذارد. در طراحی معماری شبکه باید به سؤالاتی نظیر تعداد نرون‌های هر لایه، تعداد لایه‌ها، توابع محرک شبکه در هر لایه و... پاسخ داد. در برخی مواقع برای پاسخ به هریک از سؤالات فوق، روش‌هایی وجود

با توجه به گسترش روزافزون مخاطرات جوی در کشور (وقایع حدی اقلیمی ایران در سال ۲۰۱۸ را در خزانه‌داری و همکاران (۲۰۲۰) ببینید)، پیش‌بینی دقیق ماهانه و فصلی بارش و نیز دما، می‌تواند به تصمیم‌گیران برای برنامه‌ریزی بهتر برای آینده کمک کند. در این بین، امروزه از روش‌های یادگیری ماشین به‌صورت گسترده در پیش‌بینی دما و بارش استفاده می‌شود. برای این منظور، خروجی مدل‌های آب‌وهوایی با کمک داده‌های مشاهداتی و روش‌های یادگیری ماشین پس‌پردازش شده و پیش‌بینی دقیق‌تری از دما و بارش (و یا سایر متغیرهای آب‌وهوایی) ارائه می‌شود. در این بین از روش‌های مبتنی بر شبکه‌های عصبی مصنوعی پرسپترون چندلایه، به‌صورت گسترده استفاده می‌شود. به‌عنوان مثال پاکدامن و همکاران

دارد ولی در بیشتر مواقع باید با آزمون و خطا یک معماری مناسب برای مسئله خاص مورد بررسی، پیدا کرد. یکی از مراحل مهم در استفاده از روش‌های یادگیری ماشین (در حالت کلی) و به‌خصوص استفاده از روش شبکه عصبی مصنوعی پرسپترون، مرحله آموزش است. طی فرایند آموزش شبکه عصبی که در واقع منجر به حل یک مسئله بهینه‌سازی ریاضی می‌شود، وزن‌های بهینه شبکه، به‌عنوان پارامترهای قابل تنظیم آن، محاسبه می‌شوند.

آزادی و همکاران (۱۳۹۱) یک سامانه همادی با استفاده از مدل‌های میان‌مقیاس MM5 و WRF با پیکربندی‌های گوناگون و تغییر در فیزیک آنها ساختند. سپس برون‌داد هر عضو (به‌صورت جداگانه) در سامانه همادی برای کمیت بارش با استفاده از یک شبکه عصبی مصنوعی مورد پس‌پردازش قرار گرفت و در نهایت با استفاده از روش بافت‌نگار رتبه‌ای پیش‌بینی‌های احتمالاتی بارش برای آستانه‌های گوناگون بارش برای منطقه شمال ایران محاسبه و واسنجی شد. صالح‌وند و همکاران (۱۳۹۹) برای مدل‌سازی بارش از داده‌های بارش ماهانه ایستگاه همدید و باران‌سنجی در بازه آماری (۲۰۱۴-۱۹۸۴) به مدت ۳۰ سال به‌عنوان متغیر وابسته و شاخص‌های اقلیمی، سیگنال‌های بزرگ‌مقیاس اقلیمی شامل؛ دمای سطح دریا و دمای ۱۰۰۰ میلی‌باری، هم ارتفاع تراز ۵۰۰ میلی‌باری، امگای ۲۰۰ میلی‌باری و عناصر اقلیمی به‌عنوان متغیر مستقل استفاده کردند. آنها به‌دلیل رفتار غیرخطی بارش از شبکه‌های عصبی مصنوعی جهت مدل‌سازی بهره گرفتند و جهت مشخص‌شدن بهترین معماری برای ورود به شبکه عصبی از تحلیل عاملی استفاده کردند. از روش‌های یادگیری ماشین، تنها برای پیش‌بینی استفاده نمی‌شود. به‌عنوان مثال پاکدامن و همکاران (۲۰۲۰) یک روش مبتنی بر یادگیری ماشین برای حل معادلات دیفرانسیل توازن انرژی جو در حالت فازی ارائه کردند. همچنین پاکدامن و همکاران (۲۰۲۱) رویکرد جدیدی در مدل‌سازی آلوده ارائه کردند.

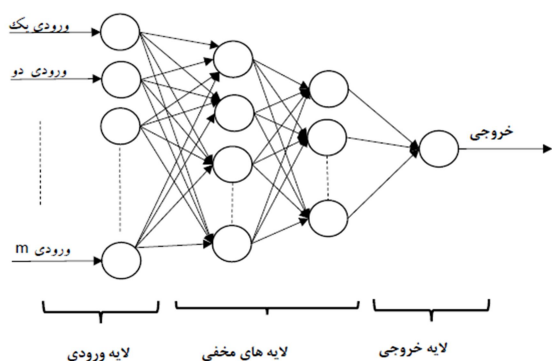
رحیمی نسب و عامریان (۱۳۹۸) با پیش‌بینی بارش ماهانه در ایران با استفاده از روش جدید ترکیب شبکه‌های عصبی مصنوعی و فیلتر کالمن توسعه یافته، توانستند با استفاده از داده‌های میانگین بارش ماهانه حدود ۱۸۰ ایستگاه سینوپتیک ایران که در سراسر کشور پراکنده هستند، و طی سال‌های ۱۹۵۱ تا ۲۰۱۶ استفاده شده بودند، به پیش‌بینی بارش ماهانه برای سال ۲۰۱۷ پردازند. عندلیب و همکاران (۱۳۹۸) پیش‌بینی چندایستگاه رواناب را با استفاده از تبدیل موجک و شبکه عصبی مصنوعی خودسازمانده و مدل‌های هوش مصنوعی انجام دادند. مدل‌سازی چند ایستگاه بارش-رواناب بر اساس خاصیت فصلی بودن انجام شد و با مدل‌سازی چند ایستگاه بر اساس خاصیت مارکف مقایسه شد. نجیب‌زاده و همکاران (۱۳۹۸) مدل‌هایی برای شبیه‌سازی فرایند بارش رواناب در زیر حوضه سد صفا رود واقع در حوضه آبریز هلیل رود با استفاده از روش‌های شبکه عصبی مصنوعی پرسپترون چند لایه (MLP)، شبکه عصبی تابع شعاع مدار (RBF)، رگرسیون ماشین بردار پشتیبان با توابع کرنل خطی (SVR Linear) و پایه شعاعی (SVR RBF) در محیط نرم‌افزاری IBM SPSS Modeler توسعه دادند. بدین‌منظور از داده‌های هیدرومتری ایستگاه هنجان و داده‌های بارش ایستگاه‌های هنجان، رابر، چشمه عروس و میدان واقع در منطقه مورد مطالعه طی دوره آماری ۱۳۹۴-۱۳۸۵ در مقیاس روزانه استفاده کردند. شافعی‌زاده و همکاران (۱۳۹۸) به شبیه‌سازی پیوسته بارش-رواناب در حوضه سد مارون با شبکه‌های عصبی مصنوعی پرداختند تا توانایی و دقت این شبکه در برآورد رواناب نیز ارزیابی کنند. در این مطالعه سعی بر ارائه یک مدل بارش-رواناب دو ضابطه‌ای شامل مدل مربوط به روزهای بارانی و مدل مربوط به روزهای غیربارانی بوده است. همچنین متغیرهای ورودی مؤثر در دبی جریان در حوضه مارون با استفاده از الگوریتم اطلاعات متقابل جزئی (PMI، Partial Mutual Information) تعیین شد. لی و همکاران (۲۰۱۸) یک مدل شبکه عصبی برای

شامل روش‌های منظم‌سازی بیزی، لونیبرگ-مارکوات و گرادیان مزدوج پاول-بل، الگوریتم شبه‌نیوتنی BFGS، گرادیان مزدوج فلچر-پاول و گرادیان مزدوج مقیاس‌شده برای پیش‌بینی ماهانه بارش کشور مورد بررسی و مقایسه قرار می‌گیرند. برای این منظور، یک شبکه عصبی پرسپترون برای پس پردازش خروجی بارش ماهانه مدل ECMWF طراحی می‌شود که برای آموزش آن از داده‌های ERA5 و روش‌های آموزش مختلف استفاده می‌شود تا علاوه بر پیش‌بینی ماهانه بارش کشور، مشخص شود، کدام‌یک از روش‌های آموزش شبکه عصبی، برای این موضوع از دقت بیشتری برخوردارند.

بخش دوم مقاله به بررسی روش‌های موردنیاز می‌پردازد و در نهایت در همین بخش، الگوریتم پیشنهادی ارائه می‌شود. بخش سوم نیز نتایج عددی حاصل از پیاده‌سازی روش پیشنهادی برای پیش‌بینی ماهانه بارش کشور با شبکه عصبی مصنوعی پرسپترون و به کمک روش‌های مختلف بهینه‌سازی ارائه می‌شود. بخش چهارم نیز شامل نتیجه‌گیری مبسوط خواهد بود.

## ۲. روش‌ها

مدل کلی یک شبکه عصبی مصنوعی پرسپترون چند لایه در شکل ۱ نشان داده شده است.



شکل ۱. نمای کلی از یک شبکه عصبی پرسپترون چندلایه.

هریک از نرون‌ها شامل یک تابع محرک هستند. توابع محرک مختلفی وجود دارد که با توجه به ساختار مسئله و نیز ساختار مدنظر برای شبکه عصبی از آنها استفاده می‌شود. یکی از توابع مرسوم در طراحی شبکه‌های عصبی

پیش‌بینی بارش در کره جنوبی توسعه دادند. در ابتدا آنها یازده ورودی برای شبکه عصبی پیشنهادی در نظر گرفتند؛ اما با بررسی همبستگی، در نهایت بهترین شبکه عصبی پیشنهادی آنها مدلی با ۵ ورودی بود که از نظر مقدار ریشه میانگین مربعات خطا، کمترین مقدار را داشت. ذباح و همکاران (۲۰۱۸) کارایی متفاوت شبکه‌های عصبی را به‌منظور پیش‌بینی بارش بررسی کردند و در نهایت الگوریتمی به‌منظور ترکیب روش‌های شبکه‌های عصبی خطی و غیرخطی پیشنهاد دادند. احمدی (۲۰۱۱) مدلی را بر اساس یک شبکه عصبی پیش‌خور که به‌وسیله الگوریتم رقابتی بهینه شده بود، پیشنهاد کرد.

فن و همکاران (۲۰۲۱) از تکنیک‌های شبکه عصبی برای مدل‌سازی و بهبود پیش‌بینی بارش و دمای مدل CFSv2 استفاده کردند. آنها نشان دادند که نتایج مربوط به شبکه عصبی مصنوعی در مقایسه با روش‌های رگرسیونی چندگانه عملکرد بهتری دارند. انگوین و همکاران (۲۰۲۱) از روش شبکه عصبی مصنوعی با آموزش از طریق روش پس‌انتشار خطا، برای پیش‌بینی بارش در مقیاس ماهانه و روزانه استفاده کردند. در این مطالعه مشخص شد عملکرد شبکه عصبی در مقیاس ماهانه بهتر از مقیاس روزانه است. آنوچی و ولهو (۲۰۲۰) نیز از شبکه عصبی پرسپترون چند لایه برای پیش‌بینی ماهانه بارش در

برزیل استفاده کردند. لیو و همکاران (۲۰۱۹) از شبکه عصبی پس‌انتشار خطا برای پیش‌بینی کوتاه‌مدت بارش استفاده کردند. آنها برای تنظیم معماری شبکه عصبی از قضیه کولموگروف و نیز آزمون و خطا استفاده کردند. همان‌طور که در مقالات و پیشینه فوق بررسی شد، نویسندگان از روش‌های مختلف برای آموزش شبکه عصبی با معماری‌های گوناگون استفاده کردند. لذا، علاوه بر اهمیت پیش‌بینی ماهانه بارش کشور (که در این مقاله از طریق پیش‌پردازش خروجی مدل ECMWF انجام می‌شود) انتخاب روش‌های کارا تر در آموزش شبکه عصبی نیز بسیار حائز اهمیت است. در این مقاله، شش روش مختلف آموزش شبکه عصبی پرسپترون چند لایه

مصنوعی، تابع زیگموئید است:

$$\sigma(x) = \frac{1}{1+e^{-x}} \quad (۱)$$

دلیل استفاده از این تابع، قدرت تقریب‌زنی مناسب آن است که در قضیه تقریب جهانی شبکه‌های عصبی پرسپترون چند لایه نیز به آن اشاره شده است (سینکو، ۱۹۸۹).

از سوی دیگر، روش‌های متعددی برای آموزش شبکه‌های عصبی ارائه شده است. عموماً در آموزش شبکه‌های عصبی، با یک مسئله بهینه‌سازی ریاضی مواجه هستیم که باید با حل آن، مقادیر بهینه وزن‌های شبکه عصبی که در واقع، پارامترهای قابل تنظیم آن هستند را به دست آورد. عموماً تابع هدف در این نوع مسائل بهینه‌سازی، یک تابع محدب نبوده و همواره روش‌هایی برای بهبود مقدار تابع هدف ابداع و ارائه می‌شود.

اکنون فرض کنید،  $N(x, \Lambda)$  نمایانگر خروجی شبکه عصبی پرسپترون چند لایه است که در آن  $x$  ورودی‌های شبکه عصبی و  $\Lambda$  برداری شامل تمام وزن‌های قابل تنظیم شبکه عصبی اعم از وزن‌های لایه ورودی و خروجی و لایه‌های میانی هستند. این وزن‌ها باید طوری تنظیم شوند تا حتی‌الامکان،  $N(x, \Lambda)$  به مقادیر مورد انتظار (در این مقاله میزان بارش ماهانه) نزدیک‌تر باشد. با فرض این که  $O(x)$  مقادیر موردانتظار متغیری که قرار است پیش‌بینی شود را نمایش دهد (در این مقاله، مقادیر مشاهداتی بارش ماهانه و یا داده‌های بازتحلیل)، در این صورت می‌توان مسئله بهینه‌سازی زیر را برای یافتن مقادیر بهینه وزن‌های  $\Lambda$  تعریف کرد:

$$\min_{\Lambda} E(\Lambda) = \sum_x [N(x, \Lambda) - O(x)]^2 \quad (۲)$$

که یک مسئله بهینه‌سازی غیرخطی بدون قید با تابع هدف  $E(\Lambda)$  است. با تعیین مقادیر بهینه وزن‌های  $\Lambda$  که آنها را با  $\Lambda^*$  نشان می‌دهیم،  $N(x, \Lambda^*)$  تقریب قابل‌قبولی از  $O(x)$  خواهد بود.

در این قسمت شش روش مختلف منظم‌سازی بی‌زی، لونیبرگ-مارکوات، گرادیان مزدوج پاول-بل، شبه‌نیوتنی

BFGS، گرادیان مزدوج فلچر-پاول و گرادیان مزدوج مقیاس شده به اختصار شرح داده می‌شوند. سپس از این روش‌ها برای پیش‌بینی بارش ماهانه کشور استفاده شده و نتایج گزارش می‌شود.

## ۲-۱. روش‌های آموزش شبکه عصبی

در روش‌های بهینه‌سازی ریاضی مبتنی بر مشتق و بردار گرادیان، مشتق مرتبه دوم تابع هدف، که ماتریس هسیان نامیده می‌شود و معکوس آن، در محاسبات نقش اساسی دارند. از سوی دیگر با افزایش تعداد متغیرها، اندازه ماتریس هسیان افزایش یافته و محاسبه معکوس آن از نظر محاسباتی زمانبر است. لذا در روش‌های بهبودیافته بهینه‌سازی، سعی می‌شود تا با تکنیک‌هایی، معکوس ماتریس هسیان تابع هدف تقریب زده شود. در روش لونیبرگ-مارکوات، ماتریس هسیان با رابطه زیر تخمین زده می‌شود (بازارا و همکاران، ۲۰۱۳):

$$H = J^T J \quad (۳)$$

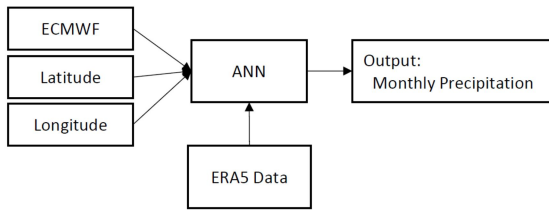
که در آن  $J$  ماتریس ژاکوبین شامل مشتقات نسبی مرتبه اول تابع خطای شبکه عصبی نسبت به وزن‌های  $\Lambda$  است. همچنین، در این الگوریتم، بردار گرادیان به صورت زیر محاسبه می‌شود:

$$g = J^T e \quad (۴)$$

که همان خطای شبکه عصبی است. ماتریس ژاکوبین نیز با کمک روش پس انتشار قابل محاسبه است. در این صورت طبق الگوریتم لونیبرگ-مارکوات، مقادیر وزن‌ها طی یک فرایند تکراری (شبه روش‌های نیوتنی) به صورت زیر به‌روزرسانی می‌شوند:

$$\Lambda_{k+1} = \Lambda_k - [J^T J + \lambda I]^{-1} g \quad (۵)$$

که در آن  $I$  ماتریس همانی و  $\lambda$  مقدار ثابتی است که باید تعیین شود. به‌عنوان مثال، وقتی  $\lambda$  صفر باشد، تکرارهای روش نیوتن به دست می‌آید (هاگان و منهاج، ۱۹۹۴).



شکل ۲. نمای کلی شبکه عصبی پیشنهادی که به شش روش مختلف مختلف منظم‌سازی بیزی، لونیگ-مارکوات، گرادیان مزدوج پاول-بل، شبه‌نیوتنی BFGS، گرادیان مزدوج فلچر-پاول و گرادیان مزدوج مقیاس شده آموزش می‌بیند.

از آنجا که هدف این مقاله مقایسه روش‌های مختلف آموزش شبکه عصبی است، لذا برای تمام آزمایشات، یک شبکه عصبی پرسپترون با یک لایه پنهان و تابع محرک زیگموئید (معادله ۱) در نظر گرفته شد. دلیل انتخاب تابع محرک زیگموئید، قضیه تقریب جهانی (سینکو، ۱۹۸۹) شبکه عصبی پرسپترون چند لایه است. این تابع توانایی تقریب بالایی را به شبکه عصبی می‌دهد. همچنین برای تعیین تعداد نرون‌های لایه پنهان از قضیه کولموگروف و نیز روش تجربی استفاده شد (لیو و همکاران، ۲۰۱۹).

سپس در شش حالت مختلف استفاده از الگوریتم‌های آموزشی، نتایج مربوطه محاسبه شد. برای ارزیابی عملکرد شبکه‌های عصبی مختلف، از سه شاخص ضریب همبستگی، میانگین مربعات خطا و نیز شاخص نش-ساتکلیف استفاده شد. روابط مربوط به این شاخص‌ها به ترتیب به صورت زیر است (شافعی‌زاده و همکاران، ۱۳۹۸):

$$R = \frac{\sum_{i=1}^n (X_i - \bar{X})(O_i - \bar{O})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2 \sum_{i=1}^n (O_i - \bar{O})^2}} \quad (6)$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (X_i - O_i)^2 \quad (7)$$

$$NS = 1 - \frac{\sum_{i=1}^n (X_i - O_i)^2}{\sum_{i=1}^n (O_i - \bar{O})^2} \quad (8)$$

که در روابط فوق  $X_i$  و  $O_i$  به ترتیب، نمایانگر  $i$ -امین مؤلفه بردارهای مدل و مشاهدات است. همچنین  $\bar{X}$  و  $\bar{O}$  به ترتیب، نشان‌دهنده میانگین مقادیر مدل و مشاهدات هستند. مقدار  $n$  نیز تعداد مشاهدات را نشان می‌دهد.

منظم‌سازی بیزی، ترکیبی از روش لونیگ-مارکوات به همراه کمینه‌سازی مضربی از وزن‌ها برای جلوگیری از افزایش بی‌رویه مقادیر آنها در طی تکرارهاست. در واقع این الگوریتم علاوه بر کمینه‌سازی خطا، به دنبال کمینه‌سازی مربعات اوزان نیز می‌باشد (فورسی و هاگان، ۱۹۹۷).

روش گرادیان مزدوج پاول-بل برای زمانی که توابع محرک شبکه و خروجی شبکه مشتق‌پذیر باشد کاربرد دارد. روش‌های گرادیان مزدوج نیازمند تعیین جهت بردار گرادیان در بازه‌های خاص هستند. در این روش از روش‌های پیشنهادی پاول-بل برای آغاز دوباره الگوریتم استفاده می‌شود. جزئیات این الگوریتم در پاول (۱۹۷۷) آمده است.

روش شبه‌نیوتنی BFGS، یک روش تکراری برای حل مسائل بهینه‌سازی نامقید است که در تکرارهای مختلف تقریبی از ماتریس هسیان تابع هدف محاسبه می‌شود. در الگوریتم گرادیان مزدوج فلچر-پاول، برای محاسبه مشتق نسبی تابع هدف نسبت به وزن‌ها و بایاس شبکه، از روش پس انتشار استفاده می‌شود. برای به‌روزرسانی وزن‌ها از یک رابطه تکراری استفاده می‌شود. جهت جست‌وجو در تکرار اول، قرینه بردار گرادیان است. در تکرارهای بعدی، بردار گرادیان مجدداً در نقاط جدید محاسبه می‌شود. در روش گرادیان مزدوج مقیاس‌شده نیز عملکرد همانند روش گرادیان مزدوج است با این تفاوت که در هر تکرار نیاز به انجام جست‌وجوی خطی نیست (بازارا و همکاران، ۲۰۱۳).

## ۲-۲. روش پیشنهادی

با وجود پیشنهاد استفاده از چند روش آموزش برای شبکه عصبی مصنوعی، می‌توان، نمای کلی روش پیشنهادی را در شکل ۲ ملاحظه کرد.

## ۳. نتایج عددی

با استفاده از مدل شبکه عصبی پیشنهادی بخش قبل، شش روش آموزش مختلف منظم‌سازی بیزی، لونیبرگ-مارکوات، گرادیان مزدوج پاول-بل، شبه‌نیوتنی BFGS، گرادیان مزدوج فلچر-پاول و گرادیان مزدوج مقیاس‌شده برای پس پردازش خروجی مدل ECMWF برای پیش‌بینی ماهانه بارش کشور مورد استفاده قرار گرفت. برای این منظور، داده‌های مدل ECMWF و نیز داده‌های بازتحلیل ERA5 برای دوره زمانی ۱۹۹۳ تا ۲۰۱۶ میلادی و برای شش زمان انتظار اول از وبسایت کوپرنیک به آدرس <https://cds.climate.copernicus.eu> برای منطقه جغرافیایی محدود به کشور ایران دریافت شد.

سپس، شبکه عصبی پیشنهادی با توجه به ساختار پیشنهادشده در بخش فوق، مورد استفاده قرار گرفت. برای این منظور داده‌های مربوط به دوره زمانی ۱۹۹۳ تا ۲۰۱۰ میلادی به‌عنوان داده‌های آموزش شبکه و داده‌های مربوط به دوره زمانی ۲۰۱۱ تا ۲۰۱۶ برای اعتبارسنجی به‌کار گرفته شد. دلیل استفاده از داده‌های بازتحلیل ERA5 اولاً تولید پیش‌بینی ماهانه به‌صورت شبکه‌ای بود و از سوی دیگر در مطالعات مختلف نشان داده شده است که این داده‌ها دقت قابل‌قبولی برای کشور دارند. با توجه به این که اخیراً، داده‌های ERA5 (نسل ۵ داده‌های بازتحلیل مرکز ECMWF) جدید در دسترس قرار گرفته‌اند و ارزیابی این داده‌ها برای منطقه مورد مطالعه، در مقالات مختلفی بررسی شده است (ایزدی و همکاران، ۲۰۲۱؛ تقی‌زاده و همکاران، ۲۰۲۱) که نشان‌دهنده دقت قابل‌قبول آنهاست، لذا در این تحقیق از این داده‌ها استفاده شد. شایان ذکر است که داده‌های ERA5 برای دوره زمانی از ۱۹۷۹ تاکنون قابل دریافت است و داده‌های دوره زمانی ۱۹۵۰ تا ۱۹۷۸ نیز به‌صورت آزمایشی در دسترس هستند. تفکیک مکانی داده‌های اصلی ERA5 حدود ۳۰ کیلومتر مربع و تا ارتفاع ۸۰ کیلومتری سطح زمین در ۱۳۷ سطح مختلف را شامل می‌شود. مؤسسه کوپرنیک این

داده‌ها را با تفکیک مکانی و زمانی مختلف در اختیار کاربران قرار داده است. از آنجاکه داده‌های ECMWF با تفکیک یک درجه هستند، داده‌های ERA5 نیز با همین تفکیک از سایت مؤسسه کوپرنیک دریافت شد.

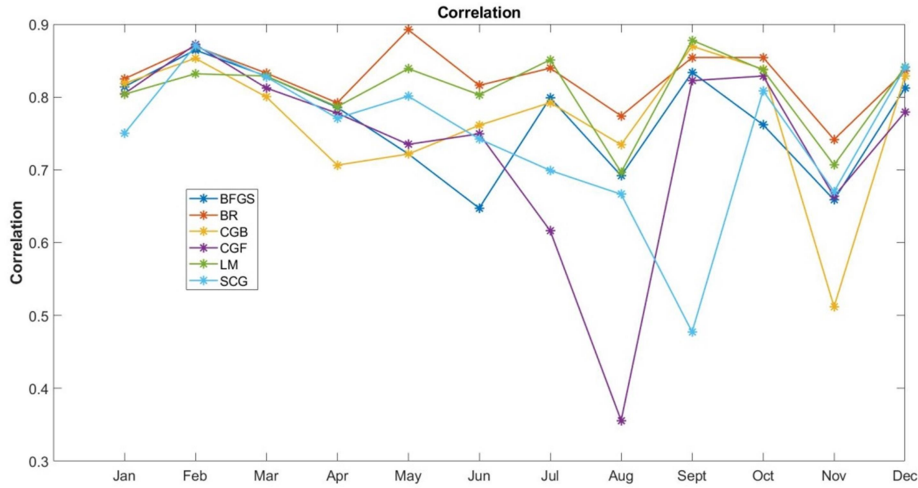
از آنجاکه مدل ECMWF دارای شش زمان انتظار مختلف است، لذا می‌توان برای ۱۲ ماه مختلف سال، ۷۲ مدل متفاوت پیشنهاد داد. در این‌صورت شبکه عصبی  $ANN(m, l)$  مدلی است که بر مبنای داده‌های ماه  $m$  و زمان تأخیر  $l$  آموزش دیده است.

در این مقاله، پس از آموزش مدل‌های مختلف، برای زمان‌های انتظار مختلف و ماه‌های متفاوت، در نهایت، مقدار سه شاخص ضریب همبستگی، میانگین مربعات خطا و نیز شاخص نش-ساتکلیف برای هر یک از ۱۲ ماه سال و برای داده‌های آزمایش ۲۰۱۶-۲۰۱۱ محاسبه شد که در شکل‌های ۳، ۴ و ۵ به تصویر کشیده شده است.

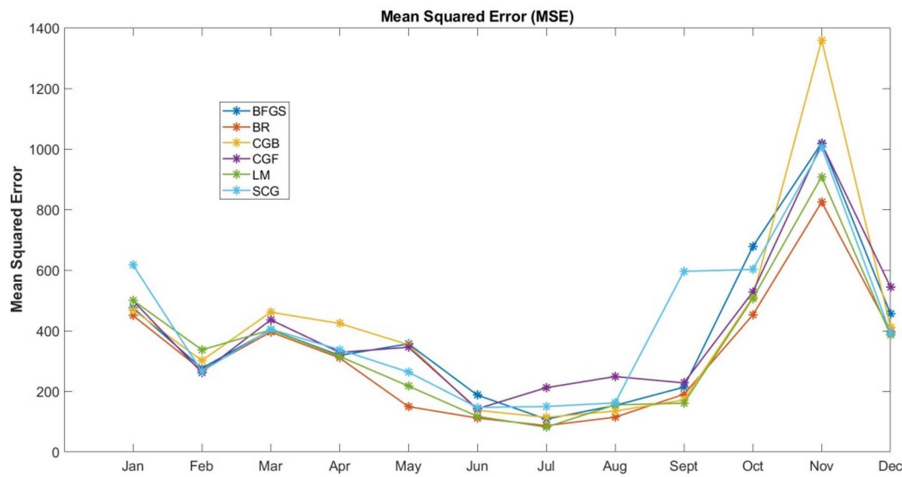
با توجه به شکل ۳، به‌طور کلی می‌توان اذعان کرد که مقدار همبستگی خروجی شبکه عصبی آموزش دیده شده با دو روش منظم‌سازی بیزی و لونیبرگ-مارکوات، با داده‌های بازتحلیل ERA5 از سایر روش‌ها بیشتر است. برای مشاهده مقادیر ضریب همبستگی به‌صورت دقیق، مقادیر عددی شکل ۳ در جدول ۱ درج شده است.

از نظر مقدار میانگین مربعات خطا، با توجه به شکل ۴، کمترین مقدار مربوط به دو روش منظم‌سازی بیزی و لونیبرگ-مارکوات است. شکل ۵ نیز شاخص نش-ساتکلیف را نشان می‌دهد. با توجه به شکل ۵ ملاحظه می‌شود، هرچند برای ماه‌های مختلف، الگوریتم‌های مختلفی عملکرد بهتری دارند؛ اما در مجموع، دو روش مذکور در این شاخص نیز از سایر روش‌های آموزش شبکه عصبی، بهتر عمل کرده‌اند.

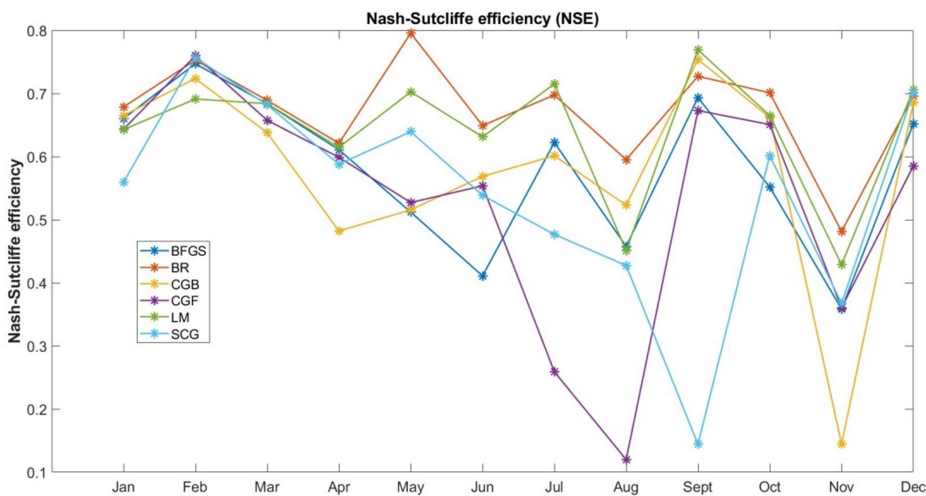
بنابراین می‌توان نتیجه گرفت که در مجموع، عملکرد دو روش منظم‌سازی بیزی و لونیبرگ-مارکوات، در دوره آزمایش ۲۰۱۶-۲۰۱۱ از دیگر روش‌های آموزش شبکه عصبی بهتر است.



شکل ۳. ضریب همبستگی برای ماه‌های مختلف دوره زمانی اعتبارسنجی و برای شش روش آموزش منظم‌سازی بیزی (BR)، لونیبرگ-مارکوات (LM)، گرادیان مزدوج پاول-بیل (CGB)، شبه‌نیوتنی BFGS، گرادیان مزدوج فلچر-پاول (CGF) و گرادیان مزدوج مقیاس‌شده (SCG).



شکل ۴. میانگین مربعات خطا برای ماه‌های مختلف دوره زمانی اعتبارسنجی و برای شش روش آموزش منظم‌سازی بیزی (BR)، لونیبرگ-مارکوات (LM)، گرادیان مزدوج پاول-بیل (CGB)، شبه‌نیوتنی BFGS، گرادیان مزدوج فلچر-پاول (CGF) و گرادیان مزدوج مقیاس‌شده (SCG).



شکل ۵. ضریب نش-ساتکلیف برای ماه‌های مختلف دوره زمانی اعتبارسنجی و برای شش روش آموزش منظم‌سازی بیزی (BR)، لونیبرگ-مارکوات (LM)، گرادیان مزدوج پاول-بیل (CGB)، شبه‌نیوتنی BFGS، گرادیان مزدوج فلچر-پاول (CGF) و گرادیان مزدوج مقیاس‌شده (SCG).

جدول ۱. ضریب همبستگی برای ماه‌های مختلف دوره زمانی اعتبارسنجی و برای شش روش آموزش منظم‌سازی بیزی (BR)، لونیبرگ-مارکوات (LM)، گرادیان مزدوج پاول-بل (CGB)، شبه‌نیوتنی (BFGS)، گرادیان مزدوج فلچر-پاول (CGF) و گرادیان مزدوج مقیاس شده (SCG).

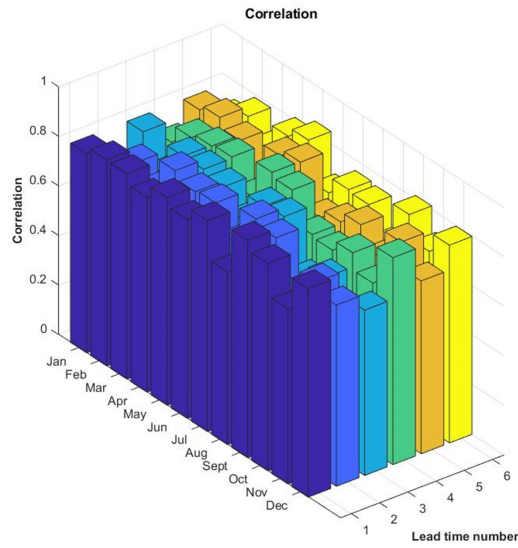
الگوریتم آموزش شبکه عصبی						ماه
BFGS	BR	CGB	CGF	LM	SCG	
۰/۸۱۴	۰/۸۲۵	۰/۸۱۹	۰/۸۰۵	۰/۸۰۴	۰/۷۵	۱
۰/۸۶۵	۰/۸۶۹	۰/۸۵۳	۰/۸۷۲	۰/۸۳۲	۰/۸۷	۲
۰/۸۲۹	۰/۸۳۳	۰/۸	۰/۸۱۲	۰/۸۲۹	۰/۸۲۸	۳
۰/۷۸۶	۰/۷۹۲	۰/۷۰۶	۰/۷۷۸	۰/۷۸۷	۰/۷۷۱	۴
۰/۷۲۲	۰/۸۹۲	۰/۷۲۲	۰/۷۳۵	۰/۸۳۹	۰/۸۰۱	۵
۰/۶۴۷	۰/۸۱۶	۰/۷۶۱	۰/۷۴۹	۰/۸۰۳	۰/۷۴۲	۶
۰/۷۹۹	۰/۸۴	۰/۷۹۲	۰/۶۱۶	۰/۸۵۱	۰/۶۹۹	۷
۰/۶۹۲	۰/۷۷۴	۰/۷۳۵	۰/۳۵۶	۰/۶۹۷	۰/۶۶۶	۸
۰/۸۳۴	۰/۸۵۴	۰/۸۷	۰/۸۲۲	۰/۸۷۸	۰/۴۷۷	۹
۰/۷۶۲	۰/۸۵۴	۰/۸۳۸	۰/۸۲۹	۰/۸۳۷	۰/۸۰۹	۱۰
۰/۶۵۹	۰/۷۴۲	۰/۵۱۲	۰/۶۶۴	۰/۷۰۷	۰/۶۷	۱۱
۰/۸۱۳	۰/۸۳۶	۰/۸۲۹	۰/۷۷۹	۰/۸۴۱	۰/۸۴	۱۲

این نکته حائز اهمیت است که زمان انتظار اول برای تمام ماه‌ها، داده مربوط به همان ماه است. زمان انتظار دوم برای هر ماه، داده مربوط به ماه بعد است و به همین ترتیب برای سایر زمان‌های انتظار. همچنین دوره زمانی شکل‌ها مربوط به دوره اعتبارسنجی است. با توجه به شکل‌های ۶، ۷ و ۸ طی ماه‌ها و زمان‌های انتظار مختلف، عملکرد شبکه عصبی آموزش دیده با روش لونیبرگ-مارکوات (LM)، متفاوت است. همان‌طور که انتظار می‌رود، در زمان‌های انتظار دورتر که از زمان ابتدایی پیش‌بینی، با توجه به ضریب همبستگی، میانگین مربعات خطا و نیز شاخص نش-ساتکلیف، عملکرد ضعیف‌تر است. با این وجود، در برخی از موارد با وجود افزایش زمان انتظار، عملکرد شبکه عصبی بهبود یافته است. به‌عنوان مثال، با توجه به شکل‌های ۶، ۷ و ۸ در زمان انتظار دوم ماه نوامبر، مقادیر ضریب همبستگی، میانگین مربعات خطا و نیز شاخص نش-ساتکلیف، بهتر از زمان انتظار اول همان ماه است.

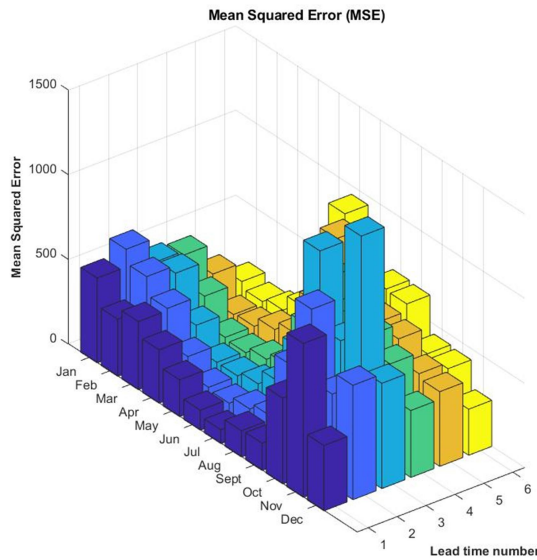
از آنجاکه در روش پیشنهادی این مقاله، برای هر یک از ۱۲ ماه سال و برای هر یک از شش زمان انتظار هر ماه سال، یک مدل شبکه عصبی مجزا توسط شش روش آموزش مختلف ایجاد شده است، که به‌عنوان نمونه، شکل‌های ۳، ۴ و ۵ مربوط به زمان انتظار اول ارائه شد، و از سوی دیگر با توجه به این سه شکل، عملکرد دو روش منظم‌سازی بیزی و لونیبرگ-مارکوات نسبت به سایر روش‌های آموزش شبکه عصبی بهتر بود، برای بررسی عملکرد این دو روش در سایر زمان‌های انتظار نیز مقادیر سه شاخص ضریب همبستگی، میانگین مربعات خطا و نیز شاخص نش-ساتکلیف برای این دو روش آموزش، و برای تمام زمان‌های انتظار و تمام ماه‌های سال، در شکل‌های ۶ تا ۱۱ محاسبه شده است.

شکل‌های ۶، ۷ و ۸ به ترتیب مقادیر سه شاخص ضریب همبستگی، میانگین مربعات خطا و نیز شاخص نش-ساتکلیف را برای روش لونیبرگ-مارکوات (LM) و برای ماه‌ها و زمان‌های انتظار مختلف نشان می‌دهند. یادآوری

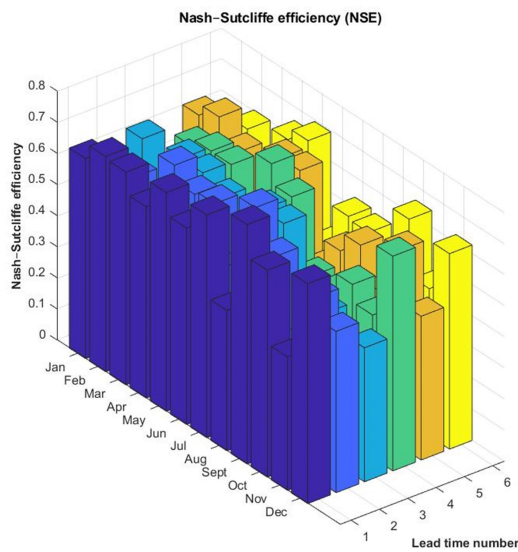




شکل ۶. ضریب همبستگی برای ماه‌های مختلف دوره زمانی اعتبارسنجی و برای زمان انتظارهای متفاوت روش لونیبرگ-مارکوات (LM).



شکل ۷. میانگین مربعات خطا برای ماه‌های مختلف دوره زمانی اعتبارسنجی و برای زمان انتظارهای متفاوت روش لونیبرگ-مارکوات (LM).

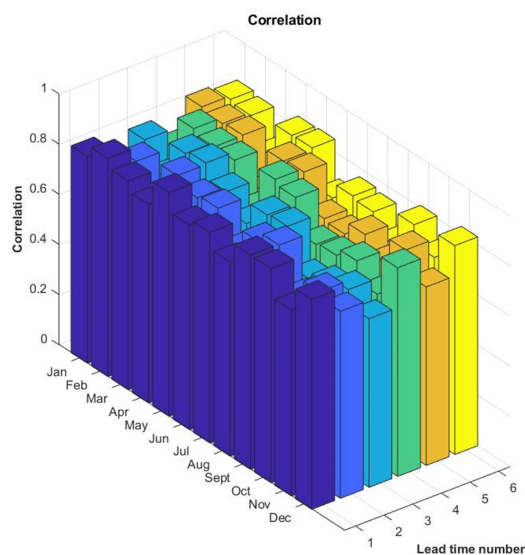


شکل ۸. ضریب نش-ساتکلیف برای ماه‌های مختلف دوره زمانی اعتبارسنجی و برای زمان انتظارهای متفاوت روش لونیبرگ-مارکوات (LM).

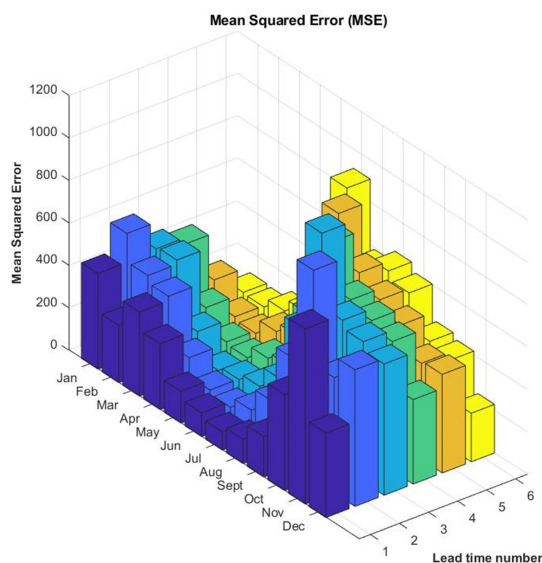
مشاهده می‌شود که با وجود افزایش زمان انتظار، برخلاف تصور، عملکرد روش منظم‌سازی بیزی در برخی موارد بهتر نیز می‌باشد. به‌عنوان مثال، با توجه به شکل‌های ۹، ۱۰ و ۱۱، در زمان انتظار ششم از ماه دسامبر، به‌صورت تقریبی، مقدار هر سه شاخص بهتر از زمان انتظار اول همان ماه است.

برای نشان دادن عملکرد الگوریتم‌های پیشنهادی، می‌توان توانایی آنها را بر مبنای مقدار شاخص‌های ارزیابی عملکرد، در دوره اعتبارسنجی ۲۰۱۱-۲۰۱۶ روی نقشه جغرافیایی کشور ترسیم کرد تا مشخص شود در مناطق مختلف جغرافیایی کشور، عملکرد هر یک چگونه است.

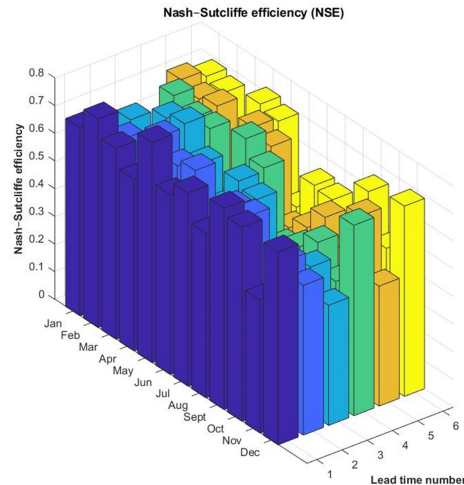
به‌طور مشابه، شکل‌های ۹، ۱۰ و ۱۱ به‌ترتیب مقادیر سه شاخص ضریب همبستگی، میانگین مربعات خطا و نیز شاخص نش-ساتکلیف را برای روش منظم‌سازی بیزی و برای ماه‌ها و زمان‌های انتظار مختلف نشان می‌دهند. با توجه به شکل‌های ۹، ۱۰ و ۱۱، طی ماه‌ها و زمان‌های انتظار مختلف، عملکرد شبکه عصبی آموزش دیده با روش منظم‌سازی بیزی، متفاوت است. با افزایش زمان‌های انتظار، که از زمان ابتدایی پیش‌بینی دورتر هستند، عملکرد در هر سه شاخص ضریب همبستگی، میانگین مربعات خطا و نیز شاخص نش-ساتکلیف ضعیف‌تر می‌شود. اما در برخی از موارد نیز



شکل ۹. ضریب همبستگی برای ماه‌های مختلف دوره زمانی اعتبارسنجی و برای زمان انتظارهای متفاوت روش منظم‌سازی بیزی (BR).



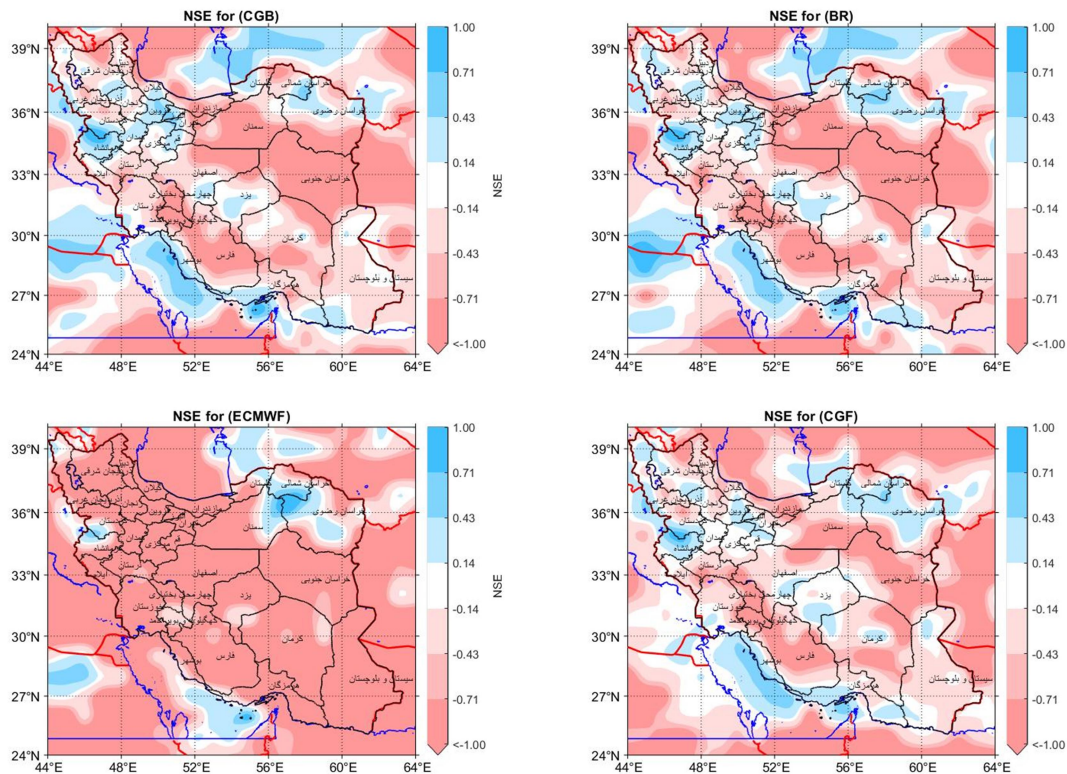
شکل ۱۰. میانگین مربعات خطا برای ماه‌های مختلف دوره زمانی اعتبارسنجی و برای زمان انتظارهای متفاوت روش منظم‌سازی بیزی (BR).

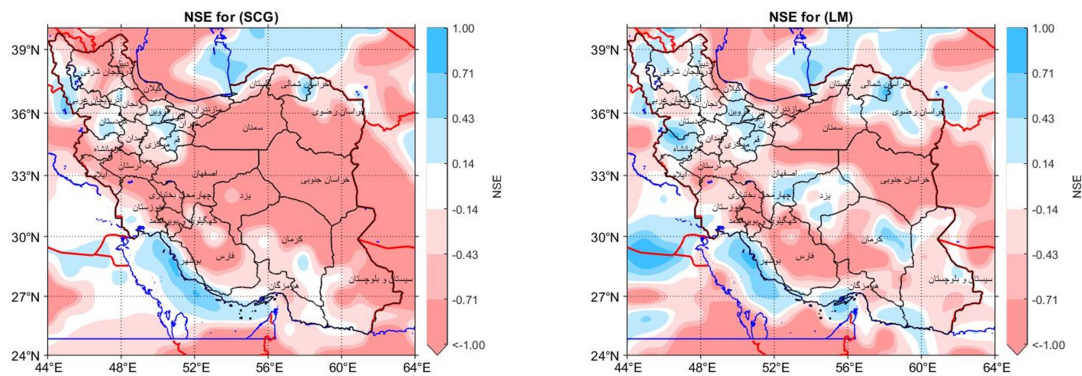


شکل ۱۱. نش-ساتکلیف برای ماه‌های مختلف دوره زمانی اعتبارسنجی و برای زمان انتظارهای متفاوت روش منظم‌سازی بیزی (BR).

است (به‌طور کلی و برای تمام مناطق کشور). این موضوع در شکل ۱۲ نیز مشخص است. پس از روش منظم‌سازی بیزی (BR)، روش گرادیان مزدوج پاول-بل (CGB)، از نظر مقدار شاخص نش-ساتکلیف، در رتبه دوم قرار می‌گیرد. از نظر مقدار شاخص نش-ساتکلیف، روش BR (و البته روش CGB)، در مقایسه با مدل پس‌پردازش‌نشده ECMWF توانسته است، به‌خصوص برای نیمه غربی کشور، باعث بهبود شود.

به‌عنوان نمونه، شکل ۱۲ مقدار شاخص نش-ساتکلیف را در مناطق مختلف جغرافیایی کشور برای هر یک از روش‌های پیشنهادی نشان می‌دهد. در این شکل، همچنین عملکرد مدل ECMWF نیز مشخص شده است. این شکل به‌عنوان نمونه، مربوط به ماه‌های ژانویه دوره اعتبارسنجی و در زمان انتظار اول است. همان‌طور که در شکل ۵ مشخص است، روش آموزش منظم‌سازی بیزی (BR) در ماه ژانویه بیشترین مقدار ضریب NSE را کسب کرده





شکل ۱۲. ضریب نش-ساتکلیف برای ماه‌های ژانویه دوره زمانی اعتبارسنجی و برای شش روش آموزش منظم‌سازی بیزی (BR)، لونیبرگ-مارکوات (LM)، گرادیان مزدوج پاول-بل (CGB)، شبه‌نیوتنی BFGS، گرادیان مزدوج فلچر-پاول (CGF) و گرادیان مزدوج مقیاس شده (SCG) و برای زمان انتظار اول.

روش منظم‌سازی بیزی و لونیبرگ-مارکوات، در مقایسه با چهار روش دیگر برای آموزش شبکه عصبی بهتر است. به این معنا که دو روش منظم‌سازی بیزی و لونیبرگ-مارکوات، توانایی بیشتری (در تعداد بیشتری از ماه‌های مختلف سال از بقیه الگوریتم‌ها بهتر عمل می‌کنند) در پیش‌بینی دارند. این موضوع می‌تواند به محققین کمک کند تا در هنگام ساخت معماری شبکه عصبی دلخواه خود، برای پیش‌بینی بارش، استفاده از این دو الگوریتم را برای آموزش شبکه عصبی در اولویت قرار دهند.

#### مراجع

آزادی، م.، واشانی، س. و حجام، س.، ۱۳۹۱، پیش‌بینی احتمالاتی بارش با استفاده از پس‌پردازش برون‌داد یک سامانه همادی، م. فیزیک زمین و فضا، ۳۸(۳)، ۲۰۳-۲۱۶.

صالحوند، ا.، گندمکار، ا. و فتاحی، ا.، ۱۳۹۹، پیش‌آگاهی بلندمدت بارش با استفاده از سیستم شبکه عصبی مطالعه موردی: حوضه کارون بزرگ، م. تحقیقات کاربردی علوم جغرافیایی (علوم جغرافیایی)، ۲۰(۵۹)، ۸۱-۹۷.

رحیمی نسب، م. و عامریان، ی.، ۱۳۹۸، پیش‌بینی بارش ماهانه در منطقه ایران با استفاده از ترکیب شبکه‌های عصبی مصنوعی و فیلتر کالمن توسعه یافته، م.

#### ۴. نتیجه‌گیری

در این مقاله، علاوه بر پیش‌بینی بارش ماهانه کشور از طریق پس‌پردازش خروجی مدل ECMWF به کمک شبکه عصبی مصنوعی، الگوریتم‌های مختلف آموزش شبکه عصبی در پیش‌بینی ماهانه بارش کشور مورد آزمایش قرار گرفت.

در این مقاله، شش روش مختلف آموزش شبکه عصبی پرسپترون چند لایه شامل روش‌های منظم‌سازی بیزی، لونیبرگ-مارکوات و گرادیان مزدوج پاول-بل، الگوریتم شبه‌نیوتنی BFGS، گرادیان مزدوج فلچر-پاول و گرادیان مزدوج مقیاس‌شده برای پیش‌بینی ماهانه بارش کشور مورد بررسی و مقایسه قرار گرفت. برای این منظور، یک شبکه عصبی پرسپترون برای پس‌پردازش خروجی بارش ماهانه مدل ECMWF طراحی شد که برای آموزش آن از داده‌های ERA5 و روش‌های آموزش مختلف استفاده شد. برای این منظور، تعداد ۷۲ مدل مختلف شبکه عصبی، برای هر یک از ۱۲ ماه سال و هر یک از شش زمان انتظار، ارائه شد که با شش روش مختلف آموزش، وزن‌های آن بهینه‌سازی شدند.

برای بررسی عملکرد شش روش آموزش شبکه عصبی، مقدار سه شاخص ضریب همبستگی، میانگین مربعات خطا و نیز شاخص نش-ساتکلیف برای هر مدل محاسبه شد. نتایج حاصل از مقایسه شاخص‌ها نشان داد که عملکرد دو

- اطلاعات جغرافیایی، ۲۸(۱۱۰)، ۷۷-۹۰.
- عندلیب، غ.، نورانی، و.، منیری فر، ح. و شرقی، ا.، ۱۳۹۸، توسعه روش ترکیبی موجک-شبکه عصبی مصنوعی و ماشین بردار پشتیبان برای مدل‌سازی چند ایستگاه بارش-رواناب با استفاده از ابزارهای خوشه‌بندی و اطلاعات مشترک، م. رویکردهای نوین در مهندسی عمران، ۳(۲)، ۴۹-۶۲.
- نجیب‌زاده، ن.، و قادری، ک. و احمدی، م. ۱۳۹۸، بهره‌گیری از روش‌های رگرسیون بردار پشتیبان و شبکه
- Pakdaman, M., Falamarzi, Y., Babaeian, I. and Javanshiri, Z., 2020a, Post-processing of the North American multi-model ensemble for monthly forecast of precipitation based on neural network models. *Theoretical and Applied Climatology*, 141(1), 405-417.
- Pakdaman, M., Naghab, S. S., Khazanedari, L., Malbousi, S. and Falamarzi, Y., 2020b, Lightning prediction using an ensemble learning approach for northeast of Iran. *Journal of Atmospheric and Solar-Terrestrial Physics*, 209, 105417.
- Pakdaman, M., Falamarzi, Y., Yazdi, H. S., Ahmadian, A., Salahshour, S. and Ferrara, F., 2020c, A kernel least mean square algorithm for fuzzy differential equations and its application in earth's energy balance model and climate. *Alexandria Engineering Journal*, 59(4), 2803-2810.
- Pakdaman, M., Habibi Nokhandan, M. and Falamarzi, Y., 2021, Revisiting albedo from a fuzzy perspective, *Kybernetes*, <https://doi.org/10.1108/K-12-2020-0834>.
- Cybenko, G., 1989, Approximation by superpositions of a sigmoidal function. *Mathematics of control, signals and systems*, 2(4), 303-314.
- Hagan, M. T. and Menhaj, M. B., 1994, Training feedforward networks with the Marquardt algorithm. *IEEE transactions on Neural Networks*, 5(6), 989-993.
- Powell, M. J. D., 1977, Restart procedures for the conjugate gradient method. *Mathematical programming*, 12(1), 241-254.
- Lee, J., Kim, C. G., Lee, J. E., Kim, N. W. and Kim, H., 2018, Application of artificial neural networks to rainfall forecasting in the Geum River basin, Korea. *Water*, 10(10), 1448.
- Zabbah, I., Roshani, A. R. and Khafage, A., 2018, Prediction of monthly rainfall using artificial neural network mixture approach, Case Study: Torbat-e Heydariyeh, *Journal of the Earth and Space Physics*, 44(4), 115-126.
- Ahmadi, M. A., 2011, Prediction of asphaltene precipitation using artificial neural network optimized by imperialist competitive algorithm. *Journal of Petroleum Exploration and Production Technology*, 1(2-4), 99-106.
- Khazanedari, L., Malbosi, S., Samadi Neghab, S., Pakdaman, M. and Javanshiri, Z., 2020, Extreme Climate Events in Iran during 2018. *Nivar*, 44(108-109), 68-78.
- Izadi, N., Karakani, E. G., Saadatabadi, A. R., Shamsipour, A., Fattahi, E. and Habibi, M., 2021, Evaluation of ERA5 Precipitation Accuracy Based on Various Time Scales over Iran during 2000-2018. *Water*, 13(18), 2538.
- Taghizadeh, E., Ahmadi-Givi, F., Brocca, L. and Sharifi, E., 2021, Evaluation of satellite/reanalysis precipitation products over Iran. *International Journal of Remote Sensing*, 42(9), 3474-3497.
- Foresee, F. D. and Hagan, M. T., 1997, Gauss-Newton approximation to Bayesian learning. In *Proceedings of international conference on neural networks (ICNN'97) (Vol. 3, pp. 1930-1935)*. IEEE.
- Bazaraa, M. S., Sherali, H. D. and Shetty, C. M., 2013, *Nonlinear programming: theory and algorithms*. John Wiley & Sons.
- Fan, Y., Krasnopolsky, V., van den Dool, H., Wu, C. Y. and Gottschalck, J., 2021, Using Artificial Neural Networks to Improve CFS Week 3-4 Precipitation and 2-Meter Air Temperature Forecasts. *Weather and Forecasting*.
- Nguyen, H. N., Nguyen, T. A., Ly, H. B., Tran, V. Q., Nguyen, L. K., Nguyen, M. V. and Ngo, C. T., 2021, Prediction of daily and monthly rainfall using a backpropagation neural network. *Journal of Applied Science and Engineering*, 24(3), 367-379.
- Anochi, J. A. and de Campos Velho, H. F., 2020, Neural network for seasonal climate precipitation prediction on the Brazil. *Ciência e Natura*, 42, 15.

Liu, Y., Zhao, Q., Yao, W., Ma, X., Yao, Y. and Liu, L., 2019, Short-term rainfall forecast model based on the improved Bp-nn algorithm. Scientific reports, 9(1), 1-12.

## The Effect of the type of training algorithm for multi-layer perceptron neural network on the accuracy of monthly forecast of precipitation over Iran, case study: ECMWF model

Pakdaman, M.\*

*Assistant Professor, Climate Research Institute, Atmospheric Science and Meteorological Research Center, Mashhad, Iran*

*(Received: 1 Aug 2021, Accepted: 10 Jan 2022)*

### Summary

Due to increasing atmospheric disasters in the Iran, accurate monthly and seasonal forecasts of rainfall as well as temperature, can help decision makers to better plan for the future. Meanwhile, machine learning methods are widely used today in predicting temperature and precipitation. For this purpose, the outputs of climate models are processed with the help of observational data and machine learning methods and a more accurate forecast of temperature and precipitation (or other climatic variables) are provided. In the meantime, methods based on multilayer perceptron artificial neural networks are widely used.

In a multi-layer perceptron artificial neural network, the design of the network architecture is very important and this design can directly affect the ability of the neural network to solve the problem. In designing network architecture, questions such as the number of neurons in each layer, the number of layers, network activity functions in each layer, etc. must be answered. In some cases, there are methods to answer each of the above questions, but in most cases, a suitable architecture for the specific problem under study must be found by trial and error. One of the important steps in using machine learning methods (in general) and especially the use of perceptron artificial neural network method is the training stage. During the neural network training process, which actually leads to solving a mathematical optimization problem, the optimal network weights are calculated as its adjustable parameters.

Today, various types of artificial neural networks are used in various fields of atmospheric science and climatology for purposes such as classification, regression and prediction. But the fundamental question in the use of artificial neural networks is how they are designed and built. One of the important points in using artificial neural networks that should be considered by designers is choosing the right algorithm for network training. In this paper, six different methods are considered for training a multilayer perceptron neural network including: Bayesian Regularization algorithm, Levenberg-Marquatt algorithm, Conjugate Gradient with Powell/Beale Restarts, BFGS Quasi-Newton algorithm, Scaled Conjugate Gradient and Fletcher-Powell Conjugate Gradient methods for monthly forecasting of precipitation that are reviewed and compared. In mathematical optimization methods based on derivatives and gradient vectors, the second-order derivative of the objective function, called the Hessian matrix, and its inverse, play an essential role in the calculations. On the other hand, with increasing the number of variables, the size of the matrix increases and its inverse calculation is computationally time consuming. Therefore, in the improved optimization methods, it is tried to approximate the inverse matrix of the objective function with some tricks.

Because the ECMWF model has six different lead times, 72 different models can be proposed for 12 different months of the year. For this purpose, data for the period 1993 to 2010 were used as network training data and data for the period 2011 to 2016 for testing. To evaluate the performance of different neural networks, three indices of correlation coefficient, mean square error and Nash-Sutcliffe index were used. Results indicated that the Bayesian Regularization and Levenberg-Marquatt, Conjugate Gradient with Powell/Beale Restarts outperforms other training algorithms.

**Keywords:** Bayesian Regularization, Levenberg-Marquardt algorithm, Multi-Layer perceptron neural network, ECMWF.

---

\* Corresponding author:

pakdaman.m@gmail.com