

پیش‌بینی بارش و دمای متوسط ماهانه با استفاده از الگوهای پیوند از دور به کمک شبکه‌های عصبی مصنوعی (مطالعه موردی: ایستگاه سینوپتیک مشهد)

مریم عرفانیان*، دانشجوی دوره دکتری آبیاری و زهکشی، دانشگاه فردوسی مشهد
دکتر حسین انصاری، دانشیار آبیاری و زهکشی، گروه مهندسی آب، دانشگاه فردوسی مشهد
دکتر امین علیزاده، استاد آبیاری و زهکشی، گروه مهندسی آب، دانشگاه فردوسی مشهد

چکیده

دما و بارش به عنوان دو پارامتر مهم هواشناسی، خصوصاً در مناطق خشک و نیمه خشک مطرح هستند. در نتیجه شناخت لازم از میزان این پارامترها، تغییرات آن‌ها و پیش‌بینی این پدیده‌ها، در جهت داشتن برنامه‌ریزی دقیق تر در مدیریت بخش‌های کشاورزی، اقتصادی و اجتماعی ضروری می‌نماید. از طرف دیگر، تغییرات بسیاری از متغیرهای هواشناسی به شدت به پارامترهای گردش جوی، زمین و سطح اقیانوسی بزرگ مقیاس وابسته‌اند. با توجه به این موضوع، در تحقیق حاضر الگوهای بزرگ مقیاس اقلیمی (پیوند از دور) مؤثر بر پارامترهای دمای متوسط و بارش ماهانه ایستگاه سینوپتیک مشهد در دوره آماری ۵۵ ساله ۲۰۱۰-۱۹۵۶ م. به صورت هم‌زمان و با تأخیرات زمانی مورد بررسی قرار گرفته است. علاوه بر این، جهت پیش‌بینی دو پارامتر فوق، از شبکه‌های عصبی مصنوعی بهره گرفته شده و اثر پیش‌پردازش داده‌های ورودی شبکه عصبی به چهار روش (رگرسیون، رگرسیون با داده‌های محدود، تست گاما و بدون پیش‌پردازش) مورد بررسی قرار گرفته است. نتایج حاصل حاکی از آن است که پیش‌پردازش به روش‌های رگرسیون، رگرسیون محدود و آزمون گاما به ترتیب بهترین نتایج را برای هر دو پارامتر ارائه کرده و پیش‌بینی با شبکه عصبی بدون پیش‌پردازش داده‌های ورودی به ویژه برای پارامتر بارندگی، نتایج مناسبی نداشته است. هم‌چنین مشخص شد که در هر دو روش رگرسیون مورد استفاده، شاخص اقلیمی نینو ۱/۲ با تأخیرات زمانی به ترتیب ۵ و ۴ ماهه مؤثرترین شاخص بر پارامترهای بارش و دمای متوسط بوده است که این نتیجه در آزمون گاما نیز صدق می‌کند.

واژگان کلیدی

شاخص‌های اقلیمی، رگرسیون، تست گاما، شبکه عصبی مصنوعی، مشهد.

۱ - مقدمه

درک علل و ماهیت تغییرات اقلیمی از اهم اهداف جمع آوری داده های هوا و اقلیم و نیز پایش پدیده های اقلیمی است (خورشید دوست و همکاران، ۱۳۸۹: ۹۶). در این راستا، دو پارامتر بارش و دما همواره مورد توجه محققان بوده است. چرا که، این دو پارامتر، مهم ترین پارامترهای اقلیمی محسوب شده که در پهنه‌بندی و طبقه‌بندی‌های اقلیمی نقش عمده‌ای داشته و از عوامل مؤثر بر فعالیت های انسانی، طبیعی، اقتصادی و کشاورزی به شمار می‌رود. علاوه بر این، تغییرات قابل ملاحظه دمای کره زمین یا گرمایش جهانی نیز، یکی از مهم ترین جنبه‌های تغییر اقلیم بوده که در قرن اخیر رخ داده است. هم چنین پدیده های سیل، خشکسالی، طوفان و افزایش و تکثیر حشرات موذی نیز از آثار تغییر اقلیم و خصوصاً افزایش دما بوده که بر جنبه‌های مختلف زندگی انسان تأثیرگذار است. با توجه به این موضوع، پیش بینی بلند مدت این پدیده ها در حوزه های مختلف جهت به حداقل رساندن ریسک، عدم قطعیت و خسارات مالی همواره مورد توجه بوده است (للوید هوگز و ساندرز^۱، ۲۰۰۲: ۱).

اخیراً در این راستا، توجه رفتار اقلیم بر اساس سازوکارهای الگوهای ارتباط از دور مورد توجه محققان قرار گرفته است. والاس و گزلر^۲ (۱۹۸۱: ۷۸۵) پیوند از دور را ارتباط هم زمان بین نوسانات عناصر اقلیمی یک مکان با تغییرات الگوهای فشار و درجه حرارت سطح دریا در نقاط جغرافیایی دیگر تعریف کرده اند. در تعریفی دیگر بیان شده است که الگوهای پیوند از دور، معیارها یا سنجه هایی هستند که به وسیله آن ها تغییرات زمانی شدت و تغییرات مکانی الگوهای گردش جوی- اقیانوسی زمین (یا بخش هایی از آن) اندازه گیری می شود (علیجانی، ۱۳۹۰: ۱۵۸). به عبارت دیگر، با تعریف این نوع شاخص ها، وضعیت، کیفیت و در نهایت توان یک الگوی گسترده سینوپتیک در یک معیار عددی خلاصه می‌شود.

سیگنال های هواشناسی که در اثر تغییرات میدان های فشار در مناطق مختلف زمین شناخته شده اند، می‌توانند با تغییر توزیع فشار در سطح زمین، الگوهای بارش و به ویژه دمای سطح زمین را تعیین کنند (معمدی و همکاران، ۱۳۸۶: ۷۶). محققین علم هواشناسی توانسته‌اند با استفاده از تغییرات میدان های فشار در مناطق مختلف کره زمین، سیگنال های متعددی را بیابند که هر کدام می‌تواند باعث اثرگذاری بر اقلیم مناطقی از کره زمین شود (از جمله تحقیقات وانگ و انفیلد^۳، ۲۰۰۱؛ تامسون و والاس^۴، ۲۰۰۰ و چوینگ^۵ و همکاران، ۲۰۰۲). در این راستا تحقیقات مختلفی چه در سطح کشوری و یا در سطح جهانی در استفاده از این شاخص ها برای بررسی اقلیم مناطق مختلف صورت گرفته است. از جمله جدیدترین این مطالعات می‌توان به خورشیددوست و قویدل رحیمی (۱۳۸۵)، ناظم السادات و همکاران (۱۳۸۶)، یاراحمدی و عزیز (۱۳۸۶)، صلاحی و همکاران (۱۳۸۶)، فاتحی مرج و همکاران (۱۳۸۵)، موسوی بایگی و همکاران (۱۳۸۷)، خورشید دوست و همکاران (۱۳۸۹)، فلاح قالهری و خوشحال (۱۳۸۸)، برهانی داریان و فاتحی مرج (۱۳۸۷)، ناظم السادات و قاسمی (۲۰۰۴)، و بنایان و همکاران (۲۰۱۰) اشاره کرد.

از سوئی دیگر محققین با ابداع و پیشرفت علمی چون روش های هوشمند که ابزاری توانمند و انعطاف پذیر و مستقل از مدل های دینامیکی سیستم هستند، در جست و جوی راه هایی برای پیشرفت در شناخت و

¹ - Lloyd-Hughes and Saunders

² - Wallace and Gutzler

³ - Wang and Enfield

⁴ - Thompson and Wallace

⁵ - Schwing

پیش‌بینی پارامترهای مهم هواشناسی می‌باشند (خلیلی و همکاران، ۱۳۸۷: ۹۰). یکی از مؤلفه‌های هوش مصنوعی، شبکه‌های عصبی هستند که با توجه به توانایی آن‌ها در محاسبه و تقریب هر تابع حسابی و منطقی، مورد توجه محققان قرار گرفته است. لذا، در سال‌های اخیر، محققان (از جمله خلیلی و همکاران (۱۳۸۷)، اصغری مقدم و همکاران (۱۳۸۷)، اسفندیاری درآباد و همکاران (۱۳۸۹) در ایران) با استفاده از تکنیک‌های جدید، به بررسی و پیش‌بینی پدیده‌های هواشناسی به تنهایی و یا با ترکیبی از شاخص‌های اقلیمی پرداخته‌اند.

با مطالعه تحقیقات صورت گرفته بر روی سیگنال‌های هواشناسی مشخص می‌شود که در اکثر آن‌ها، توجه محققان عموماً به بررسی اثر شاخص‌های انسو بر پدیده بارش (فصلی یا سالانه) معطوف بوده و بررسی جامعی از سیگنال‌های مختلف اقلیمی مؤثر بر عناصر مهم هواشناسی (دما و بارش) صورت نگرفته است. از طرف دیگر بررسی سازوکار این شاخص‌ها، مستلزم استفاده از روش‌های کارآمد نوین در این حوزه می‌باشد. لذا، این تحقیق، با هدف بررسی جامعی از شاخص‌های مختلف اقلیمی مؤثر بر پارامترهای دمای متوسط و بارش ایستگاه سینوپتیک مشهد صورت گرفته است. علاوه بر این با توجه به گستردگی و پیچیدگی مطلب، در بررسی فوق از شبکه‌های عصبی مصنوعی استفاده شده است. هم‌چنین با توجه به حجم زیاد داده‌های ورودی (داده‌های ماهانه به همراه لگ‌های زمانی تا ۱۲ ماه قبل) روش‌های مختلف پیش‌پردازش پارامترهای ورودی به شبکه مورد بررسی قرار گرفته و کارایی شبکه‌های عصبی در حالات مختلف مقایسه شده است.

۲- مواد و روش‌ها

۲-۱- مواد

۲-۱-۱- پارامترهای هواشناسی

جهت انجام این تحقیق از داده‌های روزانه بارش، دمای حداکثر و دمای حداقل (برای یافتن دمای متوسط) ایستگاه سینوپتیک مشهد با طول ۵۹ درجه و ۳۸ دقیقه شرقی و عرض جغرافیایی ۳۶ درجه و ۱۶ دقیقه شمالی و ارتفاع ۹۹۹ متر از سطح دریا استفاده شده است. داده‌های روزانه جهت تولید داده‌های ماهانه با حفظ دقت محاسبات در دوره آماری ۵۵ ساله (۲۰۱۰-۱۹۵۶) م. مورد استفاده قرار گرفته است. کلیه آمار فوق مورد بررسی کیفی قرار گرفته و موارد مفقود به روش تفاضل‌ها و نسبت‌ها بازسازی شده است (علیزاده، ۱۳۸۰: ۶۲۴). بر اساس مطالعه روزکوف^۱ و همکاران (۲۰۰۷: ۱۳۳۷) این ایستگاه در ناحیه خشک استان خراسان رضوی با متوسط بارش ۲۴۰ میلی‌متر در سال قرار گرفته است. لذا، با توجه به این مطلب و نیز واقع شدن این ایستگاه در دشت مشهد-چناران، که یکی از مناطق مهم کشاورزی استان به شمار می‌رود، لزوم بررسی و پیش‌بینی پارامترهای اقلیمی این ایستگاه مشخص می‌شود. در جدول ۱ خلاصه‌ای از شاخص‌های آماری این پارامترها در دوره آماری فوق آورده شده است.

جدول ۱: خلاصه‌ای از مشخصات آماری پارامترهای دمای متوسط و بارش ایستگاه سینوپتیک مشهد

پارامتر	حداقل	حداکثر	متوسط	انحراف استاندارد
دمای متوسط (سانتی‌گراد)	-۸/۵	۲۸/۹	۱۴/۳	۸/۸
بارندگی (میلی‌متر)	۰/۰	۱۳۲/۰	۲۰/۸	۲۵/۱

^۱ - Rozhkov

۲-۱-۲- شاخص‌های اقلیمی

مقادیر متوسط ماهانه شاخص‌های اقلیمی جهت بررسی بر روی پارامترهای دما و بارش، از سایت نو^۱ از شبکه‌هایی با قدرت تفکیک $2/5 \times 2/5$ درجه برای سال‌های آماری مورد استفاده در تحقیق اخذ شد. لیست این شاخص‌ها در جدول ۲ آورده شده است. لازم به ذکر است که در بررسی این شاخص‌ها از لگ (تأخیر زمانی) ۲ تا ۱۲ ماه قبل برای بررسی‌های ماهانه استفاده شده است.

۲-۲- روش‌ها

انتخاب و ارائه ورودی‌های مدل شبکه عصبی مصنوعی عمدتاً بر اساس داشتن دانش قبلی از مسأله مورد بررسی صورت می‌گیرد (والورد رامیرز^۲ و همکاران، ۲۰۰۵: ۱۴۸). با توجه به استفاده از شاخص‌های مختلف اقلیمی به همراه تأخیرات زمانی و نیز تعدد ورودی‌ها و نداشتن زمینه‌ای قبلی در انتخاب بهترین پارامترهای ورودی مؤثر بر بارش و دما، از چهار روش پیش‌پردازش متغیرهای ورودی در این تحقیق استفاده شده است. در ادامه توضیح مختصری از این روش‌ها ارائه شده است.

الف- مدل رگرسیون خطی چند متغیره بدون محدودیت داده‌های ورودی

هنگامی که تعداد داده‌های ورودی زیاد باشد، ارزیابی کلیه ترکیبات ممکن نیاز به حجم محاسبات زیادی دارد. بنابراین، روش‌های مختلفی ایجاد شده که صرفاً تعداد کمی از مدل‌های رگرسیون دارای زیرمجموعه‌ای از متغیرها را در یک زمان بررسی می‌کند. این روش‌ها عموماً به روش‌های گام به گام^۳ معروف هستند (قبائی سوق و همکاران، ۱۳۸۹: ۶۱۴). با توجه به تعدد ورودی‌ها در این تحقیق (۳۳۰ عدد)، از روش گام به گام برای انتخاب پارامترهای مؤثر بر بارش و دمای ماهانه استفاده شده است.

جدول ۲: لیست شاخص‌های پیوند از دور مورد استفاده در این تحقیق

نام خلاصه	نام کامل	معادل فارسی
PNA	Pacific North American Index	شاخص اقیانوس آرام آمریکای شمالی
WP	Western Pacific Index	شاخص اقیانوس آرام غربی
NAO	North Atlantic Oscillation	نوسانات اقیانوس اطلس شمالی
SOI	Southern Oscillation Index	شاخص نوسانات جنوبی
ONI	Oceanic Nino Index	شاخص نینوی اقیانوسی
TNA	Tropical Northern Atlantic Index	شاخص حاره‌ای اقیانوس اطلس شمالی
TSA	Tropical Southern Atlantic Index	شاخص حاره‌ای اقیانوس اطلس جنوبی
WHWP	Western Hemisphere Warm Pool	استخر گرم نیمکره غربی
۳NINO	Eastern Tropical Pacific SST	دمای سطح آب در ناحیه نینو ۳
NINO ۱/۲	Extreme Eastern Tropical Pacific SST	دمای سطح آب در ناحیه نینو ۱/۲
۴NINO	Central Tropical Pacific	دمای سطح آب در ناحیه نینو ۴
NINO ۳/۴	East Central Tropical Pacific SST	دمای سطح آب در ناحیه نینو ۳/۴

^۱ - <http://www.esrl.noaa.gov> (2011)

^۲ - Valverde Ramí' rez

^۳ - stepwise regression

نام خلاصه	نام کامل	معادل فارسی
PDO	Pacific Decadal Oscillation	نوسان دهه‌ای اقیانوس آرام
NOI	Northern Oscillation Index	شاخص نوسانات شمالی
NP	North Pacific pattern	الگوی آرام شمالی
TNI	Trans-Nino Index	شاخص انتقالی نینو
AO	Antarctic Oscillation	نوسان قطب جنوب
AMO	Atlantic multidecadal Oscillation	نوسانات چنددهه‌ای اقیانوس اطلس
AMM	Atlantic Meridional Mode	حالت نصف النهاری اقیانوس اطلس
NTA	North Tropical Atlantic SST Index	شاخص دمای سطح آب حاره‌ای اطلس شمالی
CAR	Caribbean SST Index	شاخص دمای سطح آب دریای کاریب
QBO	Quasi-Biennial Oscillation	نوسانات شبه دوسالانه
Monsoon	SW Monsoon Region rainfall	بارش ناحیه جنوب غربی مانسون
Flux	Solar Flux	شار خورشیدی
EA	East Atlantic Pattern	الگوی اطلس شرقی
EA/WR	East Atlantic/West Russia Pattern	الگوی اطلس شرقی/ غرب روسیه
SCA	Scandinavia Pattern	الگوی اسکاندیناوی
POL	Polar/ Eurasia Pattern	الگوی قطبی/ اروپا و آسیا (اوراسیا)
PMM	Pacific Meridional Mode	حالت نصف النهاری اقیانوس آرام
MEI	Multivariable Enso Index	شاخص چند متغیره انسو

به طور خلاصه در توضیح این روش می‌توان گفت که در هر گام، متغیر مستقلی (در ابتدا فقط مقدار ثابت یا عرض از مبدأ) به مدل اضافه می‌شود. به طوری که برای هر مدل رگرسیون، مقدار آماره‌ی F محاسبه شده و هر متغیر مستقلی که دارای F بزرگ تری باشد، برای افزوده شدن به مدل کاندید می‌شود. چنانچه F مربوط به این متغیر مستقل کاندید شده از $F_{\alpha}(1, n-2)$ بزرگ تر باشد، آن گاه متغیر مستقل مربوط را به مدل می‌افزاییم، در غیر این صورت افزودن این متغیر به مدل چندان مفید نخواهد بود. سپس چندین مرحله افزودن و یا حذف متغیرهای مستقل بر اساس محاسبات بر مبنای F انجام شده، که شرح کامل آن در بازگان لاری (۱۳۸۵: ۲۹۲) آورده شده است. در نهایت مجموعه متغیرهای با شرایط فوق به عنوان ورودی مدل سازی انتخاب می‌شوند.

ب- مدل رگرسیون خطی چند متغیره با اعمال محدودیت داده های ورودی

با توجه به مطالعات محققان مشخص شده است که رگرسیون یکی از قوی ترین ابزارها در تبیین روابط بین متغیرها به شمار می‌رود. برخی از روش ها و مدل ها مانند شبکه عصبی، بیشتر برای پیش‌بینی استفاده شده و کم تر در تبیین و آزمون فرضیه ها مورد استفاده قرار می‌گیرند، اما، رگرسیون به خوبی هر دو وظیفه پیش‌بینی و آزمون فرضیه را انجام می‌دهد (یاراحمدی و عزیز، ۱۳۸۶: ۱۶۵). لذا، تهیه مدل های پیش‌بینی

(فصلی یا ماهانه) بر اساس رگرسیون خطی چند متغیره روشی ساده و متداول می‌باشد (کیم^۱ و همکاران، ۲۰۰۷: ۹۲۶). لکن، این پروسه به سادگی اضافه کردن متغیرهای پیش‌بینی کننده به مدل برای رسیدن به یک رابطه مناسب نیست. چرا که، ممکن است باعث ایجاد مدل رگرسیونی فرابرازش داده شده،^۲ شود.

با توجه به این مطلب در تحقیق حاضر، به دلیل وجود متغیرهای ورودی زیاد و به منظور جلوگیری از فرابرازش یافتن مدل، تعداد متغیرهای پیش‌بینی کننده ورودی به شبکه عصبی به حداکثر ۸ عدد (تقریباً ۲۵٪ تعداد پارامترهای پیش‌بینی کننده اصلی) محدود شده است (کیم و همکاران، ۲۰۰۷: ۹۲۶). هم چنین قائل شدن دو دوره آموزش و تست برای یافتن مهارت آموزش مدل انتخابی تاکید شده است (للوید هوگز و ساندرز، ۲۰۰۲: ۳). لذا، در این تحقیق، مدل های پیش‌بینی کننده به طور جداگانه بر اساس تقسیم داده‌ها به دو گروه تهیه شده است. در گروه اول، دوره آموزش ۴۰ سال اول دوره آماری (۷۵٪ کل داده‌ها) و دوره تست ۱۵ سال آخر (۲۵٪ کل داده‌ها) و گروه دوم؛ دوره آموزش ۴۰ سال آخر و دوره تست ۱۵ سال اول دوره آماری انتخاب شده است.

جهت تهیه مدل برای هر ماه، در گام اول کلیه متغیرهای مستقل بر روی داده‌های آموزش گروه اول، به روش رگرسیون گام به گام برآزش داده شده و داده‌های مؤثر مشخص شدند. به ترتیب مجموعه‌های هشت تایی از این پارامترها بر روی داده‌های تست گروه اول برآزش یافته و معنی داری آن‌ها و نیز مناسبت مدل ایجاد شده بررسی شده و در نهایت بهترین مدل انتخاب گردید. سپس در مرحله بعد، داده‌های معنی دار بررسی فوق، بر روی داده‌های آموزشی گروه دوم برآزش داده شده و از لحاظ معنی داری بررسی شدند. سپس، همان پارامترها (و یا دیگر پارامترها در صورت عدم معنی داری پارامترهای فوق) بر روی داده‌های تست گروه دوم مورد بررسی قرار گرفتند. لازم به ذکر است که این مرحله به جهت کنترل پیش‌بینی کنندگی مدل تعیین شده در گروه اول برای گروه دوم صورت گرفته است. بدیهی است، چنان چه مدل انتخابی برای گروه دوم مناسب نباشد، باید فرآیند مرحله یک را برای گروه دوم با پارامترها و ترکیبات مختلف آن‌ها آن قدر تکرار کرد که مدل مناسبی برآزش یابد.

ج- تست گاما

تست گاما یک ابزار مدل‌سازی غیرخطی است؛ که به کمک آن می‌توان ترکیب مناسب از پارامترهای ورودی برای مدل‌سازی داده‌های خروجی و ایجاد یک مدل هموار را بررسی نمود. هم چنین یک ابزار توسعه یافته برای تخمین میانگین مربعات خطای حاصل از مدل‌سازی پدیده‌های مختلف با استفاده از مجموعه داده‌های مشاهداتی از آن پدیده است. این آزمون یک روش کارآمد بوده و می‌تواند در مباحث منابع آب، شامل پیش‌بینی جریان رودخانه‌ها و سایر موارد مدل‌سازی غیرخطی هیدرولوژی به کار رود (نوری و همکاران، ۲۰۱۱: ۱۷۷).

چنان چه فرض شود که مجموعه‌ای از داده‌های ورودی؛ $x_i \in R^m$ (که m نشان دهنده بعد آن) و داده‌های خروجی؛ $y_i \in R$ مشاهده شده از پدیده‌ای به صورت $\{(x_i, y_i), 1 \leq i \leq M\}$ ، که در آن M تعداد سری مجموعه مشاهده شده از آن پدیده باشد، آن گاه تست گاما با فرض عدم قطعیت و پذیرش مقدار خطا در به

¹ - Kim

² - over fit regression model

دست آوردن خروجی از روی ورودی، به دلیل پیچیدگی و غیرخطی بودن پدیده‌های مدل سازی این خطا را به صورت رابطه ۱ بین مجموعه داده‌های ورودی و خروجی با تابع f نشان می‌دهد.

$$(۱) y = f(x_1, \dots, x_m) + r$$

که در رابطه فوق f نشان دهنده تابع هموار استفاده شده، برای مدل سازی و Γ معرف متغیر تصادفی است که برای نشان دادن خطا به کار می‌رود. چنان چه فرض شود که متوسط توزیع Γ برابر صفر بوده و واریانس خطا کراندار باشد، مدل فوق به مدلی با مشتقات جزئی درجه اول محدود می‌شود. در این حالت تست گاما نشان دهنده آن بخشی از واریانس داده‌های خروجی است؛ که نمی‌تواند توسط مدل هموار محاسبه گردد.

آزمون گاما بر اساس $N[i,k]$ می‌باشد؛ که k امین $(1 \leq k \leq p)$ نزدیک ترین همسایه برای هر بردار x_i که $(1 \leq i \leq M)$ است می‌باشد. تست گاما از رابطه ۲ به دست می‌آید.

$$(۲) \delta_M(k) = \frac{1}{M} \sum_{i=1}^M |x_{N[i,k]} - x_i|^2 \quad (1 \leq k \leq p)$$

که در آن $|\dots|$ فاصله اقلیدسی بوده و مقدار متناظر آن برای خروجی عبارت است از:

$$(۳) \gamma_M(k) = \frac{1}{2M} \sum_{i=1}^M |y_{N[i,k]} - y_i|^2 \quad (1 \leq k \leq p)$$

که در رابطه فوق $y_{N[i,k]}$ مقدار متناظر خروجی برای نزدیک ترین همسایه x_i در معادله ۳ است. با ایجاد رابطه رگرسیون خطی برای p بین $(\delta_M(k), \gamma_M(k))$ مقدار آماره گاما برابر عرض از مبدأ معادله رگرسیونی به صورت زیر خواهد بود:

$$(۴) \gamma = A\delta + \Gamma$$

نتایج آزمون گاما می‌تواند توسط عبارت V_{ratio} ، که یک معیار بدون بعدی با مقادیر بین صفر و یک است، استاندارد شود. V_{ratio} به صورت زیر تعریف می‌شود.

$$(۵) V_{ratio} = \frac{\Gamma}{\sigma^2(y)}$$

در رابطه فوق $\sigma^2(y)$ واریانس خروجی y می‌باشد. هرچه این معیار به صفر نزدیک تر باشد، نشان دهنده توانایی پیش‌بینی شدن بهتر خروجی‌های y است. هم چنین یکی دیگر از معیارهای سنجش مدل، شیب خط رگرسیون است که هر چه این شیب تندتر باشد، نشان دهنده پیچیدگی بیشتر مدلی است که از روی ورودی و خروجی ساخته می‌شود.

اعتبار آماره Γ می‌تواند توسط اجرای چندین آزمون گاما با افزایش دادن M ، جهت یافتن تعداد داده مورد نیاز برای رسیدن به مقدار ثابت انجام شود. با انجام آزمون M -test، به طوری که متوسط انحراف استاندارد داده‌های آموزشی از واریانس Γ کوچک تر باشد، می‌توان از فرابرازش یافتن مدل جلوگیری کرد.

برای انجام آزمون گاما در این تحقیق از نرم افزار winGamma 1.98 استفاده شده است (دورانت ۱، ۲۰۰۱). در نهایت بهترین ترکیب ارائه شده توسط نرم افزار به عنوان داده‌های ورودی شبکه عصبی مورد استفاده قرار می‌گیرد.

د- تحلیل شبکه عصبی بدون پیش پردازش داده‌های ورودی

به عنوان چهارمین حالت در تحلیل شبکه عصبی، کلیه شاخص‌های اقلیمی به همراه تأخیرات زمانی بدون انجام پیش‌پردازش به شبکه وارد شده و مورد بررسی قرار می‌گیرند.

۳-۲- شبکه‌های عصبی مصنوعی

نظریه شبکه‌های عصبی مصنوعی برای بار نخست در سال ۱۹۴۳ م. توسط مک کلاچ و پیتز^۱ ارائه شد که در حقیقت از ساختار عصبی مغز انسان الهام گرفته بودند. شبکه‌های عصبی از یک سری لایه‌هایی شامل اجزای ساده‌ای به نام نرون تشکیل شده است؛ که به صورت موازی با هم عمل می‌کنند. اکثر شبکه‌های عصبی دارای ۳ لایه و یا بیشتر می‌باشند. لایه ورودی، که برای ورود داده‌ها به شبکه استفاده می‌شود، لایه خروجی، برای تولید پاسخ‌های مناسب به ورودی‌های داده شده و یک یا تعداد بیشتر لایه حد واسط که از گره‌های پردازشگر تشکیل شده‌اند و محل پردازش داده‌ها به شمار می‌روند. تعداد نرون‌های موجود در لایه‌های ورودی و خروجی با توجه به ماهیت مسأله مورد بررسی مشخص می‌شود، حال آن‌که شمار لایه‌های مخفی و تعداد نرون در هر لایه مخفی معمولاً به وسیله روش آزمون و خطا در جهت کاهش مقدار خطا توسط طراح مشخص می‌شود (مقدم نیا و همکاران، ۲۰۰۹: ۸۹). با این حال توصیه شده است که تعداد لایه‌های پنهان تا حد امکان باید کم باشد (اسفندیاری درآباد و همکاران، ۱۳۸۹: ۵۵). به طوری که ابتدا شبکه با یک لایه پنهان آموزش داده شده و در صورت عملکرد نامناسب، تعداد لایه‌ها افزوده می‌شود. این موضوع در مورد تعداد نرون‌های لایه‌های پنهان نیز صدق می‌کند، به گونه‌ای که در ابتدا تعداد نرون‌های کم مورد بررسی قرار گرفته و در صورت همگرا نشدن به جواب مطلوب، آن‌ها را افزایش می‌دهند.

گره‌های لایه‌های مجاور در شبکه به طور کامل با هم در ارتباط هستند (اصغری مقدم و همکاران، ۱۳۸۷: ۴). ورودی‌های هر گره ممکن است از متغیرهای ورودی یا از خروجی گره‌های دیگر حاصل شود. هر گره دارای یک تابع تبدیل گر (محرک) است که سه تابع معروف فوق عبارت است از: تابع تانژانت سیگموئید، خطی و لگاریتم سیگموئید (دموت و بیله^۲، ۲۰۰۰: ۱۰). ورودی‌ها به شکل یک بردار $x(x_1, x_2, \dots, x_n)$ هستند و هر ورودی توسط یک وزن به گره پردازشگر مربوط می‌شود و در نهایت تسلسلی از وزن‌ها به شکل بردار وزن $w(w_1, w_2, \dots, w_n)$ به گره مورد نظر مرتبط می‌گردد. خروجی گره که y نامیده می‌شود طبق رابطه زیر محاسبه می‌شود:

$$(۶) \quad y = f(x.w - b)$$

که در رابطه فوق x داده‌های ورودی، w بردار وزن و b مقدار آستانه یا بایاس^۳ می‌باشد. به طور کلی شبکه‌های عصبی به دو نوع پیشرو^۴ و پسرو^۵ تقسیم می‌شوند. تفاوت آن‌ها در این است که در شبکه‌های پسرو، حداقل یک سیگنال برگشتی از یک نرون به همان نرون یا نرون‌های همان لایه و یا لایه قبل وجود دارد. در بیشتر موارد، شبکه‌های عصبی پسرو می‌توانند بسیار مفید واقع شوند، ولی با این حال در ۸۰ درصد کاربردها از شبکه‌های عصبی پیشرو استفاده می‌شود (اسفندیاری درآباد و همکاران، ۱۳۸۹: ۵۲).

^۱ - McCulloch and Pitts

^۲ - Demuth and Beale

^۳ - Bias

^۴ - Feed Forward Neural Network (FFN)

^۵ - Recurrent/ Feedback Neural Network (RNN)

با گذشت بیش از ۶۰ سال از مطرح شدن ایده اولیه شبکه‌های عصبی، نخستین کاربرد عملی آن با معرفی شبکه‌های پرسپترون چند لایه^۱ انجام شد. در شبکه‌های فوق، هر نرون در هر لایه به تمام نرون‌های لایه قبل متصل است. این شبکه‌ها از نوع پیشرو بوده که یکی از پرکاربردترین مدل‌های شبکه عصبی مصنوعی در مدل‌سازی و پیش‌بینی عناصر اقلیمی است (حسینی، ۱۳۸۸: ۲۸).

۱-۳-۲- نرمال کردن داده‌ها

وارد کردن داده‌ها به صورت خام باعث کاهش سرعت و دقت شبکه می‌شود. لذا، با استفاده از رابطه زیر کلیه داده‌های ورودی و خروجی در ابتدا نرمال شده و سپس وارد شبکه عصبی شدند (قربانی و همکاران، ۱۳۸۸: ۲۴۶).

$$(Y) x_n = \frac{x_i - x_{\min}}{x_{\max} - x_{\min}}$$

در رابطه فوق، x_n = داده نرمال شده و اندیس‌های i ، \min و \max به ترتیب مربوط به ردیف، حداقل و حداکثر آن داده در مجموعه خود است.

۲-۳-۲- روش‌های آموزش

جهت آموزش شبکه‌های عصبی چهار الگوریتم آموزشی مرسوم که بر اساس ساختار پرسپترون چند لایه می‌باشد، وجود دارد. معروف‌ترین این روش‌ها عبارتند از: انتشار خطا به عقب^۲، گرادیان مزدوج^۳ و لونیبرگ-مارکوات^۴. الگوریتم لونیبرگ-مارکوات از سال ۱۹۹۳ م. تا امروز به عنوان سریع‌ترین روش آموزش شبکه‌های عصبی شناخته شده است. در این بررسی نیز از این الگوریتم استفاده شده است.

۴-۲- ارزیابی عملکرد شبکه

جهت ارزیابی عملکرد شبکه و انتخاب بهترین حالت، از چهار معیار متوسط خطا، متوسط قدر مطلق خطا، ریشه میانگین مربعات خطا و ضریب همبستگی که به ترتیب در روابط ۲ تا ۵ آورده شده، استفاده شده است:

$$(۸) BIAS = \frac{1}{M} \sum_{i=1}^M (x_{i,m} - x_{i,o})$$

$$(۹) MAE = \frac{1}{M} \sum_{i=1}^M |x_{i,m} - x_{i,o}|$$

$$(۱۰) RMSE = \sqrt{\frac{1}{M} \sum_{i=1}^M (x_{i,m} - x_{i,o})^2}$$

$$(۱۱) \rho = \frac{M \sum_{i=1}^M x_{i,m} \cdot x_{i,o} - \sum_{i=1}^M x_{i,m} \cdot \sum_{i=1}^M x_{i,o}}{\sqrt{[M \sum_{i=1}^M x_{i,m}^2 - (\sum_{i=1}^M x_{i,m})^2] \cdot [M \sum_{i=1}^M x_{i,o}^2 - (\sum_{i=1}^M x_{i,o})^2]}}$$

که در روابط فوق M = تعداد کل داده‌های ورودی، $x_{i,m}$ = معرف i امین داده پیش‌بینی شده با استفاده از یکی از چهار مدل فوق و $x_{i,o}$ = معرف i امین داده مشاهده شده است.

^۱ - Multi Layer Perceptron (MLP)

^۲ - Back Propagation

^۳ - conjugate gradient

^۴ - Levenberg-Marquardt

۳- بحث و نتایج

همان‌طور که در قسمت‌های قبل ذکر شد چهار روش برای پیش‌پردازش داده‌های دمای ورودی به شبکه عصبی و پیش‌بینی عناصر اقلیمی بارش و دمای متوسط استفاده شد. این روش‌ها عبارت بودند از: رگرسیون بدون محدودیت داده‌ها، رگرسیون با اعمال محدودیت داده‌های ورودی، تست گاما و استفاده از کلیه پارامترها بدون پیش‌پردازش اولیه. در ادامه به نتایج حاصل از هر روش به‌طور خلاصه اشاره می‌شود.

۳-۱- مدل رگرسیون خطی بدون محدودیت داده‌های ورودی

جهت به دست آوردن مدل پیش‌بینی‌کننده دمای متوسط و بارش با استفاده از پارامترهای بزرگ مقیاس اقلیمی، در ابتدا کلیه این پارامترها به همراه تأخیرات زمانی وارد مدل رگرسیونی شدند. با استفاده از روش گام به گام داده‌های ورودی مؤثر برای هر پارامتر مشخص شده تا جایی که ورودی مؤثر دیگری شناخته نشود. جهت انجام بررسی از کلیه داده‌های دوره آماری استفاده شده و دوره‌های آموزش و تست قائل نشده است. لازم به ذکر است که برای بررسی پارامتر بارش، با توجه به این که در فصل تابستان منطقه مورد مطالعه بارش‌ها کم و ناچیز است، لذا، مقادیر مربوط به ماه‌های تیر تا شهریور (که عموماً صفر و یا نزدیک به صفر بوده است) از مجموعه داده‌های ورودی حذف شده است. هم‌چنین برای هر پارامتر، نردبان پله‌ای ضریب همبستگی و انحراف معیار برای توابع مختلف آن پارامتر (شامل لگاریتم نمایی، توان‌های مختلف، ...) به دست آمده و بر اساس آماره‌های مختلف از جمله دوربین-واتسون^۱، باقی‌مانده‌های استاندارد، منحنی پخش باقی‌مانده‌های استاندارد و ... در ابتدا تصمیم‌گیری شد که چه تابعی برای آن پارامتر به کار گرفته شود. در بررسی حاضر، تابع لگاریتم نمایی برای بارش منجر به نتایج بهتری شده است. نهایتاً معادله رگرسیونی حاصل از کلیه پارامترهای مؤثر (با سطح اطمینان ۰.۹۵) بر بارش و دما به صورت زیر استخراج شده است:

$$\begin{aligned} \ln(\text{Rain}) = & -19.96 - 0.50 * \text{NINO1.2}(5) - 0.29 * \text{MANSOON}(11) - 0.28 * \text{PNA}(2) - 0.25 * \text{PNA}(3) + \\ & 0.51 * \text{MEI}(3) + 0.06 * \text{NOI}(10) - 0.66 * \text{ONI}(7) + 0.34 * \text{TNI}(5) - 0.11 * \text{NP}(3) + 1.28 * \text{NINO4}(6) - \\ & 0.09 * \text{NOI}(5) - 0.09 * \text{NP}(10) + 0.11 * \text{WP}(11) - 0.20 * \text{PNA}(10) - 0.23 * \text{MANSOON}(12) - \\ & 0.12 * \text{SCA}(2) - 0.04 * \text{NP}(2) - 0.12 * \text{SCA}(5) + 0.51 * \text{NINO3.4}(10) - 0.48 * \text{NINO4}(9) + 0.15 * \text{SOI}(7) + \\ & 0.11 * \text{POL}(12) \end{aligned}$$

$$(R^2=0.75, SE=1.22, D-W=1.81)$$

$$\begin{aligned} \text{Mean temperature} = & 72.95 + 0.71 * \text{NINO1.2}(4) - 0.65 * \text{NINO1.2}(10) + 1.03 * \text{MEI}(10) + 1.20 * \\ & \text{AMO}(12) - 1.34 * \text{NINO4}(5) - 0.67 * \text{NINO1.2}(11) + 0.08 * \text{NP}(12) - 0.53 * \text{NINO1.2}(9) + 1.06 * \\ & \text{MEI}(8) + 1.18 * \text{NINO4}(11) - 0.24 * \text{MANSOON}(5) + 0.30 * \text{PNA}(5) - 2.33 * \text{TNI}(6) + 1.01 * \\ & \text{NINO3}(2) - 0.73 * \text{MEI}(2) + 0.43 * \text{NINO1.2}(6) + 2.47 * \text{TNI}(9) - 0.02 * \text{QBO}(5) - 0.22 * \text{SCA}(7) - \\ & 1.26 * \text{NINO3}(8) - 0.09 * \text{NOI}(10) + 0.15 * \text{WP}(6) + 0.14 * \text{SCA}(12) + 0.40 * \text{NINO1.2}(5) - 0.99 * \\ & \text{NINO4}(4) - 0.54 * \text{MEI}(12) + 0.14 * \text{NAO}(4) + 2.38 * \text{AMO}(5) - 0.19 * \text{EAWR}(2) - 0.19 * \text{POL}(4) \\ & + 1.26 * \text{ONI}(8) - 0.15 * \text{WP}(3) - 1.62 * \text{CAR}(2) + 0.17 * \text{EA}(10) - 0.55 * \text{NINO3}(10) + 0.16 * \text{WP}(8) + \\ & 0.07 * \text{NOI}(5) \end{aligned}$$

$$(R^2=0.980, SE=1.79, D-W=1.783)$$

¹ - Durbin- Watson

در روابط فوق، R^2 = ضریب تعیین^۱، SE = انحراف از استاندارد^۲ و $D-W$ = آماره دوربین- واتسون می‌باشد. مقادیر داخل پرانتز نیز نشان دهنده تأخیر زمانی شاخص اقلیمی است. لازم به یادآوری است که هرچه مقدار این آماره به عدد ۲ نزدیک تر باشد، می‌توان نتیجه گرفت خطای مدل (اجزای اخلاص) با یک دیگر بستگی پیاپی ندارند (یاراحمدی و عزیزی، ۱۳۸۶: ۱۶۹). در روابط فوق، آماره‌های دوربین- واتسون به دست آمده نشان دهنده عدم وجود مشکل همبستگی پیاپی در مدل است.

با توجه به مدل های فوق، مشخص می‌شود که شاخص اقلیمی NINO 1.2 با تأخیرات زمانی ۵ و ۴، به ترتیب مؤثرترین شاخص بر پارامترهای بارش و دما می‌باشد. همان طور که در قسمت مقدمه و بررسی مطالعات ذکر شد تا کنون بررسی جامع مشابهی در خصوص استفاده از کلیه شاخص های اقلیمی مؤثر بر پارامترهای هواشناسی صورت نگرفته است؛ که بتوان نتایج این تحقیق را با آن مقایسه نمود. با این حال در تحقیق خورشیددوست و همکاران (۱۳۸۹: ۱۲۱)، بر مؤثر بودن شاخص NINO 1.2 بر بارش های بهاره و پائیزه ایستگاه اهر گزارش شده است. هم چنین در مطالعه قویدل رحیمی (۱۳۸۴: ۷۱) نیز این شاخص مؤثر بر بارش های بهاره آذربایجان شرقی بوده است. در نهایت، پارامترهای مؤثر در هر مدل به عنوان ورودی شبکه عصبی در نظر گرفته شده است.

۲-۳- مدل رگرسیون خطی با اعمال محدودیت داده های ورودی

جهت بررسی بهترین مدل بر روی بارش و دمای متوسط، مدل های مختلفی با ترکیب ۸ تایی از پارامترهای پیوند از دور برازش داده شدند که نتایج حاصل از چهار مدل مناسب تر با پارامترهای مؤثر (در سطح خطای نوع اول ۰/۵٪) بر روی گروه اول در جدول ۳ آورده شده است. لازم به ذکر است که ممکن است، پارامترهای گروه ۱ و ۲ متفاوت باشند. چرا که، در بررسی ها هدف تعیین بهترین پارامترهایی است؛ که بیشترین همبستگی را بین مقادیر مشاهدات و پیش بینی ها در دوره تست و نه آموزش هر گروه ارائه کنند.

جدول ۳- بهترین مدل های برازش داده شده بر روی داده های لگاریتم بارندگی (گروه ۱)

شماره	ورودی ها								R^2_{test}	SE
۱	NINO ۱/۲ (۵)	MEI (۳)	TNI(۵)	NINO ۴(۶)	MANSOON (۱۱)	NP(۳)	PNA (۲)	ONI (۷)	۰/۷۶۱	۰/۴۰۹
۲	NINO ۱/۲ (۵)	MEI (۳)	NINO ۴(۶)	TNI (۵)	MANSOON (۱۱)	NINO ۱/۲(۲)	WP (۱۱)	ONI (۷)	۰/۷۵۹	۰/۴۰۶
۳	NINO ۱/۲ (۵)	PNA (۲)	NINO ۱/۲(۲)	NINO ۴(۶)	TNI(۵)	MANSOON (۱۱)	WP (۵)	ONI (۷)	۰/۷۴۵	۰/۴۱۶
۴	NINO ۱/۲ (۲)	PNA (۲)	POL (۱۲)	WP (۵)	MANSOON (۱۱)	NINO ۱/۲(۵)	PMM (۶)	NINO ۴(۶)	۰/۷۲۶	۰/۴۲۹

لازم به ذکر است که R^2_{test} و SE، درحقیقت مقدار ضریب تعیین و انحراف استاندارد به دست آمده از برازش مدل با پارامترهای ارائه شده بر روی داده های دوره تست گروه اول می‌باشد.

¹ - Determination Coefficient

² - Standard Error

با توجه به مقادیر آماره‌های ضریب تعیین، انحراف استاندارد و نیز ضریب دوربین-واتسون مشخص شد که مدل اول، بهترین نتایج را داشته است. لذا، باید با استفاده از ضرایب این پارامترها، کارایی پیش‌بینی این مدل و نیز مدل‌های دیگر در گروه دوم نیز کنترل گردد. نتایج حاصل به شرح جدول ۴ است.

جدول ۴- بهترین مدل‌های برازش داده شده بر روی داده‌های لگاریتم بارندگی (گروه ۲)

شماره	ورودی‌ها								R ² _{test}	SE
۱	NINO ۱/۲(۵)	MEI (۳)	TNI(۵)	NIN O ۴(۶)	MANSOO N (۱۱)	NP (۳)	PNA(۲)	ONI (۷)	۰/۷۵۶	۰/۴۰۶
۲	NINO ۱/۲ (۵)	POL (۱۲)	TNI(۵)	NIN O ۴(۶)	MANSOO N (۳)	NIN O ۱/۲ (۶)	WP(۱۱)	PMM (۶)	۰/۷۵۳	۰/۴۱۰
۳	NINO ۱/۲ (۴)	TNI (۵)	NINO ۱/۲ (۳)	WP (۵)	WHWP (۱۲)	MEI (۳)	MANS OON (۸)	ONI (۷)	۰/۷۳۳	۰/۴۲۴
۴	MANSOO N (۱۱)	NIN O ۱/۲ (۵)	PMM(۴)	NIN O ۴(۵)	TNI(۵)	WP (۵)	ONI(۸)	NIN O ۱/۲ (۳)	۰/۷۴۸	۰/۴۱۸

با توجه به مقادیر جدول‌های بالا، مشخص می‌شود که مدل اول، که در واقع بهترین مدل انتخابی مرحله قبل است، در گروه دوم نیز بهترین نتایج را ارائه کرده است. به عبارت دیگر، این امر به این معنی است که مدل انتخابی برازش داده شده بر روی گروه اول (با داده‌های آموزشی ۴۰ سال اول)، قابلیت پیش‌بینی و کارایی مناسبی برای استفاده در گروه دوم (با داده‌های آموزشی ۴۰ سال آخر) را دارا بوده و به عنوان مدل مناسب برای پیش‌بینی بارش ماهانه با استفاده از شاخص‌های پیوند از دور مؤثر بر آن انتخاب می‌شود. لذا، جهت برآورد بارش ماهانه می‌توان معادله زیر را ارائه کرد:

$$LN(Rain) = -14.23 - 0.64 * NINO1.2(5) + 0.57 * MEI(3) + 0.42 * TNI(5) + 1.12 * NINO4(6) - 0.44 * MANSOON(11) - 0.05 * NP(3) - 0.19 * PNA(2) - 0.52 * ONI(7)$$

$$(R^2 = 0.75, SE = 1.30)$$

۳-۲-۱- دمای متوسط

با توجه به روش ارائه شده فوق، پارامترهای مؤثر بر دمای متوسط ماهانه نیز مورد بررسی قرار گرفتند که به دلیل پرهیز از طولانی شدن مطلب، در زیر به ارائه نتایج اکتفا شده است.

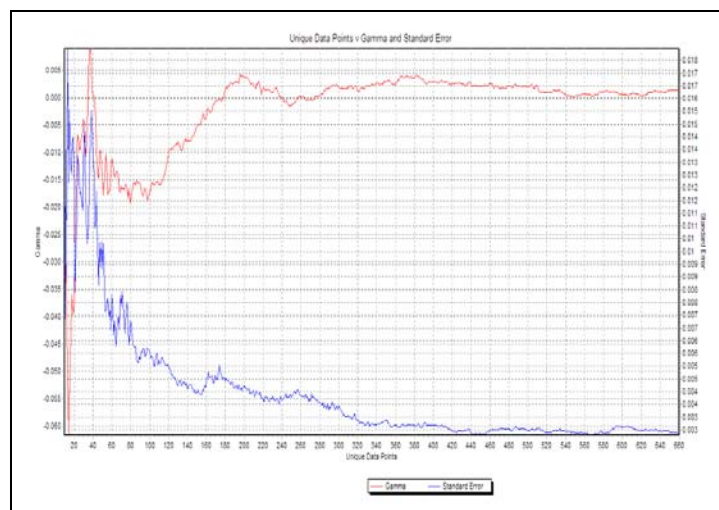
$$T_{mean} = 93.740 + 1.621 * NINO 1.2 (4) - 2.433 * NINO 1.2 (10) + 2.174 * MEI (10) + 1.286 * TNI (10) - 1.471 * TNI (4) + 2.785 * AMO (12) - 2.143 * NINO 4(5) + 0.348 * WHWP (4)$$

$$(R^2 = 0.96, SE = 2.29)$$

با توجه به معادلات ارائه شده فوق، پارامترهای موجود در آن‌ها به عنوان ورودی شبکه عصبی با پیش-پردازش به روش رگرسیون محدود استفاده شده است.

۳-۳- تست گاما

به عنوان سومین روش پیش پردازش ورودی‌های شبکه عصبی و تعیین مؤثرترین شاخص های اقلیمی بر روی پارامترهای بارش و دمای متوسط، تست گاما مورد استفاده قرار گرفته است. پیش از انجام هر گونه آنالیزی، باید بهترین همسایه (P) برای هر پارامتر تعیین گردد. لذا، تحلیل‌های مربوط به افزایش همسایه از ۲ تا ۱۰۰ برای هر پارامتر انجام شده است. نتایج نشان داد که در همسایه ۲۵، مقدار انحراف استاندارد و گاما به حداقل خود رسیده و منحنی انحراف استاندارد تقریباً ثابت شده است. لذا، برای پارامتر دمای متوسط، این عدد به عنوان نزدیک ترین همسایه برای تحلیل‌های بعدی استفاده می‌شود. بر اساس آنالیز انجام شده تعداد مناسب ورودی برای پیش‌بینی متغیر مورد نظر با استفاده از آزمون M (M-test) مورد بررسی قرار گرفت. با استفاده از این آزمون، مشخص می‌شود که چه تعداد داده برای ایجاد مقدار آماره گاما ثابت و متعاقباً تهیه یک مدل مناسب کافی می‌باشد. در شکل ۱ نتایج این آزمون برای دمای متوسط آورده شده است.



شکل ۱: نتایج آزمون M برای آماره‌های گاما و انحراف خطا پارامتر دمای متوسط (نزدیک ترین همسایه = ۲۵)

با توجه به شکل فوق مشخص می‌شود که در نزدیکی ۴۰۰ امین داده (یعنی $M = 400$) مقدار آماره گاما به حدود $0/0025$ ثابت شده است. هم چنین مقدار انحراف معیار متناظر با این حجم داده نیز بسیار کم و در حدود $0/00075$ بوده، که نشان دهنده دقت و صحت نتایج آماره گاما است. لذا، می‌توان از این تعداد داده برای یافتن پارامترهای مؤثر بهره جست.

در این تحقیق، ترکیبات مختلف از ورودی‌ها، که تعداد کل آن‌ها $(2^n - 1)$ می‌باشد، مورد بررسی قرار گرفته است. بهترین حالت از این ترکیبات مربوط به ترکیبی است که دارای کم ترین مقدار گاما (Γ)، کاهش سریع در نمودار انحراف معیار M-test، مقدار کم Vratio، که نشان دهنده مدل دقیق هموار است، شیب کوچک خط رگرسیون، که حاکی از ساده بودن مدل غیرخطی با پیچیدگی کم است و هم چنین مقدار انحراف معیار کم در نظر گرفته می‌شود.

در نهایت با توجه به آماره‌های فوق، آنالیزهای مربوط برای پارامترهای بارش و دمای متوسط انجام شده و بهترین ترکیبات ممکن برای هر یک مشخص شد. مقادیر آماره‌های فوق برای بهترین ترکیبات دمای متوسط و

بارش در جدول ۵ آورده شده است. با توجه به طولانی بودن بهترین ترکیبات از ذکر آن‌ها در این قسمت خودداری شده است.

جدول ۵: نتایج مقدار آماره‌های آزمون گاما برای ترکیبات منتخب دمای متوسط و بارش

پارامتر	مقدار گاما	گرادیان	Vratio	خطای استاندارد
دمای متوسط	۰/۰۲۴۴	۰/۰۰۳۵	۰/۳۹۰	۰/۰۰۳۴
بارش	۰/۰۰۰۰۸	۰/۰۰۴۰	۰/۰۰۰۱	۰/۰۰۲۳

۳-۴- داده‌های ورودی بدون پیش‌پردازش

به عنوان حالت چهارم، هیچ‌گونه پیش‌پردازشی بر روی داده‌های ورودی به شبکه عصبی انجام نشده و کلیه پارامترها به همراه تأخیرات زمانی آن‌ها (۳۳۰ ورودی) وارد شبکه عصبی شدند.

۳-۵- نتایج تحلیل روش‌های پیش‌پردازش داده‌های ورودی بر روی شبکه عصبی

بر اساس روش‌های مختلف پیش‌پردازش داده‌های ورودی ذکر شده در بالا، پارامترهای مؤثر در هر روش انتخاب گردیده و وارد شبکه عصبی گردید. جهت ایجاد ساختار شبکه عصبی، توابع محرک مختلف با الگوریتم لونیبرگ-مارکوات مورد بررسی قرار گرفت. ۷۵٪ از داده‌ها، که به طور تصادفی انتخاب شده‌اند، برای آموزش و ۲۵٪ برای تست مدل استفاده شده است. تعداد مختلف لایه‌های پنهان با نرون‌های مختلف، که از مقدار کم تا زیاد (۲ تا ۲۰) افزایش می‌یافت برای ایجاد بهترین شبکه‌ها مورد بررسی قرار گرفتند. با توجه به در نظر گرفتن حالات مختلف (تعداد لایه، تعداد نرون، توابع انتقال و روش‌های پیش‌پردازش مختلف)، برای هر پارامتر جدولی تهیه شد که در آن مقدار آماره‌های ذکر شده در فوق برای هر حالت محاسبه و تعیین گردید. با توجه به طولانی بودن این جدول و پرهیز از طولانی شدن مطلب، از ذکر ریز محاسبات انجام شده صرف‌نظر گردیده و تنها نتایج بهترین حالات انتخاب شده از بین موارد فوق در ادامه ارائه شده است.

۳-۵-۱- مقایسه آماری

بهترین حالت‌های شبکه عصبی که به روش سعی و خطا برای هر روش پیش‌پردازش انتخاب شده است در جدول ۶ آورده شده است.

جدول ۶- نتایج آنالیز شبکه عصبی برای بارندگی و متوسط دما نرمال شده

پارامتر	مدل (روش پیش‌پردازش)	تابع انتقال	تعداد لایه پنهان	R	BIAS	MAE	RMSE
دمای متوسط	رگرسیون کلی	PURELIN	۳	۰/۹۷۸	۰/۰۰۵۱	۰/۰۳۶۲	۰/۰۴۸۳
	رگرسیون محدود	PURELIN	۳	۰/۹۶۶	-۰/۰۰۱۵	۰/۰۴۶۷	۰/۰۶۲۰
	تست گاما	PURELIN	۷	۰/۹۴۴	۰/۰۰۳۹	۰/۰۶۶۲	۰/۰۸۱۵
	بدون پیش‌پردازش	TANSIG	۴	۰/۹۱۶	۰/۰۰۸۸	۰/۰۷۵۸	۰/۰۹۴۷
بارش	رگرسیون کلی	PURELIN	۵	۰/۸۴۹	-۰/۰۱۲۷	۰/۱۳۷۶	۰/۱۸۲۰
	رگرسیون محدود	PURELIN	۴	۰/۷۸۹	-۰/۰۱۶۹	۰/۱۴۵۳	۰/۱۸۶۲
	تست گاما	LOGSIG	۳	۰/۶۰۶	۰/۰۴۳۴	۰/۱۵۹۵	۰/۲۱۵۶
	بدون پیش‌پردازش	TANSIG	۳	۰/۵۵۷	-۰/۰۲۳۹	۰/۱۷۲۶	۰/۲۴۰۱

در جدول فوق، R مقدار ضریب همبستگی بین مقادیر مشاهده شده و برآورد شده در دوره تست است. مقدار BIAS معرف متوسّط خطای پیش‌بینی برای مقادیر برآورد شده است که مقادیر مثبت (و منفی) آن نشان دهنده فرابراورد (فروبرآورد) نسبت به مقادیر مشاهده شده است. با توجه به این که هر کدام از مقادیر برآورد شده می‌تواند خطای مثبت یا منفی داشته باشد؛ لذا، در میانگین ممکن است این خطاها سرشکن شوند. در نتیجه این آماره به تنهایی نمی‌تواند معرف دقت پیش‌بینی‌ها باشد. از طرف دیگر، MAE که یک مقدار نامنفی است متوسّط انحراف مقادیر پیش‌بینی از مشاهدات را ارائه می‌کند. علاوه بر این، RMSE نیز که مقداری مثبت است، نسبت به MAE به خطاهای بیشتر پیش‌بینی حساس تر می‌باشد. برای یافتن بهترین نتایج پیش‌بینی باید مقدار ضریب همبستگی به ۱ نزدیک بوده و آماره‌های MAE، RMSE و BIAS کوچک باشند.

با توضیح فوق و مقادیر ارائه شده در جدول، برای پارامتر دمای متوسّط مشخص می‌شود که بهترین نتایج مربوط به شبکه عصبی با پیش‌پردازش داده‌های ورودی به روش رگرسیون (بدون محدودیت ورودی) است؛ که منجر به بالاترین مقدار ضریب همبستگی و حداقل مقدار آماره‌های MAE و RMSE شده است. ولیکن، مقدار BIAS آن از مقدار مربوط به این آماره در روش رگرسیون با داده‌های محدود بیشتر می‌باشد. با این حال، این دو روش اختلاف قابل ملاحظه‌ای با هم ندارند و می‌توان از هر دو روش با اطمینان استفاده کرد. در بین چهار روش بدترین نتایج مربوط به استفاده از کلیه داده‌ها بدون پیش‌پردازش و پس از آن مربوط به روش تست گاما است؛ که اگرچه دارای مقدار ضریب همبستگی قابل قبول بوده، ولی مقادیر آماره‌های MAE و RMSE آن تقریباً دو برابر مقادیر مربوط به روش‌های رگرسیون است.

نتایج مربوط به پیش‌بینی بارندگی ماهانه نیز نشان دهنده این است که رگرسیون (بدون محدودیت) بهترین روش تعیین اولیه متغیرهای مؤثر است. در مراتب بعدی، روش‌های رگرسیون با داده‌های محدود، روش گاما و روش بدون پیش‌پردازش بهترین نتایج را ارائه کرده است. متوسّط مقادیر بهترین پیش‌بینی‌ها و مشاهدات برای بارندگی و دما با آزمون مقایسه میانگین‌ها در سطح ۹۵٪ آزمون شده‌اند. نتایج در جدول ۷ آورده شده است.

جدول ۷: نتایج مقایسه میانگین آماری مقادیر نرمال شده داده‌های مشاهداتی و پیش‌بینی شده در دوره تست

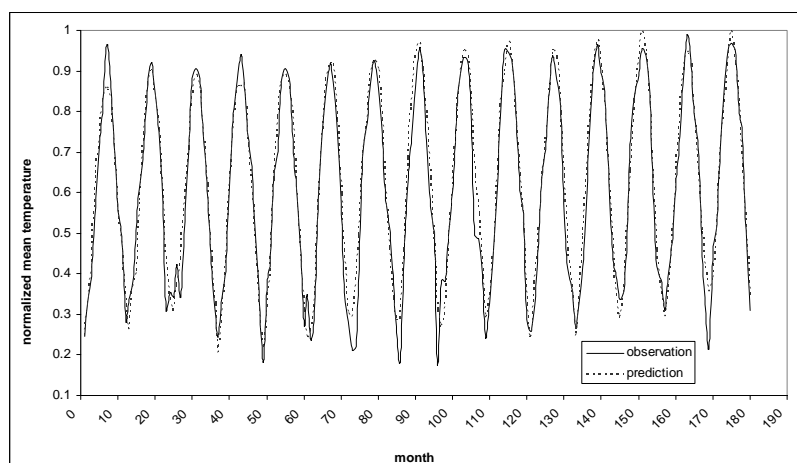
پارامتر	متوسّط مشاهدات	متوسّط پیش‌بینی	p-value
دمای متوسّط	۰/۶۱۱۳	۰/۶۱۶۴	۰/۱۵۴
بارش	۰/۶۶۵۸	۰/۶۵۳۱	۰/۴۲۱

با توجه به مقایسه میانگین‌های انجام شده، می‌توان نتیجه گرفت که اختلاف معنی‌داری بین مشاهدات و پیش‌بینی‌ها در هر دو مورد بارش و دما دیده نشده است. لذا، می‌توان نتیجه گرفت که به لحاظ آماری نیز پیش‌بینی‌ها و مشاهدات بسیار به هم نزدیک می‌باشند.

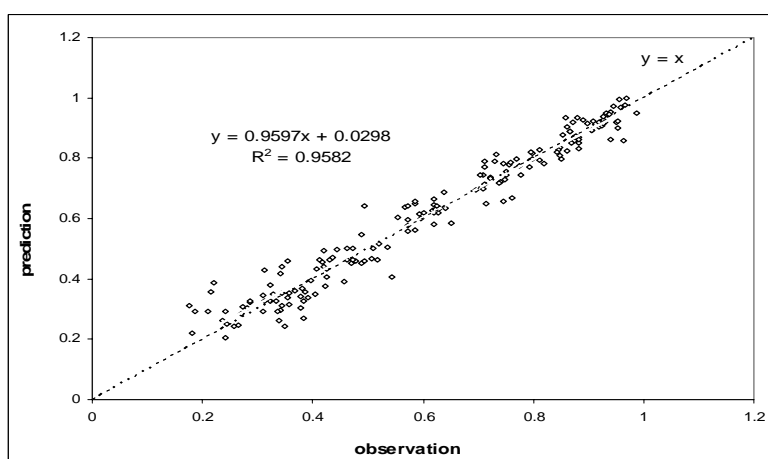
۲-۵-۳- مقایسه گرافیکی

شکل ۲ نشان دهنده مقادیر مشاهده شده و پیش‌بینی شده (در دوره تست) برای دمای متوسّط می‌باشد. همان طور که در این شکل نشان داده شده، مقادیر پیش‌بینی توانسته است به خوبی مقادیر مشاهداتی، به استثنای چند مورد از مقادیر حدّی، را پوشش دهد و می‌توان نتیجه گرفت که مدل ارائه شده توانایی خوبی

برای پیش‌بینی دمای ماهانه دارد. هم چنین خط رگرسیون بین مقادیر فوق در شکل ۳ رسم شده است. شیب این خط ۰/۹۶ بوده که به مقدار ۱ بسیار نزدیک است و نشان دهنده دقت مدل حاضر است.

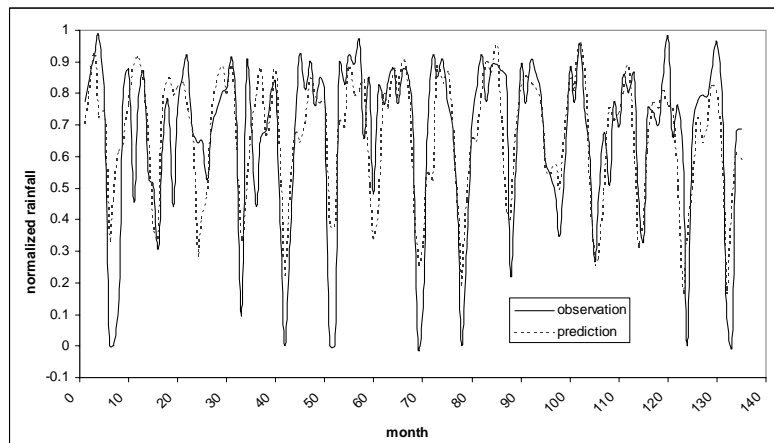


شکل ۲: مقایسه بین مقادیر مشاهداتی و پیش‌بینی دمای متوسط (دوره آزمایش)

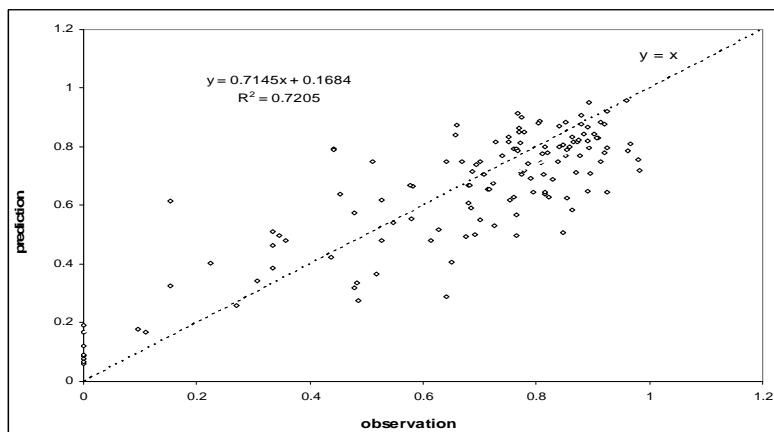


شکل ۳: خط رگرسیون بین مقادیر مشاهداتی و پیش‌بینی شده دمای متوسط (دوره آزمایش)

هم چنین مقایسه مقادیر مشاهداتی و پیش‌بینی بارندگی در شکل ۴ آورده شده است. واضح است که مدل فوق، توانسته است روند کلی مشاهدات را دنبال کند. ولیکن باید در استفاده از این مدل در موارد حدی احتیاط کرد. چرا که، در این نقاط عموماً فروبرآورد و به ندرت فرابراورد نشان داده است. هم چنین می‌توان خط رگرسیون بین مقادیر پیش‌بینی و مشاهداتی در شکل ۵ نشان داده شده است. همان طور که مشخص است در اطراف خط یک به یک پراکندگی نقاط مشاهده می‌شود و در نتیجه مقدار ضریب همبستگی آن به مقدار ۰/۸۴ رسیده است. این امر هم چنین نشان دهنده نتایج بالاست که مجدداً عدم دقت مدل بارش را در نقاط حدی نشان می‌دهد.



شکل ۴: مقایسه بین مقادیر مشاهداتی و پیش‌بینی بارش (دوره آزمایش)



شکل ۵: خط رگرسیون بین مقادیر مشاهداتی و پیش‌بینی شده بارش (دوره آزمایش)

۴- نتیجه‌گیری و پیشنهادات

دما و بارش به عنوان دو پارامتر مهم هواشناسی، خصوصاً در مناطق خشک و نیمه خشک مطرح هستند. تغییرات بسیاری از متغیرهای هواشناسی به شدت به پارامترهای گردش جوی، زمین و سطح اقیانوسی بزرگ مقیاس وابسته‌اند. برای بررسی تأثیر پارامترهای بزرگ مقیاس بر دو متغیر عمده هواشناسی (دما و بارش) با کمک شبکه‌های عصبی مصنوعی این تحقیق انجام شده است. علاوه بر این یکی دیگر از اهداف این تحقیق، بررسی تأثیر پیش‌پردازش داده‌های ورودی به شبکه عصبی با روش‌های مختلف در کارایی آن می‌باشد. در شبکه‌های مورد بررسی، توابع محرک مختلف با تعداد لایه‌های پنهان و نرون متفاوت به روش سعی و خطا مورد بررسی قرار گرفته است.

نتایج نشان دهنده این است که در هر دو پارامتر دما و بارش، روش پیش‌پردازش داده‌های ورودی رگرسیون بدون اعمال محدودیت بهترین نتایج را ارائه کرده است. این کارایی در مورد پارامتر دما بیشتر از بارش ماهانه است. پس از این روش، رگرسیون با داده‌های محدود، که نتایج نزدیکی به روش قبلی ارائه کرده است، در رتبه دوم بهترین روش‌ها قرار گرفته است. در نهایت، نتایج مربوط به هیچ‌گونه پیش‌پردازشی در داده‌های ورودی آخرین رتبه را، خصوصاً در مورد بارندگی به خود اختصاص داده است. این نتیجه نشان می‌-

دهد که روش رگرسیون می‌تواند به عنوان یک روش کارا، برای پیش‌پردازش شبکه عصبی مورد استفاده قرار گیرد. علاوه بر این، نتایج رگرسیون بدون محدودیت و بدون در نظر گرفتن دوره آموزش و تست، مشابه با نتایج رگرسیون با اعمال محدودیت و نیز جا به جایی دوره های آموزش و تست (برای آزمایش قابلیت پیش‌بینی مدل ارائه شده) است. این امر گویای آن است که نتایج رگرسیون عام، کارایی مناسبی برای پیش‌بینی و آزمون فرضیه ارائه نموده است.

با توجه به نتایج فوق می‌توان نتیجه گرفت که انجام پیش‌پردازش، خصوصاً در تحقیقات مشابه می‌تواند باعث بهبود کارایی و دقت پیش‌بینی نتایج شبکه‌های عصبی شود. هم‌چنین یکی دیگر از مزایای انجام پیش‌پردازش داده‌های ورودی صرفه‌جویی در زمان انجام تحلیل‌های شبکه عصبی است، که در مورد ساختن و مرحله آموزش شبکه برای داده های بدون پیش‌پردازش بسیار وقتگیر بوده و حتی در بعضی از کامپیوترهای موجود قابل اجرا نیست.

نتایج ارائه شده، چه به صورت معادلات و یا شبکه‌های عصبی در این تحقیق، می‌تواند به صورت کارایی برای پیش‌بینی مقادیر ماهانه دمای متوسط و بارش ایستگاه مورد استفاده واقع شود. با این حال لازم به تذکر است که مقدار برآورد این مدل‌ها در موارد حدی این پارامترها گاهاً کم تر و یا بیشتر از موارد مشاهده شده است. این امر لزوم انجام تحقیقات مشابه و بیشتر بر روی پارامترهای مهم حدی هواشناسی را می‌رساند.

علاوه بر این لازم است، که کارایی شاخص‌های بزرگ مقیاس در برآورد مقادیر فصلی پارامترهای هواشناسی مورد بررسی قرار گیرد و پارامترهای مؤثر بر آن‌ها مشخص شود. هم‌چنین پیشنهاد می‌شود که در تحقیقات آتی، از مقادیر پارامترهای هواشناسی با تأخیرات مختلف به عنوان یک پیش‌بینی کننده نیز استفاده شود. به طور مثال در برآورد دمای ماهانه، استفاده از این پارامتر به همراه تأخیرات مختلف آن، علاوه بر شاخص‌های بزرگ مقیاس استفاده شود. هم‌چنین می‌توان از روش‌های دیگر، از جمله به کارگیری سری‌های زمانی در تحلیل پارامترهای مختلف بهره جست. در نهایت با توجه به اهمیت شناخت پارامترهای هواشناسی در حوزه‌های مختلف کشاورزی، اقتصادی و اجتماعی و به منظور انجام مدیریت بهتر منابع، انجام تحقیق مشابه برای کلیه ایستگاه‌های استان خراسان رضوی و شناخت پارامترهای مؤثر بر آب و هوای آن‌ها ضروری می‌نماید.

۵- فهرست منابع

- ۱- اسفندیاری درآباد، فریبا، حسینی، اسعد، آزادی مبارکی، محمد و زهرا حجازی زاده (۱۳۸۹)، پیش-بینی میانگین دمای ماهانه ایستگاه سینوپتیک سنندج با استفاده از مدل شبکه‌ی عصبی مصنوعی پرسپترون چند لایه (MLP)، مجله جغرافیا، سال هشتم، شماره ۲۷.
- ۲- اصغری مقدم، اصغر، نورانی، وحید و عطاالله ندیری (۱۳۸۷)، مدل سازی بارش دشت تبریز با استفاده از شبکه‌های عصبی مصنوعی، مجله دانش کشاورزی، جلد ۱۸، شماره ۱.
- ۳- بازرگان لاری، عبدالرضا (۱۳۸۵)، رگرسیون خطی کاربردی، چاپ دوم، انتشارات مرکز نشر دانشگاه شیراز.
- ۴- برهانی داریان، علی رضا و احمد فاتحی مرج (۱۳۸۷)، کاربرد شبکه عصبی مصنوعی در پیش‌بینی جریان رودخانه با استفاده از شاخص های اقلیمی مطالعه موردی: حوضه آبریز نازلوچای، مجله دانشکده فنی دانشگاه تبریز، جلد ۳۵، شماره ۳، پیاپی ۵۱.
- ۵- حسینی، اسعد (۱۳۸۸)، برآورد و تحلیل دماهای حداکثر شهرستان اردبیل با استفاده از مدل تئوری شبکه‌های عصبی مصنوعی، پایان نامه کارشناسی ارشد جغرافیای طبیعی (اقلیم شناسی)، دانشگاه محقق اردبیلی.
- ۶- خلیلی، نجمه، خدائشناس، سعید رضا، داوری، کامران و محمد موسوی بایگی (۱۳۸۷)، پیش‌بینی بارش ماهانه با استفاده از شبکه‌های عصبی مصنوعی، مجله علوم و صنایع کشاورزی، ویژه آب و خاک، جلد ۲۲، شماره ۱.
- ۷- خورشیددوست، علی محمد و یوسف قویدل رحیمی (۱۳۸۵)، ارزیابی اثر پدیده انسو بر تغییرپذیری بارش‌های فصلی استان آذربایجان شرقی با استفاده از شاخص چند متغیره انسو، مجله پژوهش‌های جغرافیایی، شماره ۵۷.
- ۸- خورشیددوست، علی محمد، قویدل رحیمی، یوسف و کریم عباس زاده (۱۳۸۹)، کاربرد الگوهای کلان مقیاس جوی- اقیانوسی در تحلیل نوسانات بارش (مطالعه موردی: ایستگاه اهر)، مجله فضای جغرافیایی، شماره ۲۹.
- ۹- صلاحی، برومند، خورشیددوست، علی محمد و یوسف قویدل رحیمی (۱۳۸۶)، ارتباط نوسان های گردش جوی- اقیانوسی اطلس شمالی با خشکسالی‌های آذربایجان شرقی، مجله پژوهش‌های جغرافیایی، شماره ۶۰.
- ۱۰- علیجانی، بهلول (۱۳۹۰)، اقلیم شناسی سینوپتیک، انتشارات سمت، چاپ چهارم.
- ۱۱- علیزاده، امین (۱۳۸۰)، اصول هیدرولوژی کاربردی، چاپ سیزدهم، انتشارات آستان قدس رضوی، دانشگاه امام رضا (ع).
- ۱۲- فاتحی مرج، احمد، برهانی داریان، علیرضا و محمد حسین مهدیان (۱۳۸۵)، پیش‌بینی بارش فصلی با استفاده از پیوند از دور: مطالعه موردی حوزه آبریز دریاچه ارومیه، مجله علوم و فنون کشاورزی و منابع طبیعی، شماره ۳ (الف).

- ۱۳- فلاح قاهری، غلام عباس. و جواد خوشحال (۱۳۸۸)، پیش‌بینی بارش بهاره استان خراسان رضوی بر اساس سیگنال‌های بزرگ مقیاس اقلیمی با استفاده از شبکه‌های عصبی مصنوعی، مجله پژوهش‌های جغرافیای طبیعی، شماره ۶۹.
- ۱۴- قبائی سوق، محمد، مساعدی، ابوالفضل، حسام، موسی و ابوطالب هزارجریبی (۱۳۸۹)، ارزیابی تأثیر پیش‌پردازش پارامترهای ورودی به شبکه عصبی مصنوعی (ANNs) با استفاده از روش‌های رگرسیون گام به گام و گاما تست به منظور تخمین سریعتر تبخیر و تعرق روزانه، نشریه آب و خاک، جلد ۲۴، شماره ۳.
- ۱۵- قربانی، محمدعلی، فرسادی زاده، داود، فتحی، پیمان، چابک پور، جعفر و حسن جهانگیری (۱۳۸۸)، کتاب نرم افزارهای مهندسی آب، انتشارات نوپردازان.
- ۱۶- قوبدل رحیمی، یوسف (۱۳۸۴)، اثر الگوی بزرگ مقیاس گردش جوی-اقیانوسی انسو بر تغییرپذیری فصلی اقلیم در ایران: آثار ال‌نینو و لانینا بر تغییرپذیری بارش‌های بهاری در آذربایجان شرقی، فصل‌نامه مدرس علوم انسانی، دوره ۹، شماره ۴.
- ۱۷- معتمدی، محمد، احترامیان، کوروش و علی رضا شهاب‌فر (۱۳۸۶)، بررسی ارتباط از دور سیگنال هواشناسی ENSO با نوسانات بارندگی و دما در استان خراسان، مجله علوم محیطی، شماره ۴.
- ۱۸- موسوی بایگی، محمد، فلاح قاهری، غلام عباس و مجید حبیبی نوخندان (۱۳۸۷)، بررسی ارتباط نشانه‌های پهن مقیاس اقلیمی با بارش خراسان، مجله علوم کشاورزی و منابع طبیعی، جلد ۱۵، شماره ۲.
- ۱۹- ناظم السادات، محمد جعفر، انصاری بصیر، ارمان و محمدرضا پیشوایی (۱۳۸۶)، ارزیابی سطح معنی داری برای پیش‌بینی دوران خشکسالی و ترسالی فصل پائیز و شش ماهه سرد ایران بر اساس وضعیت فازهای تابستانه ENSO، مجله تحقیقات منابع آب ایران، شماره ۱.
- ۲۰- یاراحمدی، داریوش و قاسم عزیزی (۱۳۸۶)، تحلیل چند متغیره ارتباط میزان بارش فصلی ایران و شاخص‌های اقلیمی، مجله پژوهش‌های جغرافیایی، شماره ۶۲.

- 21- Bannayan, Mohammad, Sanjani, Sarah, Alizadeh, Amin, Sadeghi Lotfabadi, Sajad and Azadeh Mohamadian (2010), Association between climate indices, aridity index, and rainfed crop yield in northeast of Iran, *Field Crops Research*, 118: pp 105-114.
- 22- Demuth, Howard and Mark Beale (2000), *Neural Network Toolbox User's Guide*, Copyright 1992-2002, By The Math Works, Inc, Version 4, 840pp.
- 23- Durrant Peter (2001), *winGamma: A non-linear data analysis and modeling tool with applications to flood prediction*. PhD thesis, Department of Computer Science, Cardiff University, Wales, UK.
- 24- Kim, Maeng-Ki, Kim, Yeon-Hee and Woo-Seop Lee (2007), Seasonal prediction of Korean regional climate from preceding large-scale climate indices. *Int. J. Climatol*, 27: pp 925-934.
- 25- Lloyd-Hughes, Benjamin and Mark A. Saunders (2002), Seasonal prediction of European spring precipitation from El Nino-Southern oscillation and local sea-surface temperatures, *Int. J. Climatol*, 22: pp 1-14.
- 26- McCulloch, Warren and Walter Pitts (1943), a logical calculus of the ideas immanent in nervous activity. *Bull Math Biophys*. 5: pp115-133. Reprinted in Anderson & Rosenfeld, 1988, pp 18-28.
- 27- Moghaddamia, Ali Reza, Ghafari Gousheh, Mostafa, Piri, Jamshid, Amin, S. and Dawei Han (2009), *Evaporation estimation using artificial neural networks and*

- adaptive neuro-fuzzy inference system techniques, Advances in Water Resources, 32: pp 88-97.*
- 28- Nazemosadat, Mohammad Jafar and Ahmad Reza Ghasemi (2004), *quantifying the ENSO related shifts in the intensity and probability of drought and wet periods in Iran, Journal of Climate, 17 (20): pp 4005-4018.*
- 29- Noori, Roohollah, Karbassi, Ali Reza, Moghaddamnia, Ali Reza, Han, Dawei, Zokaei-Ashtiani, Mohammad Hossein, Farokhnia, Ali and Mostafa Ghafari Gousheh (2011), *Assessment of input variables determination on the SVM model performance using PCA, Gamma test, and forward selection techniques for monthly stream flow prediction, Journal of Hydrology, 401: pp 177-189.*
- 30- Rozhkov, Vide A., Nurberdiev, M. and A. Rangavar (2007), *agro ecological bases for raising the productivity of degraded soils in Khorasan Province of Iran, Eurasian Soil Sci, 40 (12): pp 1335-1342.*
- 31- Schwing, Fing B., Murphree, Tom and Patrick M. Green (2002), *The Northern Oscillation Index (NOI): a new climate index, Progress in Oceanography, 53: pp 115-139.*
- 32- Thompson, David W.J. and John M. Wallace (2000), *annular modes in the extra tropical circulation, Part I: month-to-month variability, Journal of Climate, 13: pp 1000-1016.*
- 33- Valverde Ramí' rez, Maria Cleofe, Campos Velho, Haroldo Fraga and Nelson Jesus Ferreira (2005), *artificial neural network technique for rainfall forecasting applied to the Sa' o Paulo region, Journal of Hydrology, 301: pp146-162*
- 34- Wallace, John M. and David S. Gutzler (1981), *Tele-connections in the geopotential height field during the northern hemisphere winter, Monthly weather Review, 109: pp 784-812.*
- 35- Wang, Chunzai and David B. Enfield (2001), *the tropical western hemisphere warm pool, Geophysical research letters, 28: pp1635-1638.*