

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ



دانشکده مهندسی عمران

پایان نامه کارشناسی ارشد مهندسی عمران-محیط زیست

استفاده از الگوریتم جنگل تصادفی و شاخص بارش استاندارد به منظور پیش بینی خشکسالی و ارزیابی پیامدهای زیست محیطی آن

نگارنده: زینب ترابی

استاد راهنما: دکتر رمضان واقعی

استاد مشاور: مهندس علیرضا قائمی

دی ۱۴۰۰



مدیریت تحصیلات تکمیلی

باسمه تعالی

فرمهای ارزشیابی پایان نامه کارشناسی ارشد

مربوط به ورودی های ۹۴ به بعد

شماره: ۵۸۹/ع = ت

تاریخ: ۱۴۰۱/۰۱/۲۷

ویرایش:

تاریخ و امضاء و مهر دانشکده

نام و نام خانوادگی رئیس دانشکده: دکتر رضا نادری

فرم شماره (۳) صورتجلسه نهایی دفاع از پایان نامه دوره کارشناسی ارشد

با نام و یاد خداوند متعال، ارزیابی جلسه دفاع از پایان نامه کارشناسی ارشد خانم زینب ثرابی با شماره دانشجویی ۹۸۰۳۴۰۴ رشته مهندسی عمران گرایش محیط زیست تحت عنوان استفاده از الگوریتم جنگل تصادفی و شاخص بارش استاندارد به منظور پیش بینی خشکسالی و ارزیابی پیامدهای زیست محیطی آن که در تاریخ ۱۴۰۰/۱۰/۸ با حضور هیأت محترم داوران در دانشگاه صنعتی شاهرود برگزار شد به شرح ذیل اعلام می گردد:

الف) درجه عالی: نمره ۱۹-۲۰ ☒ (ب) درجه خیلی خوب: نمره ۱۸/۹۹ - ۱۸ ☐
 ج) درجه خوب: نمره ۱۷/۹۹ - ۱۶ ☐ (د) درجه متوسط: نمره ۱۵/۹۹ - ۱۴ ☐
 ه) کمتر از ۱۴ غیر قابل قبول و نیاز به دفاع مجدد دارد ☐
 نوع تحقیق: نظری ☒ عملی ☐

عضو هیأت داوران	نام و نام خانوادگی	مرتبه علمی	امضاء
۱- استاد راهنمای اول	دکتر رمضان واقعی	استادیار دانشکده مهندسی عمران	
۲- استاد راهنمای دوم			
۳- استاد مشاور	مهندس علیرضا قائمی	کارشناس ارشد مهندسی و مدیریت منابع آب	
۴- استاد داور اول	دکتر طاهره تقی زاده فیروزجایی	استادیار دانشکده مهندسی عمران	
۵- استاد داور دوم	دکتر صمد امامقلی زاده	دانشیار دانشکده مهندسی عمران	
۶- نماینده تحصیلات تکمیلی	دکتر جلیل شفائی	استادیار دانشکده مهندسی عمران	

نام و نام خانوادگی رئیس دانشکده: دکتر رضا نادری



تقدیم به پدر و مادر عزیزم

خدای رابی شاکرم که از روی کرم، پدر و مادری فداکار نصیبم ساخته تا از سایه وجودشان در راه کسب علم و دانش تلاش نمایم. والدینی که بودنشان تاج افتخاری است بر سرم و نامشان دلیلی است بر بودنم، چرا که این دو وجود، پس از پروردگار، مایه هستی ام بوده اند و ستم را اگر قند و راه رفتن را در وادی زندگی آموختند. آموزگارانی که برایم زندگی، بودن و انسان بودن را معنا کردند...

قدردانی

شکرشایان نثار ایزدمنان که توفیق رارفتق را هم ساخت تا این پایان نامه را به پایان برسانم. از محضر ارزشمنند پروماد عزیزم به خاطر همه ی تلاشهای محبت آمیزی که در دوران مختلف زندگی ام انجام داده اند و بامهربانی چگونه زیستن را به من آموخته اند، کمال شکر و قدردانی را دارم.

همچنین از اساتذراهنمافاضل و اندیشمندجناب آقای دکتر رمضان واقعی و مشاور محترم جناب آقای مهندس علیرضا قانعی که همواره مرا مورد لطف و محبت خود قرار داده اند، کمال شکر و قدردانی را دارم. بدون شک انجام این تحقیق مدیون حمایت و پشتیبانی ایشان است.

تهدنامه

اینجانب زینب ترابی دانشجوی دوره کارشناسی ارشد رشته مهندسی عمران-محیط زیست دانشکده عمران دانشگاه صنعتی شاهرود نویسنده پایان نامه استفاده از الگوریتم جنگل تصادفی و شاخص بارش استاندارد به منظور پیش بینی خشکسالی و ارزیابی پیامدهای زیست محیطی آن تحت راهنمایی دکتر رمضان واقعی و مهندس علیرضا قائمی متعهد می شوم.

- تحقیقات در این پایان نامه توسط اینجانب انجام شده است و از صحت و اصالت برخوردار است .
- در استفاده از نتایج پژوهشهای محققان دیگر به مرجع مورد استفاده استناد شده است .
- مطالب مندرج در پایان نامه تاکنون توسط خود یا فرد دیگری برای دریافت هیچ نوع مدرک یا امتیازی در هیچ جا ارائه نشده است .
- کلیه حقوق معنوی این اثر متعلق به دانشگاه صنعتی شاهرود می باشد و مقالات مستخرج با نام « دانشگاه صنعتی شاهرود » و یا « Shahrood University of Technology » به چاپ خواهد رسید .
- حقوق معنوی تمام افرادی که در به دست آمدن نتایج اصلی پایان نامه تأثیرگذار بوده اند در مقالات مستخرج از پایان نامه رعایت می گردد.
- در کلیه مراحل انجام این پایان نامه ، در مواردی که از موجود زنده (یا بافتهای آنها) استفاده شده است ضوابط و اصول اخلاقی رعایت شده است .
- در کلیه مراحل انجام این پایان نامه، در مواردی که به حوزه اطلاعات شخصی افراد دسترسی یافته یا استفاده شده است اصل رازداری ، ضوابط و اصول اخلاق انسانی رعایت شده است .

تاریخ

امضای دانشجو

مالکیت نتایج و حق نشر

کلیه حقوق معنوی این اثر و محصولات آن (مقالات مستخرج ، کتاب ، برنامه های رایانه ای ، نرم افزار ها و تجهیزات ساخته شده است) متعلق به دانشگاه صنعتی شاهرود می باشد . این مطلب باید به نحو مقتضی در تولیدات علمی مربوطه ذکر شود .
استفاده از اطلاعات و نتایج موجود در پایان نامه بدون ذکر مرجع مجاز نمی باشد.

خشکسالی باعث تشدید بحران آب و ایجاد خسارت های جبران ناپذیر به جوامع می شود. در سال های اخیر استفاده از روش های یادگیری ماشین در ارزیابی خشکسالی مورد توجه محققین قرار گرفته است. هدف این تحقیق، ارزیابی خشکسالی در شهرهای زابل و زاهدان طی بازه زمانی (۲۰۲۰-۱۹۹۰) است که برای این منظور شاخص بارش استاندارد (SPI) در مقیاس فصلی و سالانه، تحلیل خودهمبستگی جزئی (PACF) و الگوریتم جنگل تصادفی (RF) استفاده شده است. پس از محاسبه SPI، نتایج تحلیل PACF برای SPI، به عنوان ورودی های مدل در نظر گرفته شده اند. داده های آموزش و آزمایش با دو ورودی متفاوت بررسی شده اند. طبق نتایج تحقیق در مقیاس فصلی و سالانه SPI، طبقه خشکسالی تقریباً نرمال (N) بیشترین فراوانی وقوع را در هر دو ایستگاه دارد و شدت خشکسالی در ایستگاه زاهدان در مقیاس فصلی و سالانه نسبت به ایستگاه زابل زیادتر بوده و اوج خشکسالی سالانه هر دو ایستگاه در سال ۲۰۰۱ رخ داده است، همچنین اوج خشکسالی هر دو ایستگاه در مقیاس فصلی در فصولی رخ داده است که فاقد بارش بوده اند. بر اساس تحلیل PACF، مناطق مورد مطالعه در طی بازه زمانی ۳۰ ساله دچار تغییر اقلیم شده اند. نتایج پیش بینی SPI توسط چندین پارامتر آماری بررسی شده اند. شاخص توافق (IOA) در ایستگاه زابل برای داده های آموزش با در نظر گرفتن چهار تاخیر زمانی (۲، ۴، ۶ و ۱۲ ماهه) و سه تاخیر زمانی (۲، ۴ و ۶ ماهه) به ترتیب برابر ۰/۹۶۴۸ و ۰/۹۲۵۶ بوده و برای داده های آزمایش به ترتیب ۰/۸۵۵۶ و ۰/۸۶۷۳ می باشد. IOA در ایستگاه زاهدان برای داده های آموزش با چهار تاخیر زمانی (۲، ۴، ۶ و ۸ ماهه) و سه تاخیر زمانی (۲، ۴ و ۶ ماهه) به ترتیب برابر ۰/۹۴۹۵ و ۰/۹۲۰۵ است و برای داده های آزمایش به ترتیب ۰/۷۴۰۸ و ۰/۶۳۰۳ می باشد. سایر پارامتر های آماری بررسی شده نیز نشان دهنده قابلیت مطلوب RF در تخمین SPI می باشند.

کلمات کلیدی: خشکسالی، شاخص بارش استاندارد (SPI)، الگوریتم جنگل تصادفی (RF)، تحلیل خودهمبستگی جزئی (PACF)

لیست مقالات مستخرج از پایان نامه

- ۱- ارزیابی خشکسالی به کمک شاخص بارش استاندارد و الگوریتم جنگل تصادفی. ترابی، ز.، واقعی، ر.، قائمی، ع. (۱۳۹۹). نشریه علمی پژوهشی هواشناسی و علوم جو، جلد ۳، شماره ۴، صفحات ۳۹۰-۴۰۵.

فهرست مطالب

۱	فصل اول: کلیات
۲	۱-۱- مقدمه.....
۳	۲-۱- بیان موضوع.....
۴	۳-۱- ضرورت انجام تحقیق.....
۵	۴-۱- اهداف تحقیق.....
۶	۵-۱- ساختار پایان نامه.....
۹	فصل دوم: مروری بر مطالعات پیشین
۱۰	۱-۲- مقدمه.....
۱۱	۲-۱-۱- شاخص های خشکسالی هواشناسی.....
۱۱	۲-۱-۲- شاخص های خشکسالی کشاورزی.....
۱۱	۲-۱-۳- شاخص های خشکسالی هیدرولوژیکی.....
۱۲	۲-۲- کاربرد الگوریتم جنگل تصادفی در علوم مختلف.....
۱۳	۳-۲- کاربرد الگوریتم جنگل تصادفی در محیط زیست.....
۱۵	۴-۲- کاربرد شاخص های خشکسالی.....
۱۷	۵-۲- پیش بینی شاخص های مختلف خشکسالی با روش های یادگیری ماشین.....
۱۹	۶-۲- جمع بندی مرور بر مطالعات پیشین.....
۲۱	فصل سوم: مواد و روش ها
۲۲	۱-۳- مقدمه.....
۲۲	۲-۳- محدوده مورد مطالعه.....

- ۳-۳- ویژگی ایستگاه های مورد مطالعه..... ۲۴
- ۳-۴- داده های مورد استفاده ۲۵
- ۳-۵- آزمون و بازسازی داده ها..... ۲۵
- ۳-۵-۱- بررسی تکمیل بودن داده ها ۲۶
- ۳-۵-۲- بررسی همگن بودن داده ها..... ۲۶
- ۳-۵-۳- بررسی روند داده ها یا تصادفی بودن داده ها ۲۷
- ۳-۶- شاخص بارش استاندارد (SPI)..... ۲۷
- ۳-۶-۱- تاریخچه SPI ۲۷
- ۳-۶-۲- تعریف SPI..... ۲۷
- ۳-۶-۳- محاسبه SPI..... ۲۸
- ۳-۷- تابع خودهمبستگی جزئی (PACF)..... ۳۱
- ۳-۸- روش های یادگیری ماشین..... ۳۲
- ۳-۹- روش بگینگ..... ۳۳
- ۳-۱۰- الگوریتم جنگل تصادفی (RF)..... ۳۳
- ۳-۱۰-۱- تاریخچه الگوریتم RF..... ۳۳
- ۳-۱۰-۲- تعریف الگوریتم RF..... ۳۳
- ۳-۱۰-۳- پارامتر های الگوریتم RF..... ۳۴
- ۳-۱۰-۳-۱- تعداد متغیر های تصادفی در هر گره (m_{try})..... ۳۴
- ۳-۱۰-۳-۲- تعداد درختان جنگل (n_{tree})..... ۳۴
- ۳-۱۰-۴- نحوه اجرا RF..... ۳۴
- ۳-۱۰-۵- قابلیت های RF..... ۳۵

- ۳-۱۰-۱-ارزیابی الگوریتم به کمک نمونه های خارج از کیسه(OOB).....۳۵
- ۳-۱۰-۲- اندازه گیری اهمیت متغیر ها.....۳۵
- ۳-۱۰-۶- مزایای RF.....۳۶
- ۳-۱۱- پارامتر های آماری.....۳۸
- ۳-۱۱-۱- ضریب تبیین (R^2).....۳۸
- ۳-۱۱-۲- جذر میانگین مربعات (MSE).....۳۸
- ۳-۱۱-۳- خطای جذر میانگین مربعات (RMSE).....۳۸
- ۳-۱۱-۴- میانگین خطای مطلق (MAE).....۳۹
- ۳-۱۱-۵- شاخص توافق (IOA).....۳۹
- ۳-۱۱-۶- نرمال شده خطای جذر میانگین مربعات (NRMSE).....۳۹
- ۳-۱۲- مراحل انجام تحقیق.....۴۰
- ۳-۱۲-۱- اطمینان یافتن از همگن بودن و تصادفی بودن داده ها.....۴۰
- ۳-۱۲-۲- تشکیل سری زمانی و آماده سازی داده های بارش برای برآورد SPI.....۴۰
- ۳-۱۲-۳- برازش تابع.....۴۰
- ۳-۱۲-۴- محاسبه SPI.....۴۰
- ۳-۱۲-۵- انجام تحلیل PACF.....۴۰
- ۳-۱۲-۶- پیش بینی SPI به کمک الگوریتم RF.....۴۱
- ۳-۱۲-۷- اعتبار سنجی الگوریتم RF.....۴۱
- ۳-۱۲-۸- ارزیابی اثرات زیست محیطی خشکسالی.....۴۱

فصل چهارم: نتایج و بحث

۴۳

۴۴	۱-۴- مقدمه.....
۴۵	۲-۴- بررسی همگن بودن داده های بارش.....
۴۶	۳-۴- بررسی وضعیت روند داده های بارش.....
۴۷	۴-۴- نتایج برآورد پارامتر های تابع گاما.....
۴۸	۵-۴- نتایج SPI در مقیاس ۳ و ۱۲ ماهه.....
۵۹	۶-۴- نتایج تحلیل PACF در مقیاس ۳ و ۱۲ ماهه.....
۶۳	۷-۴- نتایج پیش بینی SPI به کمک الگوریتم RF.....
۷۱	۸-۴- اعتبار سنجی الگوریتم RF.....
۷۴	۹-۴- ارزیابی اثرات زیست محیطی خشکسالی در منطقه مورد مطالعه.....
۷۵	۱-۹-۴- کیفیت هوا و عوامل اقلیمی.....
۷۶	۲-۹-۴- خاک ، زمین شناسی و ژئومورفولوژی.....
۷۶	۳-۹-۴- آب.....
۷۶	۴-۹-۴- بوم شناسی.....

فصل پنجم: نتیجه گیری و پیشنهادات

۷۹

۸۰	۱-۵- مقدمه.....
۸۰	۲-۵- نتیجه گیری.....
۸۱	۳-۵- مقایسه نتایج تحقیق با تحقیقات مشابه پیشین.....
۸۲	۴-۵- پیشنهادات جهت مطالعات آتی.....

۸۳

مراجع

فهرست اشکال

- شکل ۳-۱ نقشه عوارض طبیعی کشور ایران..... ۲۲
- شکل ۳-۲ نقشه استان سیستان و بلوچستان و محدوده مطالعاتی زابل و زاهدان..... ۲۴
- شکل ۳-۳ نمایی از روش یادگیری ماشین..... ۳۲
- شکل ۳-۴ نمایی از به کارگیری الگوریتم جنگل تصادفی در این مطالعه..... ۳۷
- شکل ۳-۵ فلوجارت پایان نامه..... ۴۲
- شکل ۴-۱ نمودار درصد رخداد طبقات مختلف خشکسالی SPI 3 در ایستگاه زابل ۵۴
- شکل ۴-۲ نمودار درصد رخداد طبقات مختلف خشکسالی SPI 12 در ایستگاه زابل..... ۵۴
- شکل ۴-۳ نمودار درصد رخداد طبقات مختلف خشکسالی SPI 3 در ایستگاه زاهدان..... ۵۵
- شکل ۴-۴ نمودار درصد رخداد طبقات مختلف خشکسالی SPI 12 در ایستگاه زاهدان..... ۵۵
- شکل ۴-۵ نمودار SPI ایستگاه زابل و زاهدان در فصل زمستان..... ۵۶
- شکل ۴-۶ نمودار SPI ایستگاه زابل و زاهدان در فصل بهار..... ۵۶
- شکل ۴-۷ نمودار SPI ایستگاه زابل و زاهدان در فصل تابستان..... ۵۷
- شکل ۴-۸ نمودار SPI ایستگاه زابل و زاهدان در فصل پاییز..... ۵۷
- شکل ۴-۹ نمودار SPI ایستگاه زابل و زاهدان در مقیاس سالانه..... ۵۷
- شکل ۴-۱۰ نمودار تحلیل PACF در ایستگاه زابل برای SPI 3..... ۵۹
- شکل ۴-۱۱ نمودار تحلیل PACF در ایستگاه زابل برای SPI 12..... ۶۰
- شکل ۴-۱۲ نمودار تحلیل PACF در ایستگاه زاهدان برای SPI 3..... ۶۱
- شکل ۴-۱۳ نمودار تحلیل PACF در ایستگاه زاهدان برای SPI 12..... ۶۲
- شکل ۴-۱۴ ضریب تبیین SPI فصلی زابل با تاخیر زمانی ۲و۴و۶و۱۲ در مجموعه آموزش..... ۶۵

- شکل ۴-۱۵ ضریب تبیین SPI فصلی زابل با تاخیر زمانی ۲و۴و۶و۱۲ در مجموعه آزمایش.....۶۵
- شکل ۴-۱۶ ضریب تبیین SPI فصلی زابل با تاخیر زمانی ۲و۴و۶ در مجموعه آموزش.....۶۶
- شکل ۴-۱۷ ضریب تبیین SPI فصلی زابل با تاخیر زمانی ۲و۴و۶ در مجموعه آزمایش.....۶۶
- شکل ۴-۱۸ ضریب تبیین SPI فصلی زاهدان با تاخیر زمانی ۲و۴و۶و۸ در مجموعه آموزش.....۶۷
- شکل ۴-۱۹ ضریب تبیین SPI فصلی زاهدان با تاخیر زمانی ۲و۴و۶و۸ در مجموعه آزمایش.....۶۷
- شکل ۴-۲۰ ضریب تبیین SPI فصلی زاهدان با تاخیر زمانی ۲و۴و۶ در مجموعه آموزش.....۶۸
- شکل ۴-۲۱ ضریب تبیین SPI فصلی زاهدان با تاخیر زمانی ۲و۴و۶ در مجموعه آزمایش.....۶۸
- شکل ۴-۲۲ نمودار مقادیر مشاهداتی و پیش بینی شده SPI فصلی زابل با تاخیر زمانی ۲و۴و۶.....۶۹
- شکل ۴-۲۳ نمودار مقادیر مشاهداتی و پیش بینی شده SPI فصلی زابل با تاخیر زمانی ۲و۴و۶و۱۲.....۶۹
- شکل ۴-۲۴ نمودار مقادیر مشاهداتی و پیش بینی شده SPI فصلی زاهدان با تاخیر زمانی ۲و۴و۶.....۷۰
- شکل ۴-۲۵ نمودار مقادیر مشاهداتی و پیش بینی شده SPI فصلی زاهدان با تاخیر زمانی ۲و۴و۶و۸.....۷۰
- شکل ۴-۲۶ تصویر ماهواره ای از طوفان های گرد و غبار در تاریخ ۱۳۸۸/۶/۱۴.....۷۴
- شکل ۴-۲۷ اثرات زیست محیطی خشکسالی در محدوده مورد مطالعه.....۷۷

فهرست جداول

- جدول ۱-۱ درجه بندی برخی از حوادث طبیعی با توجه به ویژگی و اثرات آن ها.....۲
- جدول ۱-۳ مشخصات ایستگاه زابل و زاهدان.....۲۴
- جدول ۲-۳ پارامتر های آماری سری زمانی بارش ماهانه در ایستگاه زابل و زاهدان.....۲۵
- جدول ۳-۳ طبقه بندی خشکسالی بر اساس SPI.....۳۰
- جدول ۱-۴ آزمون ران تست در ایستگاه زابل.....۴۵
- جدول ۲-۴ آزمون ران تست در ایستگاه زاهدان.....۴۶
- جدول ۳-۴ آزمون تاوکنندال در ایستگاه زابل.....۴۶
- جدول ۴-۴ آزمون تاوکنندال در ایستگاه زاهدان.....۴۶
- جدول ۵-۴ پارامتر های تابع گاما در ایستگاه زابل و مقیاس ۳ ماهه.....۴۷
- جدول ۶-۴ پارامتر های تابع گاما در ایستگاه زابل و مقیاس ۱۲ ماهه.....۴۷
- جدول ۷-۴ پارامتر های تابع گاما در ایستگاه زاهدان و مقیاس ۳ ماهه.....۴۷
- جدول ۸-۴ پارامتر های تابع گاما در ایستگاه زاهدان و مقیاس ۱۲ ماهه.....۴۷
- جدول ۹-۴ SPI در ایستگاه زابل و زاهدان در مقیاس ۳ ماهه.....۴۸
- جدول ۱۰-۴ SPI در ایستگاه زابل و زاهدان در مقیاس ۱۲ ماهه.....۵۲
- جدول ۱۱-۴ فراوانی رخداد طبقات مختلف خشکسالی ایستگاه زابل و زاهدان.....۵۳
- جدول ۱۲-۴ درصد رخداد طبقات مختلف خشکسالی ایستگاه زابل و زاهدان.....۵۳
- جدول ۱۳-۴ کمترین مقدار SPI، تاریخ وقوع اوج خشکسالی و شدت خشکسالی.....۵۸
- جدول ۱۴-۴ تحلیل PACF در ایستگاه زابل برای SPI 3.....۵۹
- جدول ۱۵-۴ تحلیل PACF در ایستگاه زابل برای SPI 12.....۶۰

جدول ۴-۱۶	تحلیل PACF در ایستگاه زاهدان برای SPI 3.....	۶۱
جدول ۴-۱۷	تحلیل PACF در ایستگاه زاهدان برای SPI 12.....	۶۲
جدول ۴-۱۸	تنظیمات الگوریتم جنگل تصادفی.....	۶۴
جدول ۴-۱۹	اعتبارسنجی پیش بینی SPI فصلی ایستگاه زابل.....	۷۱
جدول ۴-۲۰	اعتبارسنجی پیش بینی SPI فصلی ایستگاه زاهدان.....	۷۳

فصل اول: کلیات

۱-۱- مقدمه

خشکسالی^۱ پدیده ای طبیعی می‌باشد و هنگامی رخ می‌دهد که منطقه ای در درازمدت دچار کاهش بارندگی شود [۱]. خشکسالی ناشی از کمبود بارندگی و افزایش دما می‌باشد که در هر وضعیت اقلیمی ممکن است رخ دهد و توصیف زمانی و مکانی آن دشوار می‌باشد [۲]. خشکی^۲ ویژگی دائمی اقلیم می‌باشد و مختص نواحی کم باران است اما خشکسالی پدیده ای موقتی است و در هر شرایط اقلیمی ممکن است به وجود آید. اثرات خشکسالی غیرساختاری است و در سطح وسیع اتفاق می‌افتد بنابراین کمی نمودن پیامد های خشکسالی دشوار می‌باشد. خشکسالی در مقایسه با دیگر بلایای جوی بر اساس خطراتی که ایجاد می‌نماید، در جایگاه نخست قرار گرفته است که جزئیات این درجه بندی در جدول ۱-۱ آورده شده است [۳].

جدول ۱-۱ درجه بندی برخی از حوادث طبیعی با توجه به ویژگی و اثرات آن ها (پر اهمیت ترین ۱ و کم اهمیت ترین ۵) [۳].

رخداد	شدت	مدت	وسعت	تلفات جانی	خسارات اقتصادی	اثر اجتماعی	اثرات شدید درازمدت
خشکسالی	۱	۱	۱	۱	۱	۱	۱
چرخنده های حاره ای	۱	۲	۲	۲	۲	۲	۱
سیل های منطقه ای	۲	۲	۲	۱	۱	۱	۲
زلزله	۱	۵	۱	۲	۱	۱	۲
آتشفشان	۱	۴	۴	۲	۲	۲	۱
زمین لغزه	۴	۲	۲	۴	۴	۴	۵
بارش های سیل آسا	۳	۵	۴	۴	۴	۴	۵

¹Drought

² Aridity

شناسایی و مشخص کردن شروع و همچنین نقطه پایان خشکسالی مشکل بوده و به همین سبب خشکسالی یک پدیده خزنده است. در بسیاری از نقاط جهان علاوه بر کمبود بارش، دما، باد و رطوبت نسبی پایین بر روی خشکسالی موثر می‌باشد. وقوع خشکسالی عمدتاً از تغییرات ذاتی گردش هوا حاصل می‌شود. همچنین ممکن است به عواملی مانند حمل خاکستر و غبار آتشفشانی بستگی داشته باشد که باعث کاهش تابش خورشید به سطح زمین می‌شود. خشکسالی از نظر خصوصیات اقلیمی و گستردگی فضایی، متنوع می‌باشد لذا در بررسی خشکسالی، ویژگی‌های محلی باید در نظر گرفته شود. پیش‌بینی‌های کوتاه مدت و بلند مدت خشکسالی به تصمیم‌گیرندگان کمک می‌کند تا آن‌ها بتوانند به وقایع خشکسالی به شیوه‌ای دقیق‌تری پاسخ دهند [۴].

۱-۲- بیان موضوع

آب ماده‌ای حیاتی بخش و ضروری است. حدود ۷۰ تا ۷۵ درصد زمین را آب پوشانده است. حدود ۳ درصد آب به صورت برف و یخ، ۱ درصد به صورت آب شیرین مایع و تقریباً ۹۶ درصد آب مربوط به اقیانوس‌ها و دریاها می‌باشد [۵]. عواملی مانند رشد جمعیت، افزایش جمعیت شهرها، صنعتی شدن، متغیر بودن الگوهای بارش و شرایط جوی، گرمایش زمین، عدم استفاده از روش‌های مدرن در آبیاری زمین‌های کشاورزی و عدم فرهنگ سازی صحیح در مصرف بهینه آب باعث ایجاد بحران آب شده است و خشکسالی این بحران را تشدید می‌نماید. در کشور ایران حدود ۹۰ درصد مناطق به لحاظ اقلیمی خشک یا نیمه خشک هستند. رشته کوه البرز در قسمت شمالی و زاگرس در قسمت غربی، مانع ورود رطوبت به بخش‌های مختلف می‌باشد. میانگین بارش سالانه در ایران حدود ۲۵۱ میلی‌متر یا ۴۱۳ میلیارد مترمکعب است که از یک سوم میانگین بارش جهانی (۸۳۱ میلی‌متر) کم‌تر می‌باشد. تنوع مکانی و زمانی بارش در ایران بالا می‌باشد. بخش‌هایی از جنوب دریای خزر، حداکثر ۲۰۰۰ میلی‌متر بارندگی سالانه دارند در صورتی که بارندگی در بخش‌های مرکزی و شرقی کشور، کم‌تر از ۵۰ میلی‌متر در سال می‌باشد [۶].

۱-۳- ضرورت انجام تحقیق

محدوده مطالعاتی ایستگاه سینوپتیک شهرهای زابل و زاهدان، واقع در استان سیستان و بلوچستان می‌باشد. این محدوده دارای آب و هوای کویری و نیمه بیابانی است و متوسط بارندگی سالانه آن حدود ۱۱۰ میلی متر و متوسط دمای آن ۲۲ تا ۳۷ درجه سانتی گراد است. رطوبت نسبی کم، کم بودن بارندگی سالانه نسبت به بارندگی متوسط کشور و تغییرات دمایی بالا بخشی از ویژگی های این ناحیه می‌باشد [۷]. در دهه گذشته دریاچه هامون که از رودخانه هیرمند کشور افغانستان سرچشمه می‌گیرد به دلیل سدسازی های این کشور و کاهش بارندگی خشک شده است و موجب به خطر افتادن زیستگاه گیاهان و جانوران شده است. یکی دیگر از ویژگی های مهم این منطقه، باد های ۱۲۰ روزه است که معمولا از اواخر بهار تا اوایل پاییز ادامه دارد. سرعت غالب باد های ۱۲۰ روزه نزدیک به ۱۰ متر بر ثانیه می‌باشد که موجب افزایش نرخ تبخیر، افزایش فرسایش خاک، کاهش پوشش گیاهی، طوفان های گرد و غبار و مشکلات تنفسی برای افراد می‌گردد [۸]. طوفان های گرد و غبار باعث تغییر در شرایط اقلیمی، چرخه بیولوژیکی، زمین شناسی و محیط زیست انسان می‌گردد. آئروسل های معدنی ایجاد شده از طوفان های گرد و غبار بر روی تشکیل ابر و کمیت نزولات جوی اثر می‌گذارد [۹]. طوفان های گرد و غبار باعث کاهش میزان فتولیز (تجزیه شیمیایی بر اثر نیرو تابشی) و در نهایت تغییر اقلیم زمین و همچنین از بین رفتن مواد آلی و مغذی خاک و کاهش بهره وری در کشاورزی می‌گردد [۱۰]. با توجه به اثرات بسیار مخرب خشکسالی که جنبه های اقتصاد، اجتماع و محیط زیست را به طور گسترده تحت تاثیر قرار می‌دهد، لزوم توجه و پایش خشکسالی به وضوح درک می‌شود. خشکسالی بر روی کمیت و کیفیت منابع آب سطحی و زیر زمینی، پوشش گیاهی، وقوع سیل و زلزله، وقوع طوفان های شن در مناطق خشک، افزایش ریزگرد ها، بهره وری کشاورزی، امنیت تامین مواد غذایی و موارد متعدد دیگر اثرات مخربی را بر جای می‌گذارد. از طرفی با ارزیابی خشکسالی می‌توان تصمیمات موثری در زمینه مدیریت خشکسالی اتخاذ کرد.

۴-۱- اهداف تحقیق

هدف اول: ارزیابی کمی و کیفی خشکسالی

بعد از تهیه داده های بارش ایستگاه های هواشناسی در محدوده مورد مطالعه، شاخص بارش استاندارد^۱ در مقیاس فصلی و سالانه محاسبه شده است. فراوانی وقوع و درصد رخداد هر طبقه خشکسالی در طول سری زمانی مشخص شده است. وضعیت SPI در فصل های بهار، تابستان، پاییز، زمستان و مقیاس سالانه به صورت سری زمانی در دو ایستگاه نمایش داده شده است. کمترین مقدار SPI، تاریخ وقوع اوج خشکسالی و شدت خشکسالی در طول سری زمانی در دو ایستگاه مشخص گشته است.

هدف دوم: تعیین وابستگی میان SPI ها در طول سری زمانی

با هدف مشخص شدن وابستگی زمانی SPI در مقیاس فصلی و سالانه و بررسی موضوع تغییر اقلیم در طول سری زمانی، تحلیل خودهمبستگی جزئی^۲ انجام شده است.

هدف سوم: پیش بینی SPI در الگوریتم جنگل تصادفی

از شاخص های دارای همبستگی به عنوان ورودی الگوریتم جنگل تصادفی^۳ استفاده شده است و SPI پیش بینی شده است. ۷۰ درصد نمونه ها برای آموزش دیدن مدل و ۳۰ درصد نمونه ها در فرایند آزمایش استفاده شده است. پس از اجرا نمودن الگوریتم RF، نتایج به دست آمده بررسی و تحلیل شده اند. ضریب تبیین بین مقادیر مشاهداتی و پیش بینی شده SPI معین شده است و همچنین مقادیر مشاهداتی و پیش بینی شده SPI به صورت سری زمانی به تفکیک فصل نمایش داده شده اند.

^۱ Standardized Precipitation Index (SPI)

^۲ Partial Autocorrelation Function (PACF)

^۳ Random Forest (RF)

هدف چهارم: اعتبارسنجی پیش بینی SPI در الگوریتم جنگل تصادفی

دقت و اعتبار الگوریتم RF با به کارگیری پارامترهای آماری بین SPI پیش بینی شده و SPI محاسباتی تعیین شده است.

هدف پنجم: ارزیابی کیفی اثرات زیست محیطی خشکسالی

اثرات زیست محیطی خشکسالی در مناطق مورد مطالعه به صورت کیفی ذکر شده است.

۱-۵- ساختار پایان نامه

مطالب این تحقیق در پنج فصل تدوین شده است که به طور مختصر به شرح زیر بیان می گردد:

فصل اول: کلیات

در فصل اول، مقدمه، اهمیت پژوهش، اهداف و فعالیت ها بیان شده است.

فصل دوم: مروری بر تحقیقات پیشین

در فصل دوم، پیشینه و سابقه تحقیقات صورت گرفته شرح داده شده است و جمع بندی آن بیان شده است.

فصل سوم: مواد و روش ها

در فصل سوم، محدوده مورد مطالعه، داده های مورد استفاده، SPI، PACF، RF و پارامترهای آماری اعتبارسنجی و نیز مراحل انجام تحقیق شرح داده شده است.

فصل چهارم: نتایج و بحث

در فصل چهارم، نتایج محاسبه SPI در مقیاس فصلی و سالانه، نتایج PACF و پیش بینی SPI در الگوریتم RF نمایش داده شده است و به کمک پارامترهای آماری، اعتبار پیش بینی SPI با الگوریتم RF سنجیده شده است و همچنین پیامدهای زیست محیطی پدیده خشکسالی به صورت کیفی در محدوده مطالعاتی بررسی شده است.

فصل پنجم: نتیجه‌گیری و پیشنهادات

در فصل پنجم، نتایج به دست آمده از تحقیق بیان شده است و نتایج با تحقیقات دیگر مقایسه شده است و نیز پیشنهاداتی جهت پژوهش‌های آتی ارائه شده است.

فصل دوم: مروری بر مطالعات پیشین

سازمان جهانی هواشناسی^۱ خشکسالی را کمبود گسترده و تداوم دار بارش توصیف کرده است. دایره المعارف اقلیم و وضع هوا^۲ خشکسالی را کمبود بارش نسبت به متوسط چند ساله منطقه در یک دوره طولانی بیان کرده است. کنگره مقابله با خشکسالی و بیابان زایی ملل متحد^۳ بیان می کند که خشکسالی زمانی وقوع می یابد که بارش به طور معنی داری زیر سطح نرمال نزول یابد و باعث برهم زدن تعادل آبشناسی منابع منطقه گردد[۳]. در مورد انواع خشکسالی می توان به خشکسالی هواشناسی^۴، کشاورزی^۵، هیدرولوژیکی^۶ و اجتماعی - اقتصادی^۷ اشاره نمود[۱۱]. شاخص های خشکسالی^۸ گسترش فضایی و شدت و مدت خشکسالی را مشخص می کنند. شاخص های متنوعی برای ارزیابی خشکسالی توسعه داده شده است که متغیرهای اقلیمی و هیدرولوژیکی در محاسبه شاخص های خشکسالی رایج تر می باشد[۱۴]. به علت ماهیت پیچیده پدیده خشکسالی، شاخص های خشکسالی با توجه به منطقه و شرایط اقلیمی تعیین می شوند[۱۵]. بارندگی، دمای هوا، رطوبت خاک، تبخیر- تعرق و جریان در بیشتر روش ها ملاک تعیین شاخص های خشکسالی می باشد[۱۶]. شاخص های خشکسالی در سه نوع شاخص های خشکسالی هواشناسی، شاخص های خشکسالی کشاورزی و شاخص های خشکسالی هیدرولوژیکی تقسیم می شوند[۳].

¹ World Meteorological Organization

² Schneider

³ UN Convention to Combat Drought and

⁴ Meteorological Drought

⁵ Hydrologic Drought

⁶ Agricultural Drought

⁷ Socio-economic Drought

⁸ Drought Indices

۲-۱-۱- شاخص های خشکسالی هواشناسی

این شاخص ها شامل شاخص نمره Z (ZSI)^۱، شاخص درصد نرمال بارش (PNPI)^۲، شاخص دهک های بارش (PDI)^۳، شاخص بی هنجاری بارش (RAI)^۴، شاخص خشکسالی بالم و مولی (BMDI)^۵، شاخص بارش استاندارد (SPI)، شاخص بارش استاندارد اصلاح شده (SPImod)^۶، شاخص خشکسالی موثر (EDI)^۷، شاخص Z چینی (CZI)^۸، شاخص Z چینی تعدیل شده (MCZI)^۹ و شاخص بارش استاندارد چند متغیره (MSPI)^{۱۰} می باشد [۳].

۲-۱-۲- شاخص های خشکسالی کشاورزی

این شاخص ها شامل شاخص شاخص شدت خشکسالی پالمر (PDSI)^{۱۱}، شاخص شناسایی خشکسالی (RDI)^{۱۲}، شاخص خشکسالی خاص محصول (CSDI)^{۱۳}، شاخص باران موثر (ERI)^{۱۴} و شاخص تبخیر تعرق بارش استاندارد (SPEI)^{۱۵} می شود [۳].

۲-۱-۳- شاخص های خشکسالی هیدرولوژیکی

این شاخص ها شامل شاخص تامین آب سطحی (SWSI)^{۱۶}، شاخص خشکسالی جامع (ADI)^{۱۷} و شاخص کمبود توام (JDI)^{۱۸} می باشد [۳]. هر شاخص روش و متغیرهای ورودی مخصوصی را ملاک قرار می دهد. شاخص بارش استاندارد (SPI) توسط سازمان جهانی هواشناسی توصیه می شود و بیشترین کاربرد را دارد [۱۷]. مهم ترین مزیت SPI فراهم کردن امکان مقایسه خشکسالی به صورت منطقه ای می باشد [۱۸]. با توجه به موارد ذکر شده، در مطالعه حاضر SPI مد نظر قرار گرفته است.

¹ Z-score

² Percent of Normal Precipitation Index

³ Precipitation Deciles Index

⁴ Rainfall Anomaly Index

⁵ Bahlme and Mooley Drought Index

⁶ Modified Standardized Precipitation Index

⁷ Effective Drought Index

⁸ China-Z Index

⁹ Modified China-Z Index

¹⁰ Multivariate Standardized Precipitation Index

¹¹ Palmer Drought Severity Index

¹² Reconnaissance Drought Index

¹³ Crop-Specific Drought Index

¹⁴ Effective Rainfall Index

¹⁵ Standardized Precipitation Evapotranspiration Index

¹⁶ Surface Water Supply Index

¹⁷ Aggregate Drought Index

¹⁸ Joint Deficit Index

الگوریتم جنگل تصادفی یکی از روش های یادگیری ماشین می باشد. در این بخش نمونه هایی از کاربرد الگوریتم جنگل تصادفی، کاربرد شاخص های مختلف خشکسالی و روش های یادگیری ماشین در ارزیابی خشکسالی بیان شده است.

۲-۲- کاربرد الگوریتم جنگل تصادفی در علوم مختلف

Kulkarni و Lowe در سال ۲۰۱۵ در زمینه تجزیه و تحلیل تصاویر ماهواره ای چند طیفی به بررسی کارایی الگوریتم جنگل تصادفی و روش هایی مانند حداکثر احتمال، شبکه عصبی و ماشین بردار پشتیبانی پرداختند که بر اساس نتایج ایشان الگوریتم جنگل تصادفی از نظر دقت کلی و ضریب کاپا نسبت به سایر روش های ذکر شده مناسب تر بوده است [۱۹]. Belgiu و همکارش در سال ۲۰۱۶ مزایای الگوریتم جنگل تصادفی را در علم سنجش از دور و به منظور اندازه گیری اهمیت متغیر ها، کاهش ابعاد داده های فراطیفی، شناسایی مرتبط ترین داده های سنجش از دور و جغرافیایی چند منظوره و همچنین انتخاب مناسب ترین فصل برای طبقه بندی کلاس های هدف خاص را شرح دادند [۲۰]. Everingham و همکاران در سال ۲۰۱۶ در مورد اندازه محصول نیشکر در منطقه Tully، واقع در شمال شرقی استرالیا، الگوریتم جنگل تصادفی را به کار بستند. آن ها عواملی مانند شاخص های پیش بینی آب و هوای فصلی، میزان بارندگی حداکثر و حداقل دما تابش را به عنوان ورودی الگوریتم جنگل تصادفی تنظیم نمودند و مدلی جهت توضیح تنوع سالانه در مورد محصول نیشکر ارائه کردند. لازم به ذکر است وجود چنین پیش بینی هایی مدیران کارخانه ها را در امر برنامه ریزی صحیح و فعالیت بهتر بازاریابان کمک می نماید [۲۱].

۲-۳- کاربرد الگوریتم جنگل تصادفی در محیط زیست

Melesse و همکاران در سال ۲۰۱۶ با هدف نقشه برداری پتانسیل آب زیرزمینی در منطقه مهران، الگوریتم جنگل تصادفی و حداکثر انتروپی (ME)^۱ را به کار بستند. آن ها عواملی مانند ارتفاع، درصد شیب، جهت شیب، انحنای طرح، تراکم زهکشی، فاصله از رودخانه، شاخص رطوبت تپوگرافی (TWI)^۲، کاربری زمین، سنگ شناسی و بافت خاک را ورودی روش ها تنظیم نمودند و از روش های ذکر شده جهت تولید نقشه های پتانسیل آب زیرزمینی استفاده کردند. نتایج اعتبار سنجی نشان داد که نرخ قابل قبول مساحت زیر منحنی تشخیص عملکرد نسبی برای الگوریتم جنگل تصادفی و حداکثر انتروپی به ترتیب ۸۶/۵ درصد و ۹۱ درصد می باشد. مساحت زیر منحنی تشخیص عملکرد نسبی (ROC)^۳ در مورد میزان پیش بینی الگوریتم جنگل تصادفی و حداکثر انتروپی به ترتیب ۸۳/۱ درصد و ۸۷/۷ درصد تعیین شد که نشان دهنده موثر بودن دو روش فوق در نقشه برداری پتانسیل آب زیرزمینی می باشد [۲۲]. Kulkarni و Lowe در سال ۲۰۱۶ با هدف تجزیه و تحلیل پوشش زمین به وسیله سنجش از دور، قابلیت الگوریتم جنگل تصادفی را با روش هایی همانند حداکثر احتمال، حداقل فاصله، درخت تصمیم، شبکه عصبی و ماشین بردار پشتیبانی مقایسه کردند که از نظر دقت کلی و ضریب کاپا کارایی الگوریتم جنگل تصادفی نسبت به سایر روش ها بیشتر بوده است [۲۳]. Hong و همکاران در سال ۲۰۱۷ از سه روش درخت مدل لجستیک (LMT)^۴، جنگل تصادفی، درختان رگرسیون و طبقه بند (CART)^۵ برای پیش بین مکانی آسیب پذیری زمین لغزش در منطقه لانگ بهره جستند و در نهایت به این نتیجه رسیدند که هر سه مدل عملکرد نسبتاً خوبی دارند اما الگوریتم جنگل تصادفی با درصد موفقیت ۰/۸۳۷ و نرخ پیش بینی ۰/۷۸۱، روشی مناسب برای ارزیابی آسیب پذیری زمین لغزش می باشد [۲۴].

¹ Maximum Entropy

² Topographic Wetness Index

³ Receiver Operating Characteristic

⁴ Logistic Model Tree

⁵ Classification And Regression Tree

Sylvester و همکارش در سال ۲۰۱۸ به منظور تخصیص جمعیت ژنتیکی و اهمیت آن در مدیریت و حفاظت از حیات وحش، الگوریتم جنگل تصادفی را به کار گرفتند. تخصیص جمعیت ژنتیکی به پانل هایی از مارکر های ژنتیکی با اطلاعات بالا و تست های حساس تخصیصی نیاز دارد اما آن ها با به کارگیری الگوریتم جنگل تصادفی با دقت بالا و به سادگی تخصیص جمعیت ژنتیکی را انجام دادند [۲۵]. Ilia و همکاران در سال ۲۰۱۸ پدیده فرونشست زمین را در منطقه غربی تسالالی، کشور یونان به کمک روش های سنجش از دور و روش الگوریتم جنگل تصادفی انجام دادند. با استفاده از الگوریتم جنگل تصادفی نرخ تغییر شکل زمین را بر اساس سه پارامتر اصلی تخمین زدند و به این نتیجه رسیدند که همبستگی قوی ای بین تغییر شکل زمین و ضخامت رسوبات سست وجود دارد. ضریب همبستگی در اعتبار سنجی الگوریتم جنگل تصادفی برابر 0.7503 شد که دقت بالایی می باشد [۲۶]. Norouzi و همکاران در سال ۲۰۲۰ برای ارزیابی کیفیت آب آبخوان دشت میاندوآب در اکتبر ۲۰۱۸ تعداد ۷۵ نمونه برداشت کردند و تحلیل شیمیایی لازم را برای سنجش پارامتر های کیفی آب انجام دادند. آن ها از شاخص کیفیت آب زیرزمینی و الگوریتم جنگل تصادفی و نیز منطق فازی استفاده نمودند و از دقت پیش بینی زیاد الگوریتم جنگل تصادفی نسبت به منطق فازی و توانایی تحلیل روابط غیر خطی آن جهت پهنه بندی کیفی منابع آب زیرزمینی استفاده نمودند [۲۷]. Saha و همکاران در سال ۲۰۲۰ در محدوده رودخانه گومانی در کشور هند احتمال جنگل زدایی را با روش های مختلفی از جمله الگوریتم جنگل تصادفی بررسی نمودند و نتیجه کارشان بیان گر این بود که ۳۰ درصد اراضی جنگلی خطر بالا تخریب جنگل را دارد [۲۸].

۲-۴- کاربرد شاخص های خشکسالی

Moradi و همکاران در سال ۲۰۱۱ در استان فارس در مورد ۲۶ ایستگاه هواشناسی در بازه زمانی ۱۹۶۸-۱۹۹۹ و در مقیاس زمانی ۶،۹،۱۲،۱۸،۲۴ و ۳ ماهه شاخص را به کار بستند. آن ها برای ایجاد احتمال انتقال و ماتریس تداوم تئوری RUN و زنجیره مارکوف را با هم ترکیب نمودند. شرایط خشکسالی بلند مدت از نظر تعداد و مدت زمان برای ۱۰ سال آینده پیش بینی شد. با استفاده از سیستم اطلاعات جغرافیایی (GIS)^۱ و روش های آماری جغرافیایی ویژگی های مدت خشکسالی تعیین شد. بر طبق نتایج این تحقیق خشکسالی های جنوب استان فارس شدت و مدت زمان زیادتری داراست [۲۹]. Kousari و همکاران در سال ۲۰۱۴ در کشور ایران از RDI در مقیاس ۶،۹،۱۲،۱۸،۲۴ و ۳ ماهه برای ارزیابی شدت خشکسالی استفاده کردند. در این تحقیق آزمون های غیر پارامتری Mann-Kendall و Sen's slope به کار گرفته شد. بر اساس نتایج کار آن ها سری زمانی RDI روند کاهشی داشته است که خود بیانگر روند افزایشی شدت خشکسالی می باشد [۳۰]. Dashtpajardi و همکاران در سال ۲۰۱۵، ۲۵ ایستگاه هواشناسی سینوپتیک واقع در مناطق خشک و نیمه خشک ایران را برای محاسبه SPI در مقیاس زمانی ۶،۹،۱۲،۱۸،۲۴ و ۳ ماهه و بازه سال های ۱۹۷۵-۲۰۰۵ در نظر گرفتند. ایشان در مطالعه خود علاوه بر شدت خشکسالی میزان خشکسالی (DM)^۲ را نیز مورد بررسی قرار دادند. برای تشخیص روند مقادیر SPI از آزمون غیرپارامتری Mann-Kendall استفاده شد. طبق نتایج ایشان روند SPI در مناطق خشک و نیمه خشک ایران منفی بود که خود حاکی از افزایش شدت خشکسالی در این مناطق می باشد و روند میزان خشکسالی (DM) در سه دهه مورد بررسی افزایش یافته است [۶]. Huang و همکاران در سال ۲۰۱۵ برای درک بهتر خشکسالی در حوضه رودخانه زرد شروع و پایان خشکسالی و دوره های انتقال آن و مدت زمان خشکسالی را بررسی نمودند. ایشان با ترکیب اطلاعات بارش و جریان شاخص خشکسالی استاندارد چند متغیره غیر پارامتری (NMSDI)^۳ را معرفی کردند. همبستگی بین نوسان ال نینو جنوب (ENSO)^۴ و تغییرات به کمک روش موجک انجام شد.

¹ Geographic Information System

² Drought Magnitude

³ Nonparametric Multivariate Standardized Drought Index

⁴ El Niño-Southern Oscillation

طبق نتایج آن ها تغییرات NMSDI با SPI سازگار می باشد و NMSDI در توصیف خشکسالی موثر می باشد. متوسط مدت طولانی خشکسالی در کل حوضه نزدیک به ۵/۸ ماه بود که از میانگین دوره خشکسالی و دوره های خاتمه آن طولانی تر بود [۳۱]. Pramudya و همکاران در سال ۲۰۱۸، SPI را به صورت ۱۳،۶،۹،۱۲ ماهه در شهر تگال، واقع در کشور اندونزی و در بازه زمانی ۱۹۸۳-۲۰۱۵ محاسبه کردند و شدیدترین خشکسالی ها را معین نمودند. نتایج بیانگر پتانسیل بالای SPI در تعیین مدت و شدت خشکسالی بود [۳۲]. Sobral و همکاران در سال ۲۰۱۹ برای توصیف خشکسالی ایالت ریودوژانیرو از SPI سالانه در بازه زمانی ۱۹۷۹-۲۰۰۹ استفاده نمودند. آن ها از آزمون هایی مانند curvature sen, pettitt, mann-kendall برای ارزیابی شاخص سالانه SPI بهره گرفتند. روند های کاهشی و افزایشی SPI سالانه توسط آزمون mann-kendall شناسایی شد و طبق آزمون pettitt شروع افزایش روند سالانه در دو دوره متمایز رخ داد [۱۷]. Kamruzzaman و همکاران در سال ۲۰۱۹ برای بررسی روند خشکسالی، SPI و شاخص خشکسالی موثر (EDI) را در کشور بنگلادش استفاده کردند، آن ها داده های مورد نیاز را از ۲۷ ایستگاه هواشناسی در سراسر کشور برای مدت زمان ۲۰۱۵-۱۹۸۱ تهیه کردند و به این نتیجه رسیدند که شاخص EDI در ارزیابی کوتاه مدت و بلند مدت خشکسالی عملکرد بهتری نسبت به SPI دارد [۳۳]. Lingcheng و همکاران در سال ۲۰۲۰ به منظور پایش روند خشکسالی در کشور چین، SPI و SPEI را برای بازه زمانی ۱۹۸۰-۲۰۱۵ در نظر گرفتند و از الگوریتم خوشه بندی سه بعدی و نیز روش توزیعی شدت - منطقه به منظور بررسی پویایی زمانی و مکانی خشکسالی استفاده نمودند. نتایج نشان داد که رویداد های خشکسالی شناسایی شده با SPEI شدیدتر از SPI می باشد [۳۴].

۲-۵- پیش بینی شاخص های مختلف خشکسالی با روش های یادگیری ماشین

Chen و همکاران در سال ۲۰۱۲ برای برآورد SPI از دو روش الگوریتم جنگل تصادفی و ARIMA^۱ استفاده نمودند و به این نتیجه رسیدند که دقت الگوریتم جنگل تصادفی از روش ARIMA در پیش بینی خشکسالی بالاتر می باشد [۳۵]. Im و همکاران در سال ۲۰۱۶ شانزده فاکتور خشکسالی مبتنی بر سنجش از دور از طیف سنج تصویری سنجنده مودیس (MODIS) را برای نظارت بر خشکسالی هواشناسی و کشاورزی در طول بازه زمانی ۲۰۰۰-۲۰۱۲ در ایالات متحده آمریکا در نظر گرفتند. آن ها SPI را مورد استفاده قرار دادند. رابطه بین شانزده فاکتور خشکسالی از طریق سه روش یادگیری ماشین شامل جنگل تصادفی، درختان رگرسیون تقویت شده و Cbiust بررسی شد که نتایج نشان داد که جنگل های تصادفی دارای بهترین عملکرد (ضریب تعیین ۰/۹۳ و میانگین مربعات ریشه ۰/۳) برای پیش بینی SPI در بین سه روش ذکر شده است [۳۶]. Im و همکاران در سال ۲۰۱۷ بر مبنای روش های ماشین لرنینگ از جمله جنگل تصادفی و نیز داده های پیش بینی آب و هوا دور برد و سنجش از دور به منظور پیش بینی نمودن خشکسالی به کمک SPI و SPEI در مقیاس ۳،۶،۹،۱۲ ماهه استفاده کردند و در نهایت به این نتیجه رسیدند که سهم داده های پیش بینی آب و هوا دور برد زیاد نیست و درختان تصادفی از نظر دقت، عملکرد مطلوب تا ۶۴ درصد را نشان دادند [۳۷]. Park و همکاران در سال ۲۰۱۹ با به کارگیری داده های توپوگرافی و تصاویر ماهواره ای، مدل پیش بینی ناحیه شدید خشکسالی را بدون در نظر گیری داده های هواشناسی ارائه کردند و نیز الگوریتم جنگل تصادفی را هم به کار گرفتند و به دقت بالا مدل دست یافتند [۳۸]. Dikshit و همکاران در سال ۲۰۲۰ به منظور تعیین روند خشکسالی در منطقه ولز، کشور استرالیا بازه زمانی ۱۹۰۱-۲۰۱۸ و SPEI را در نظر گرفتند. آن ها روش جنگل تصادفی را بر روی داده ها پیاده نمودند و در مرحله اعتبار سنجی به ضریب همبستگی بالا و میانگین مربعات ریشه مطلوب دست یافتند و به این نتیجه رسیدند که الگوریتم جنگل تصادفی عملکرد بالایی را داراست [۳۹]. Mehr و همکاران در سال ۲۰۲۰ مدل جنگل تصادفی

¹ Autoregressive Integrated Moving Average

فازی را در مورد SPEI در پیش بینی پدیده خشکسالی به کار بستند و به دقت بالایی رسیدند [۴۰]. C. Deo و Tien Bui و همکاران در سال ۲۰۲۰ در جنوب شرقی کوئینزلند استرالیا، که در معرض خشکسالی است و طی دوره ۱۹۹۴-۲۰۱۳ و با به کارگیری الگوریتم جنگل تصادفی (RF)، ماشین بردار پشتیبانی (SVM)^۱، درختان رگرسیون تقویت شده (BRT)^۲، درختان طبقه بندی و رگرسیون (CART)، خطوط رگرسیونی تطبیقی چند متغیره (MARS)^۳ و تجزیه و تحلیل تفکیک کننده انعطاف پذیر (FDA)^۴ اقدام به پیش بینی خشکسالی نمودند و طبق نتیجه اعتبار سنجی الگوریتم جنگل تصادفی بالاترین میزان دقت را دارا بود [۴۱]. Khan و همکاران در سال ۲۰۲۰ خشکسالی کشور پاکستان را با استفاده از سه روش یادگیری ماشین یعنی ماشین بردار پشتیبانی (SVM)، شبکه عصبی مصنوعی (ANN)^۵ و نزدیک ترین همسایه (KNN)^۶ و با استفاده از SPEI پیش بینی نمودند که طبق مطالعه آن ها روش KNN نسبت به روش های SVM و ANN عملکرد پایین تری را داشته است و مدل های مبتنی بر روش SVM ویژگی های مکانی و زمانی خشکسالی را بهتر ارائه می دهند [۴۲]. Shamshirband و همکاران در سال ۲۰۲۰ سه شاخص SPI، SPEI و شاخص جریان استاندارد را به کمک روش ماشین بردار رگرسیون (SVR)^۷، برنامه نویسی بیان ژن (GEP)^۸ و درختان مدل (MT)^۹ مدل سازی کردند و نتیجه کارشان دقیق بودن SPI بود [۴۳]. Dehe Xu و همکاران در سال ۲۰۲۰، SPI را در مقیاس زمانی ۱،۳،۶،۱۲ ماهه به کمک مدل ARIMA و مدل ARIMA-SVR و با یک دیگر مقایسه نمودند و در نهایت به این نتیجه رسیدند که در اعتبار سنجی مدل ها بین داده های مشاهداتی و پیش بینی شده دقت پیش بینی مدل ARIMA-SVR از مدل ARIMA بالاتر می باشد [۴۴].

¹ Support Vector Machine

² Boosted Regression Trees

³ Multivariate Adaptive Regression Splines

⁴ Flexible Discriminant Analysis

⁵ Artificial Neural Network

⁶ k-Nearest Neighbour

⁷ Support Vector Regression

⁸ Gene Expression Programming

⁹ Model Trees

۲-۶- جمع بندی مرور بر مطالعات پیشین

طبق تحقیقات صورت گرفته چنین نتیجه می‌شود که استفاده از شاخص های متنوع موجود در زمینه پایش خشکسالی مطلوب می‌باشد و همچنین کاربرد روش های مختلف یادگیری ماشین همانند ماشین بردار پشتیبانی، درختان رگرسیون تقویت شده، درختان طبقه بندی و رگرسیون، خطوط رگرسیونی تطبیقی چند متغیره، تجزیه و تحلیل تفکیک کننده انعطاف پذیر، شبکه عصبی مصنوعی، برنامه نویسی بیان ژن و درختان مدل به همراه شاخص های خشکسالی به منظور ارزیابی خشکسالی کارا می‌باشد و دقت لازم را فراهم می‌سازد. در تحقیق حاضر ارزیابی خشکسالی هواشناسی مورد نظر است لذا از SPI که در دسته شاخص های خشکسالی هواشناسی قرار دارد، استفاده شده است. همچنین SPI توسط سازمان جهانی هواشناسی توصیه می‌شود و بیشترین کاربرد را دارد و مبنای تصمیم گیری و تخصیص منابع آب بر اساس شاخص خشکسالی هواشناسی SPI می‌باشد [۱۷]. به کارگیری شاخص SPI به همراه الگوریتم RF در زمینه پایش خشکسالی بسیار مفید واقع شده است لذا در این مطالعه از شاخص SPI و الگوریتم RF استفاده شده است. با توجه به تعیین وابستگی میان SPI ها در طول سری زمانی و بررسی موضوع تغییر اقلیم، از مقیاس فصلی و سالانه SPI استفاده شده است. در مقالات پیشین، کم تر به بحث ارزیابی شدت خشکسالی پرداخته شده است لذا در این تحقیق شدت خشکسالی در طول سری زمانی محاسبه شده است و همچنین انجام تحلیل خودهمبستگی جزئی (PACF) پیش از الگوریتم جنگل تصادفی به منظور بهبود پیش بینی SPI در نظر گرفته شده است که جنبه جدید تحقیق حاضر می‌باشد.

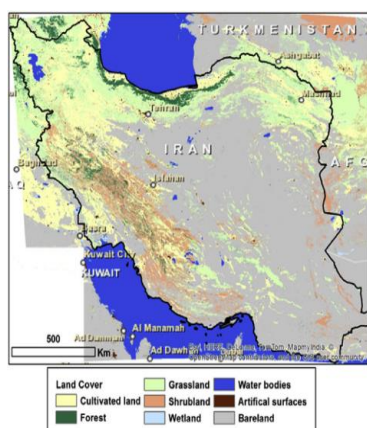
فصل سوم: مواد و روش‌ها

۳-۱- مقدمه

در این فصل محدوده مورد مطالعه، ویژگی ایستگاه های مورد مطالعه، داده های مورد استفاده، مراحل آزمون و بررسی داده ها، شاخص بارش استاندارد، تابع خودهمبستگی جزئی، روش های یادگیری ماشین، روش بگینگ، الگوریتم جنگل تصادفی، پارامتر های آماری و مراحل انجام تحقیق شرح داده شده است.

۳-۲- محدوده مورد مطالعه

کشور ایران با مساحتی حدود ۱,۶۴۸,۱۹۵ کیلومتر مربع در جنوب غربی قاره آسیا و بین عرض های جغرافیایی $25^{\circ}00'N$ و $38^{\circ}39'N$ و طول جغرافیایی $44^{\circ}00'E$ و $63^{\circ}25'E$ واقع شده است [۳۰]. به دلیل ویژگی های متنوع جغرافیایی و توپوگرافی، ایران اقلیم های مختلفی دارد. رشته کوه البرز در قسمت شمالی و زاگرس در قسمت غربی، مانع ورود رطوبت به بخش های مختلف می باشد. میانگین بارش سالانه در ایران حدود ۲۵۱ میلی متر یا ۴۱۳ میلیارد مترمکعب است که از کم تر از یک سوم میانگین بارش جهانی (۸۳۱ میلی متر) می باشد. تنوع مکانی و زمانی بارش در ایران بالا می باشد. بخش هایی از جنوب دریای خزر، حداکثر ۲۰۰۰ میلی متر بارندگی سالانه دارند در صورتی که بارندگی در بخش های مرکزی و شرقی کشور، کم تر از ۵۰ میلی متر در سال می باشد. حدود ۹۰ درصد مناطق کشور به لحاظ اقلیمی خشک یا نیمه خشک هستند [۶]. در شکل ۳-۱ نقشه عوارض طبیعی ایران قابل مشاهده است [۴۵].



شکل ۳-۱ نقشه عوارض طبیعی کشور ایران [۴۵].

محدوده مطالعاتی ایستگاه سینوپتیک شهرهای زابل و زاهدان، واقع در استان سیستان و بلوچستان می‌باشد.

وسعت استان سیستان و بلوچستان ۱۸۱۴۷۱ کیلومتر مربع می‌باشد. محدوده مطالعاتی بین عرض های جغرافیایی $25^{\circ}25'N$ و $27^{\circ}31'N$ و طول جغرافیایی $50^{\circ}58'E$ و $63^{\circ}21'E$ واقع شده است. این استان در جنوب با دریای عمان ۳۰۰ کیلومتر مرز آبی داراست و در شرق با کشور های افغانستان و پاکستان ۱۸۰۰ مرز زمینی دارد [۸]. این محدوده دارای آب و هوای کویری و نیمه بیابانی است و متوسط بارندگی سالانه آن حدود ۱۱۰ میلی متر و متوسط دمای آن ۲۲ تا ۳۷ درجه سانتی گراد است. رطوبت نسبی کم، کم بودن بارندگی سالانه نسبت به بارندگی متوسط کشور و تغییرات دمایی بالا بخشی از ویژگی های این ناحیه می‌باشد [۷]. در دهه گذشته دریاچه هامون که از رودخانه هیرمند ارتفاعات کشور افغانستان سرچشمه می‌گیرد به دلیل سدسازی های این کشور و کاهش بارندگی خشک شده است و موجب به خطر افتادن زیستگاه گیاهان و جانوران شده است. یکی دیگر از ویژگی های مهم این منطقه، باد های ۱۲۰ روزه است که معمولا از اواخر بهار تا اوایل پاییز ادامه دارد. سرعت غالب باد های ۱۲۰ روزه نزدیک به ۱۰ متر بر ثانیه می‌باشد که موجب افزایش نرخ تبخیر، افزایش فرسایش خاک، کاهش پوشش گیاهی، طوفان های گرد و غبار و مشکلات تنفسی برای افراد می‌گردد [۸]. طوفان های گرد و غبار باعث تغییر در شرایط اقلیمی، چرخه بیولوژیکی، زمین شناسی و محیط زیست انسان می‌گردد. آئروسل های معدنی ایجاد شده از طوفان های گرد و غبار بر روی تشکیل ابر و کمیت نزولات جوی اثر می‌گذارد [۹]. طوفان های گرد و غبار باعث کاهش میزان فتولیز (تجزیه شیمیایی بر اثر نیرو تابشی) و در نهایت تغییر اقلیم زمین و همچنین از بین رفتن مواد آلی و مغذی خاک و کاهش بهره وری در کشاورزی می‌گردد [۱۰]. در شکل ۲-۳ نقشه استان سیستان و بلوچستان و محدوده مطالعاتی زابل و زاهدان قابل ملاحظه است.



شکل ۳-۲ نقشه استان سیستان و بلوچستان و محدوده مطالعاتی زابل و زاهدان

۳-۳- ویژگی ایستگاه های مورد مطالعه

محدوده مطالعاتی ایستگاه سینوپتیک شهرهای زابل و زاهدان، واقع در استان سیستان و بلوچستان می باشد. اطلاعات مربوط به ایستگاه های مورد مطالعه شامل طول جغرافیایی، عرض جغرافیایی، ارتفاع از سطح دریا، متوسط بارش سالانه و اقلیم در جدول ۳-۱ بیان شده است [۶].

جدول ۳-۱ مشخصات ایستگاه زابل و زاهدان [۶].

نام ایستگاه	طول جغرافیایی درجه	عرض جغرافیایی درجه	ارتفاع از سطح دریا بر حسب متر	متوسط بارش سالانه بر حسب میلی متر	اقلیم
زابل	61.48	31.03	489.2	62.2	بسیار خشک
زاهدان	60.88	29.47	1370	75.3	خشک

۳-۴- داده های مورد استفاده

داده های بارش ماهانه ایستگاه های هواشناسی شهرستان زابل و زاهدان در بازه زمانی دسامبر ۱۹۹۰ تا نوامبر ۲۰۲۰ به صورت سری زمانی از سایت هواشناسی کشور تهیه شده است. برای محاسبه پارامتر های آماری ذکر شده در جدول ۳-۲ از نرم افزار SPSS^۱ استفاده شده است.

جدول ۳-۲ پارامتر های آماری سری زمانی بارش ماهانه در ایستگاه زابل و زاهدان

نام ایستگاه	انحراف معیار استاندارد	بر حسب میلی متر	مجموع	حداکثر بارش	حداقل بارش	تعداد داده
		بر حسب میلی متر	بر حسب میلی متر	بر حسب میلی متر	بر حسب میلی متر	
زابل	9.51284	4.457	1604.51	72.2	0	360
زاهدان	12.47609	6.5312	2351.24	85.4	0	360

میانگین بارش ماهانه ایستگاه زابل و زاهدان در طول سری زمانی به ترتیب ۴/۴۶ میلی متر و ۶/۵۳ میلی متر می باشد.

۳-۵- آزمون و بازسازی داده ها

عواملی مانند تغییر محل ایستگاه های اندازه گیری، تغییر اشخاصی که کار قرائت را انجام می دهند، خطا در قرائت، منظم نبودن زمان اندازه گیری، تعویض ادوات، احداث ساختمان در اطراف ایستگاه و عوامل دیگر باعث ایجاد خطا در داده ها می شود. باید از کمیت و کیفیت داده ها و کامل بودن سری آماری اطمینان حاصل کرد و سپس به تجزیه و تحلیل پرداخت [۲].

¹ Statistical Package for Social Science

۳-۵-۱- بررسی تکمیل بودن داده ها

تکمیل بودن داده ها بررسی می گردد.

۳-۵-۲- بررسی همگن بودن داده ها

همگن بودن به معنی مشخص بودن داده های جامعه آماری می باشد. ران تست آزمونی غیرپارامتری است و توسط والد-ولفوویتس پیشنهاد شد. برای انجام روش ران تست پنج گام لازم است :

گام اول: داده ها به همان صورتی که از نظر زمانی اتفاق افتاده اند، ردیف شوند.

گام دوم: میانگین داده ها محاسبه شود.

گام سوم: داده ها با میانگین مقایسه شود. اگر داده بزرگ تر از میانگین باشد به آن علامت a و اگر داده کوچک تر از میانگین باشد به آن علامت b نسبت می دهیم. در صورتی که داده برابر با میانگین باشد از آن داده صرف نظر شود.

گام چهارم: تعداد دنباله ها یا ران های به وجود آمده از a تا b شمارش شده و تعداد کل آن ها R نامیده می شود.

گام پنجم: تعداد a ها و b ها شمارش شده و با علائم n_a و n_b مشخص شوند. در صورتی که هر کدام از n_a یا n_b بزرگ تر از ۲۰ باشند، آزمون به روش استفاده از جدول آماری Z و اگر هم n_a و هم n_b از ۲۰ کم تر باشند، از آزمون آماری جدول بحرانی U استفاده شود [۲]. برای بررسی همگن بودن داده های بارش از آزمون Run test در محیط نرم افزار SPSS استفاده شده است [۴۶]. در این آزمون فرض صفر همگن بودن داده ها و فرض یک ناهمگن بودن داده ها می باشد. اگر مقدار P -value از مقدار درجه اطمینان در نظر گرفته شده بزرگتر باشد، فرض صفر صحیح است و در غیر این صورت فرض یک مورد پذیرش می باشد و داده ها ناهمگن هستند [۴۷].

۳-۵-۳- بررسی روند داده ها یا تصادفی بودن داده ها

در صورتی که سری زمانی داده های هیدرولوژی به صورت یکنواخت سیر صعودی یا نزولی داشته باشد داده ها دارای روند می باشند [۲]. برای بررسی وجود یا عدم وجود روند در داده ها از آزمون تاوکندال در محیط نرم افزار SPSS استفاده شده است [۴۷]. در این مطالعه فرض صفر این است که مجموعه داده دارای روند نمی باشد و داده ها تصادفی هستند و فرض یک این است که مجموعه داده بررسی شده دارای روند می باشد. مقدار value-p یا همان $\text{Asymp. Sig. (2-tailed)}$ سطح معنی داری آزمون می باشد.

۳-۶-۳- شاخص بارش استاندارد (SPI)

۳-۶-۳-۱- تاریخچه SPI

McKee و همکاران در سال ۱۹۹۳، SPI را در مقیاس زمانی کوتاه مدت ۳ و ۶ ماهه برای مقاصد کشاورزی و مقیاس زمانی بلند مدت ۱۲، ۲۴ و ۴۸ ماهه برای اهداف آب شناسی پیشنهاد نمودند [۴۸]. Hayes و همکاران در سال ۱۹۹۹، عملکرد بالای SPI را در مورد زمان شروع خشکسالی و پیشرفت آن در پایش و ارزیابی خشکسالی ایالت کلرادوی آمریکا در سال ۱۹۹۶، نتیجه گرفتند [۴۹]. Cancelliere و همکاران در سال ۲۰۰۷، با فرض نرمال بودن داده های بارندگی ماهانه و به کمک توابع انتقال احتمال خشکسالی ماهانه را با به کارگیری SPI پیش بینی نمودند [۵۰]. نمونه های بیشتری از کاربرد شاخص در فصل دوم بیان شده است.

۳-۶-۳-۲- تعریف SPI

شاخص بارش استاندارد در سال ۱۹۹۳ توسط مک کی برای ارزیابی خشکسالی اقلیمی ابداع شد که داده های ماهانه بارندگی را در دوره مورد نظر ملاک قرار می دهد و توزیع آماری گاما را به عنوان تابع توزیع مناسب برای برازش روی داده های بارش استفاده می کند. این شاخص برای مقیاس های زمانی ۱، ۳، ۶، ۹، ۱۲ و ۲۴ ماهه تعریف می شود. مقدار شاخص به دست آمده برای داده های یک دوره زمانی طولانی، تابع توزیع نرمال می باشد که میانگین صفر و انحراف معیار ۱ را داراست. بر اساس شاخص بارش استاندارد زمانی که مقدار شاخص منفی گردد، خشکسالی

وقوع می‌یابد و هنگامی که مقدار شاخص مثبت گردد، خشکسالی پایان می‌پذیرد. برای محاسبه این شاخص حداقل ۳۰ سال دوره آماری لازم می‌باشد [۳۳].

۳-۶-۳ محاسبه SPI

SPI با برازش دادن یک تابع چگالی احتمال به توزیع فراوانی بارش تعیین و محاسبه می‌شود و این عمل را برای هر مقیاس زمانی دلخواه می‌توان به کار گرفت. هر تابع چگالی احتمال به یک توزیع نرمال استاندارد تبدیل می‌شود. تابع چگالی احتمال از رابطه (۱-۳) قابل محاسبه است:

$$g(x) = \frac{1}{\beta \Gamma(\alpha)} x^{\alpha-1} e^{-x/\beta} \quad x > 0 \quad (1-3)$$

که $\alpha > 0$ و $\beta > 0$ به ترتیب فاکتورهای شکل و مقیاس هستند و $x > 0$ مقدار بارش می‌باشد و $\Gamma(\alpha)$ تابع گاما می‌باشد که از رابطه (۲-۳) محاسبه می‌شود:

$$\Gamma(\alpha) = \int_0^{\infty} y^{\alpha-1} e^{-y} dy \quad (2-3)$$

Edwards and McKee بر اساس تقریب تام برای حداکثر احتمال ضریب‌های بهینه α و β را به کارگیری روابط (۳-۳) و (۴-۳) و (۵-۳) بیان کردند.

$$\alpha = \frac{1}{4A} \left(1 + \sqrt{1 + \frac{4A}{3}} \right) \quad (3-3)$$

$$\beta = \frac{\bar{x}}{\alpha} \quad (4-3)$$

$$A = \ln(\bar{x}) - \frac{1}{n} \sum_{i=1}^n \ln(x_i) \quad (5-3)$$

در رابطه (۵-۳)، n تعداد داده‌ها یا مشاهدات است.

احتمال تجمعی بارش صورت گرفته در هر مقیاس زمانی از رابطه (۶-۳) حاصل می‌گردد:

$$G(x) = \int_0^x g(x)dx = \frac{1}{\beta^\alpha \Gamma(\alpha)} \int_0^x x^{\alpha-1} e^{-x/\beta} dx \quad (۶-۳)$$

با جایگزین نمودن t به صورت $t = \frac{x}{\beta}$ تابع گاما ناقص از معادله (۷-۳) محاسبه می‌شود.

$$G(x) = \frac{1}{\Gamma(\alpha)} \int_0^x t^{\alpha-1} e^{-t} dt \quad (۷-۳)$$

به این علت که تابع گاما برای x معادل $\bullet$ تعریف نشده است و در مناطق خشک و نیمه خشک ممکن است بارشی رخ ندهد، تابع احتمال تجمعی کل که شامل مقادیر $\bullet$ بارش باشد، به صورت رابطه (۸-۳) تعریف می‌شود [۶].

$$H(x) = q + (1 - q) G(x) \quad (۸-۳)$$

q احتمال بارش معادل $\bullet$ است که از رابطه (۹-۳) محاسبه می‌شود:

$$q = \frac{m}{n} \quad (۹-۳)$$

در رابطه (۹-۳)، n تعداد کل داده‌ها و m تعداد داده‌های بارش دارای مقدار $\bullet$ در سری زمانی می‌باشد [۳۰].

تابع احتمال تجمعی به متغیر تصادفی نرمال استاندارد با میانگین $\bullet$ و واریانس ۱ تبدیل می‌شود که در واقع مقدار SPI می‌باشد. برای دستیابی راحت به مقادیر Z یا SPI از معادله پیشنهاد شده Stegun - Abramowitz به صورت روابط (۱۰-۳) و (۱۱-۳) و (۱۲-۳) و (۱۳-۳) استفاده شده است که این روابط، احتمالات تجمعی را به متغیر تصادفی نرمال استاندارد تبدیل می‌کنند [۶]. در جدول ۳-۳ حدود SPI و طبقات خشکسالی ذکر شده است [۱۲].

به ازای $0 < H(x) \leq 0.5$

$$t = \sqrt{\ln\left(\frac{1}{H(x)^2}\right)} \quad (۱۰-۳)$$

$$Z = SPI = -\left(t - \frac{c_0 + c_1 t + c_2 t^2}{1 + d_1 t + d_2 t^2 + d_3 t^3}\right) \quad (۱۱-۳)$$

و به ازای $0.5 < H(x) \leq 1$

$$t = \sqrt{\ln\left(\frac{1}{(1-H(X))^2}\right)} \quad (۱۲-۳)$$

$$Z = \text{SPI} = \left(t - \frac{c_0 + c_1 t + c_2 t^2}{1 + d_1 t + d_2 t^2 + d_3 t^3}\right) \quad (۱۳-۳)$$

جدول ۳-۳ طبقه بندی خشکسالی بر اساس SPI [۱۲].

$$c_0 = 2.515517$$

$$c_1 = 0.802853$$

$$c_2 = 0.010328$$

$$d_1 = 1.432788$$

$$d_2 = 0.189269$$

$$d_3 = 0.00130$$

مقدار SPI	طبقه یا کلاس خشکسالی
>2	ترسالی بسیار شدید (Extremely wet)
1.5 تا 1.99	ترسالی شدید (Severely wet)
1 تا 1.49	ترسالی متوسط (Moderately wet)
-0.99 تا 0.99	تقریباً نرمال (Near normal)
-1 تا -1.49	خشکسالی متوسط (Moderately dry)
-1.5 تا -1.99	خشکسالی شدید (Severely dry)
< -2	خشکسالی بسیار شدید (Extremely dry)

۳-۷-تابع خودهمبستگی جزئی (PACF)

تابع خودهمبستگی جزئی یک روش خطی می‌باشد که معمولاً در شناسایی مدل‌های کاهنده خودکار استفاده می‌شود [۵۱]. تابع خودهمبستگی جزئی، همبستگی بین Z_t و Z_{t+k} را بعد از حذف متغیرهای همبسته $Z_{t-k+1}, \dots, Z_{t-2}$ و Z_{t-1} معین می‌کند. ضریب خودهمبستگی جزئی معمولاً با Φ_{kk} نمایش داده می‌شود که در رابطه (۳-۱۴) بیان شده است.

$$\Phi_{kk} = \text{Corr}(Z_t, Z_{t-k} \mid Z_{t-1}, Z_{t-2}, \dots, Z_{t-k+1}) \quad (۳-۱۴)$$

Φ_{kk} ضریب همبستگی بین دو متغیر تصادفی Z_t و Z_{t-k} به شرط $Z_{t-k+1}, \dots, Z_{t-2}$ و Z_{t-1} می‌باشد. در محاسبه ضرایب خودهمبستگی جزئی معمولاً از معادله Yule-Walker که در رابطه (۳-۱۵) آورده شده است، استفاده می‌شود.

$$\rho_1 = \Phi_{k1} + \Phi_{k2}\rho_1 + \dots + \Phi_{kk}\rho_{k-1}$$

$$\rho_2 = \Phi_{k1}\rho_1 + \Phi_{k2} + \dots + \Phi_{kk}\rho_{k-2}$$

تا

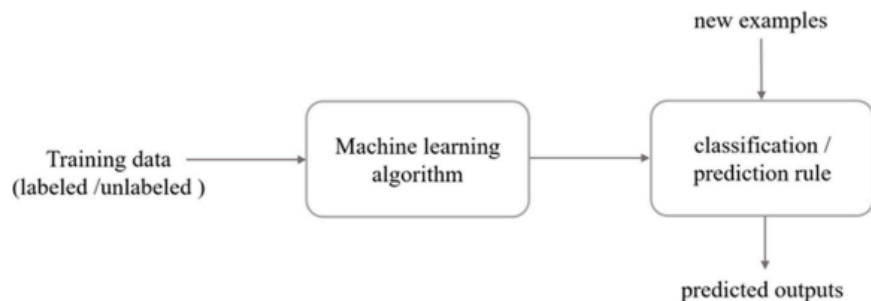
$$\rho_k = \Phi_{k1}\rho_{k-1} + \Phi_{k2}\rho_1 + \dots + \Phi_{kk} \quad (۳-۱۵)$$

ضرایب خودهمبستگی جزئی به کمک ضرایب خودهمبستگی جزئی با تغییر مقدار ρ در معادله Yule-Walker با r و شمارش $k=1,2,\dots$ می‌تواند تخمین زده شود. مقدار Φ_{kk} با به کارگیری قوانین کرامر قابل محاسبه است [۵۲].

۳-۸- روش های یادگیری ماشین

از پیشرفت های ریاضیات کاربردی، می توان به روش های یادگیری ماشین اشاره نمود. روش های یادگیری ماشین در طیف وسیعی از مطالعات پزشکی، مهندسی و محاسبات به کار می روند [۵۳]. اجرای روش های یادگیری ماشین در مقایسه با مدل های فیزیکی آسان تر بوده و امکان آموزش سریع و اعتبار سنجی و رسیدن به دقت بالا وجود دارد و همچنین می توان به وسیله آن سیستم های پیچیده و غیر خطی را با دقت بالا مدل سازی نمود [۵۴].

در روش های یادگیری ماشین، داده های ورودی آموزش داده می شود و از مدل آموزش دیده به منظور طبقه بندی، پیش بینی و یا خوشه بندی نمونه های آزمایشی استفاده می شود. بسته به نوع یادگیری ماشین، روش های یادگیری ماشین می توانند نظارت شده یا نظارت نشده باشند. در یادگیری نظارت شده هدف ساخت یک قانون کلی است که از مدل آموزش دیده به منظور پیش بینی استفاده می شود اما در یادگیری نظارت نشده تفاوتی میان مجموعه آموزش و آزمایش وجود ندارد و یادگیرنده داده های ورودی را با هدف کشف و آشکارسازی الگو های پنهان پردازش می نماید [۵۵]. برخی از مدل های یادگیری ماشین شامل شبکه های عصبی مصنوعی (ANN)، عصبی-فازی، ماشین بردار پشتیبان (SVM)، ماشین بردار رگرسیون (SVR)، دسته بندی^۱، جنگل تصادفی (RF)، رگرسیون لوجستیک و درختان تصمیم (DT)^۲ می گردد [۵۳ و ۵۴]. در شکل ۳-۳ نمایی از روش های یادگیری ماشین نمایش داده شده است [۵۵].



شکل ۳-۳ نمایی از روش یادگیری ماشین [۵۵].

¹ Clustering

² Decision trees

۳-۹- روش بگینگ

به منظور کاهش انحراف از یک تابع پیش گویی از روش بگینگ یا تجمع بوت استرپ^۱ (نمونه برداری توسط جایگذاری مجدد داده ها در مجموعه اصلی) می توان استفاده کرد. در روش بگینگ همه ی متغیر های موجود در فرایند تقسیم گره بررسی می شوند [۵۶].

۳-۱۰- الگوریتم جنگل تصادفی (RF)

۳-۱۰-۱- تاریخچه الگوریتم RF

الگوریتم جنگل تصادفی نخستین بار توسط Breiman در سال ۲۰۰۱ پیشنهاد شد [۵۷]. الگوریتم جنگل تصادفی عدم تشابه موجود در مشاهدات را اندازه گیری می کند و همچنین می تواند داده های مشاهده شده را از داده هایی که مصنوعی ساخته شدند، تشخیص دهد [۵۸]. نمونه های بیشتری از کاربرد الگوریتم جنگل تصادفی در فصل دوم بیان شده است.

۳-۱۰-۲- تعریف الگوریتم RF

الگوریتم جنگل تصادفی برای بهبود پیش بینی چندین درخت تصمیم تصادفی را ترکیب می کند و میانگین نتیجه هر درخت را به عنوان نتیجه نهایی در نظر می گیرد. الگوریتم جنگل تصادفی در رگرسیون و طبقه بندی قابل به کارگیری است [۵۷].

¹ Bootstrap

۳-۱۰-۳- پارامتر های الگوریتم RF

الگوریتم جنگل تصادفی چندین پارامتر مهم دارد که برای رسیدن به پاسخ بهینه قابل تنظیم هستند. که مهم ترین این پارامتر ها شامل تعداد متغیر های تصادفی در هر گره، تعداد درختان جنگل و می باشد :

۳-۱۰-۳-۱- تعداد متغیر های تصادفی در هر گره (m_{try})

در هر گره از میان کل متغیر های موجود (p) متغیر های تصادفی انتخاب می شود. به طور پیش فرض مقدار این پارامتر در نرم افزار ها در فرایند رگرسیون $\frac{p}{3}$ و در فرایند طبقه بندی $\sqrt{p}$ می باشد اما امکان تغییر آن وجود دارد [۲۶].

۳-۱۰-۳-۲- تعداد درختان جنگل (n_{tree})

تعداد درختان جنگل باید در حد کافی زیاد باشد تا خطای OOB پایدار گردد [۲۷].

۳-۱۰-۴- نحوه اجرا RF

حدود $\frac{2}{3}$ داده ها در فرایند آموزش و $\frac{1}{3}$ داده ها در فرایند تست استفاده می گردد [۵۶]. به کمک زیر مجموعه ای تصادفی از متغیر ها یا فاکتور های هر گره، تقسیم هر درخت انجام می گردد [۵۹]. در الگوریتم جنگل تصادفی هر درخت به کمک نمونه بوت استرپ ایجاد می شود [۵۷]. به کمک روش بوت استرپ از داده های آموزش زیر مجموعه ای N عضوی از داده ها، برای اولین تا آخرین درخت انتخاب می شود. از p متغیر کل، تعداد m متغیر به طور تصادفی انتخاب می شود و سپس از میان m متغیر تصادفی انتخاب شده، بهترین متغیر جهت شکافتگی گره، برگزیده می شود و گره والد یا اصلی بر اساس آن متغیر، به دو گره دختر شکافته می شود. هر درخت از روی داده های بوت استرپ و با تکرار مراحل فوق ساخته می شود طوری که هر گره انتهایی درخت حداقل مقدار را دارا باشد [۵۶].

۳-۱۰-۵- قابلیت های RF

۳-۱۰-۵-۱- ارزیابی الگوریتم به کمک نمونه های خارج از کیسه (OOB)^۱

از مجموعه داده های تست به کمک نمونه های خارج از کیسه (OOB) برای ارزیابی پیش بینی استفاده می شود.

مقدار $\hat{Y}^{OOB}(x_i)$ ، مقدار پیش بینی شده برای داده x_i است که از میانگین گیری نتیجه هر درخت به کمک نمونه

های OOB مربوط به آن داده به دست می آید. خطای هر درخت و خطای کل مجموعه به کمک داده های OOB

قابل محاسبه است. میزان خطای مربع میانگین^۲ در روش رگرسیون و نرخ خطا^۳ در روش طبقه بندی در الگوریتم

جنگل تصادفی به شرح روابط (۳-۱۶) و (۳-۱۷) است:

$$MSE \approx MSE^{OOB} = n^{-1} \sum_{i=1}^n \{ \hat{Y}^{OOB}(x_i) - Y_i \}^2 \quad (۳-۱۶)$$

$$ER \approx ER^{OOB} = n^{-1} \sum_{i=1}^n I(\hat{Y}^{OOB}(x_i) \neq Y_i) \quad (۳-۱۷)$$

در رابطه (۳-۱۷) I یک تابع شناساگر^۴ است [۵۶].

۳-۱۰-۵-۲- اندازه گیری اهمیت متغیر ها

اگر تفاوت مقدار واقعی با مقدار پیش بینی شده، مقداری مثبت باشد، پیش بینی الگوریتم جنگل تصادفی درست

است و اگر تفاوت مقداری منفی باشد، پیش بینی الگوریتم جنگل تصادفی نادرست می باشد. پیش بینی هر درخت

با داده های OOB انجام می گردد. هر متغیر در داده های OOB با متغیر های تصادفی دیگر، جا به جا می شود و

دوباره درخت با آن داده تصادفی، پیش بینی انجام می دهد. بر اساس اختلاف میان دو پیش بینی صورت گرفته

قبل و در انتهای آموزش، تفاوت برای هر داده محاسبه می شود. اگر M میانگین تفاوت بر اساس داده های OOB

و M_j میانگین تفاوت ها بر اساس داده های OOB تصادفی باشد که j امین متغیر به صورت تصادفی در آن جا به

جا شده باشد، از طریق رابطه (۳-۱۸) اهمیت متغیر j ام محاسبه می گردد.

¹ Out of bag

² Mean Square Error

³ Error rate

⁴ Indicator Function

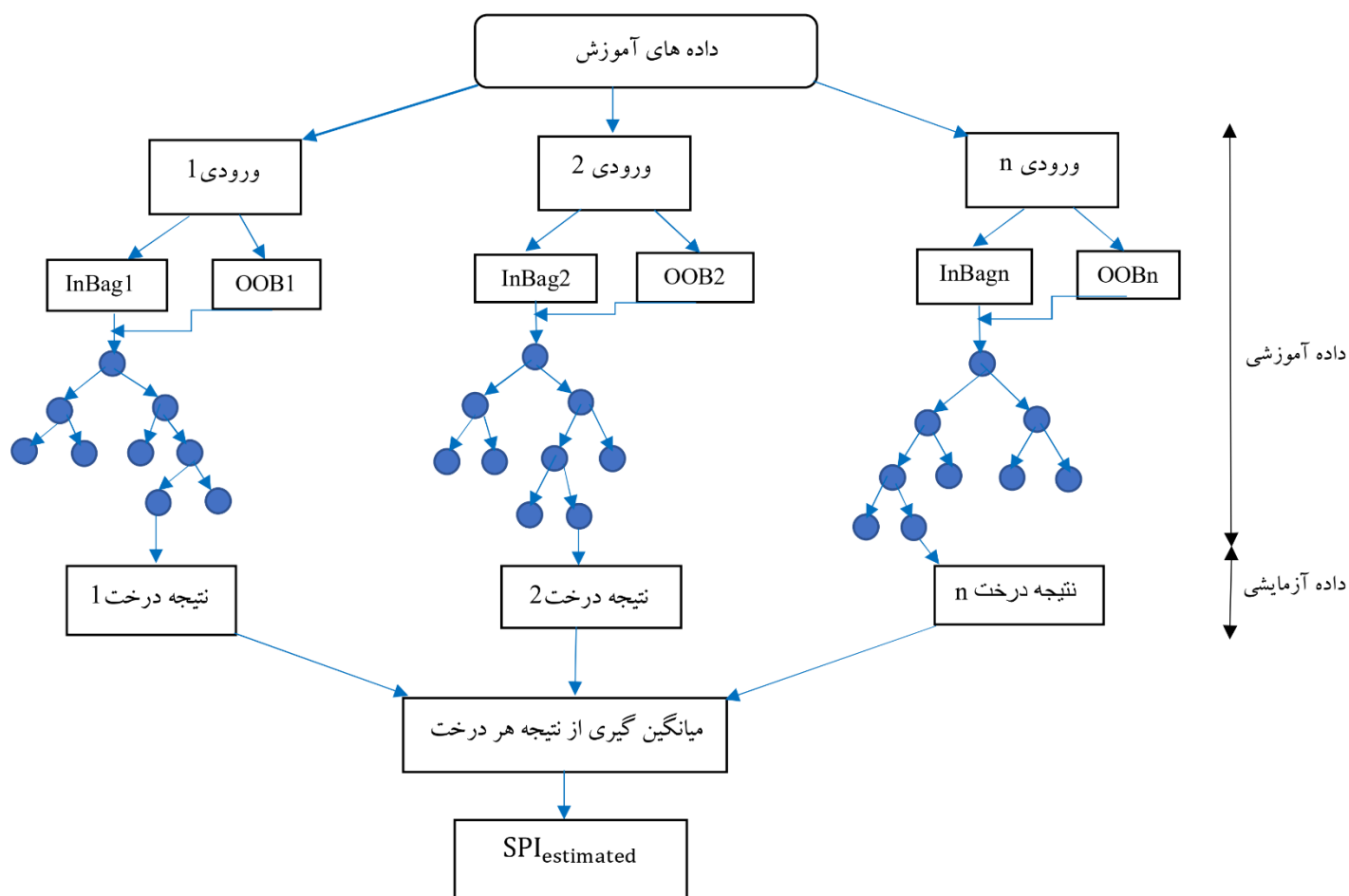
$$\text{Margin} = M - M_j$$

(۱۸-۳)

در رابطه فوق هر چه میزان تفاوت زیادتر باشد، اهمیت متغیر بیش تر است [۵۶].

۳-۱۰-۶- مزایای RF

واریانس کم، کم بودن همبستگی میان درخت ها، قابلیت بالا و مطلوب در پیش بینی، تخمین خطا با استفاده از داده OOB بخشی از مزیت های الگوریتم جنگل تصادفی است [۲۲]. در الگوریتم جنگل تصادفی تعداد محدودی توصیف کننده در شکافتگی گره استفاده می شود اما زمانی که یک درخت تصمیم به کار گرفته شود، همه توصیف کننده ها برای شکافتگی هر گره امتحان می شوند. برای افزایش دقت مدلسازی اغلب نیاز است درخت تصمیم در چند مرحله هرس شود اما الگوریتم جنگل تصادفی نیاز به هرس شدن درخت ندارد [۵۶]. در شکل ۳-۴ نمایی شماتیک از مکانیسم الگوریتم جنگل تصادفی جهت تخمین SPI نشان داده شده است که هر درخت به کمک نمونه بوت استرپ ایجاد می شود و به کمک زیر مجموعه ای تصادفی از متغیر ها یا فاکتور های هر گره، تقسیم هر درخت انجام می گردد. متغیر های تصادفی در هر گره از میان کل متغیر های موجود (ρ)، انتخاب می شود و از نمونه های خارج از کیسه (OOB) به منظور ارزیابی مدل استفاده می گردد. پس از تقسیم شدن هر درخت ، با میانگین گیری از نتیجه هر درخت ، SPI در الگوریتم جنگل تصادفی پیش بینی می شود.



شکل ۳-۴ نمایشی از به کار گیری الگوریتم جنگل تصادفی در این مطالعه

۳-۱۱- پارامتر های آماری

در تمام پارامتر های آماری ذکر شده D_n داده مشاهداتی، M_n داده پیش بینی شده، N تعداد داده های مورد بررسی، $\bar{D}_n$ میانگین داده مشاهداتی و $\bar{M}_n$ میانگین داده پیش بینی شده می باشد [۶۰].

۳-۱۱-۱ ضریب تبیین (R^2)

ضریب تبیین بیانگر انحراف نسبی میان نتایج مشاهداتی و نتایج پیش بینی شده است که از جذر ضریب همبستگی پیرسون (PCC)^۱ به دست می آید و از رابطه (۳-۱۹) به دست می آید:

$$R^2 = \frac{\{\sum_{n=1}^N (D_n - \bar{D}_n)(M_n - \bar{M}_n)\}^2}{\sum_{n=1}^N (D_n - \bar{D}_n)^2 \sum_{n=1}^N (M_n - \bar{M}_n)^2} \quad (۳-۱۹)$$

۳-۱۱-۲ جذر میانگین مربعات (MSE)

این پارامتر جذر میانگین مربعات^۲ را بین مقادیر مشاهداتی و پیش بینی شده برآورد می کند و از رابطه (۳-۲۰) به دست می آید:

$$MSE = \frac{1}{N} \sum_{n=1}^N (D_n - M_n)^2 \quad (۳-۲۰)$$

۳-۱۱-۳ خطای جذر میانگین مربعات (RMSE)

خطای جذر میانگین مربعات^۳ را بین مقادیر مشاهداتی و پیش بینی شده محاسبه می کند و از رابطه (۳-۲۱) قابل محاسبه است:

$$RMSE = \sqrt{\frac{1}{N} \sum_{n=1}^N (D_n - M_n)^2} \quad (۳-۲۱)$$

^۱ Pearson Correlation Coefficient

^۲ Mean Square Error

^۳ Root Mean Square Error

۳-۱۱-۴- میانگین خطای مطلق (MAE)

میانگین خطا مطلق^۱ را بین مقادیر مشاهداتی و پیش بینی شده نشان می‌دهد و از رابطه (۳-۲۲) قابل محاسبه است:

$$MAE = \frac{1}{N} \sum_{n=1}^N |D_n - M_n| \quad (۳-۲۲)$$

۳-۱۱-۵- شاخص توافق (IOA)

شاخص توافق^۲ مناسب بودن مدل را نشان می‌دهد و عددی بدون بعد در بازه ۰ و ۱ است که هر چه مقدار شاخص توافق زیاد تر باشد، دقت و کارایی مدل بالا تر است که این شاخص از رابطه (۳-۲۳) قابل محاسبه است [۶۰]:

$$IOA = 1 - \frac{\sum_{n=1}^N (D_n - M_n)^2}{\sum_{n=1}^N (|D_n - \bar{D}_n| + |M_n - \bar{M}_n|)^2} \quad (۳-۲۳)$$

۳-۱۱-۶- نرمال شده خطای جذر میانگین مربعات (NRMSE)

خطای جذر میانگین مربعات را به صورت نرمال شده^۳ بین مقادیر مشاهداتی و پیش بینی شده محاسبه می‌کند. هرچه مقدار این پارامتر کم تر باشد، دقت پیش بینی مدل بالاتر است. این پارامتر از رابطه (۳-۲۴) قابل محاسبه است [۶۰ و ۳۸]:

$$NRMSE = \frac{RMSE}{D_{max} - D_{min}} * 100(\%) \quad (۳-۲۴)$$

¹ Mean Absolute Error

² Index Of Agreement

³ Normalized Root Mean Square Error

۳-۱۲- مراحل انجام تحقیق

۳-۱۲-۱- اطمینان یافتن از همگن بودن و تصادفی بودن داده ها

برای تعیین همگن بودن داده ها از روش ران تست و به منظور تعیین تصادفی بودن داده ها از آزمون همبستگی تاوکندال در محیط نرم افزار SPSS استفاده شده است.

۳-۱۲-۲- تشکیل سری زمانی و آماده سازی داده های بارش برای برآورد SPI

منظور از تشکیل سری زمانی، قرار دادن داده ها به ترتیب وقوع و پشت سر هم می باشد. برای محاسبه SPI در مقیاس فصلی و سالانه، لازم است داده های بارش، به داده های مناسب برای محاسبه SPI در مقیاس مورد نظر، تبدیل گردد. پنجره زمانی ثابت در نظر گرفته شده است.

۳-۱۲-۳- برازش تابع

در این مطالعه با توجه به پیوسته بودن داده های بارش از توزیع احتمالاتی گاما که جزء توزیع های پیوسته است، استفاده شده است. به منظور برآورد پارامتر های تابع توزیع گاما نرم افزار Easy Fit ۵.۵ به کار گرفته شده است.

۳-۱۲-۴- محاسبه SPI

برای هر دو ایستگاه مورد مطالعه، SPI در مقیاس سالانه و فصلی به کمک نرم افزار Excel محاسبه گشته است.

۳-۱۲-۵- انجام تحلیل PACF

تحلیل PACF را برای SPI در مقیاس سالانه و فصلی به کمک نرم افزار SPSS انجام داده و تاخیر های زمانی معنادار مشخص شده است.

۳-۱۲-۶- پیش بینی SPI به کمک الگوریتم RF

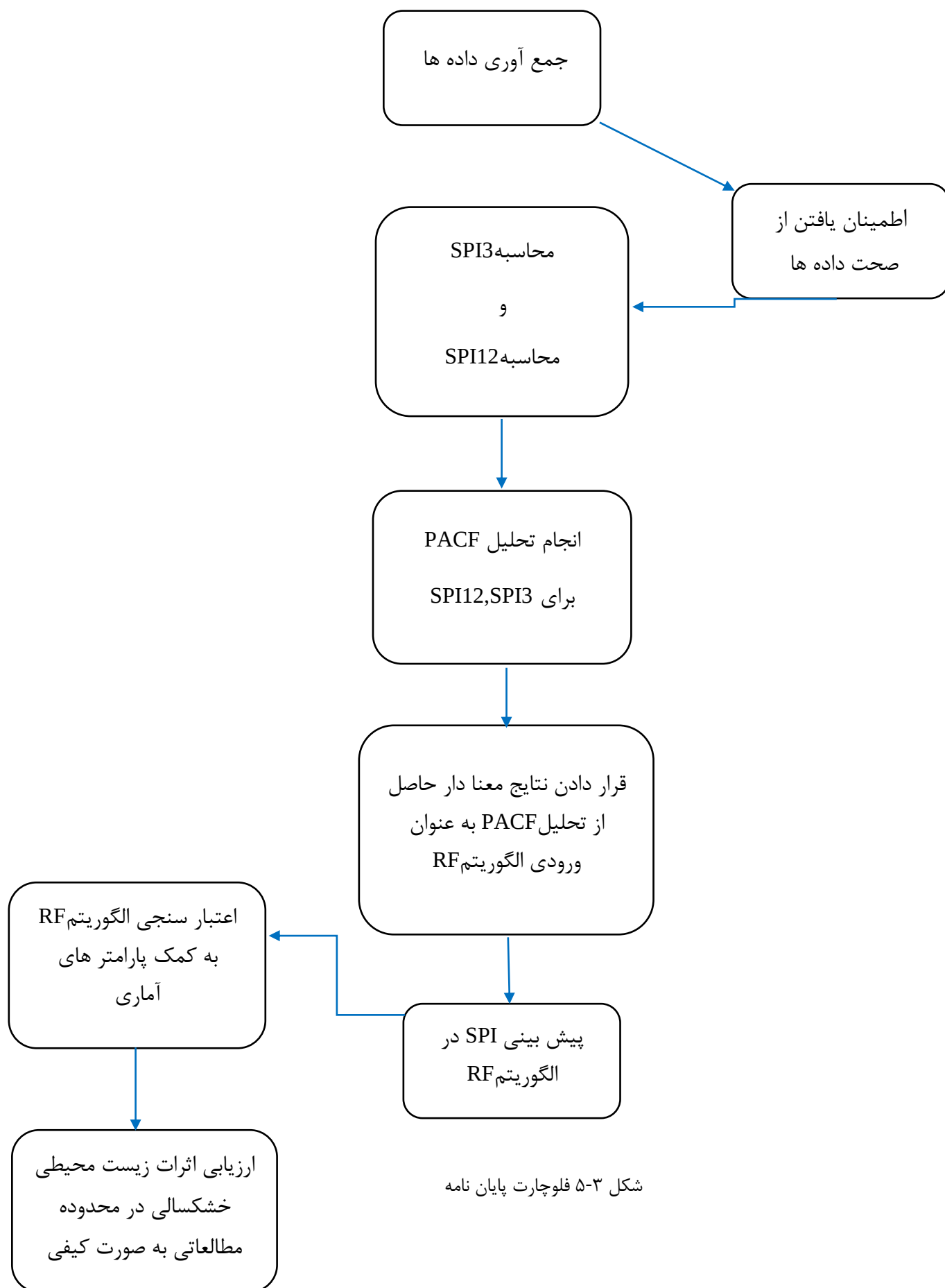
تاخیر های زمانی معنادار حاصل از تحلیل PACF به عنوان ورودی الگوریتم RF قرار داده شده است و پس از آموزش الگوریتم RF ، SPI پیش بینی شده است. داده های آموزش و آزمایش با دو ورودی متفاوت بررسی شده اند که برای پیاده سازی الگوریتم RF ، نرم افزار برنامه نویسی Rstudio به کار گرفته شده است.

۳-۱۲-۷- اعتبار سنجی الگوریتم RF

اعتبار پیش بینی در مورد شاخص بارش استاندارد مشاهداتی و پیش بینی شده در مجموعه آموزش و آزمایش به کمک شاخص های آماری مانند R^2 ، MSE، RMSE، NRMSE ، MAE و IOA سنجش شده است.

۳-۱۲-۸- ارزیابی اثرات زیست محیطی خشکسالی

اثرات زیست محیطی خشکسالی در مناطق مورد مطالعه به صورت کیفی شرح داده شده است.



شکل ۳-۵ فلوجارت پایان نامه

فصل چہارم: نتیجہ و بحث

به منظور ارزیابی کمی و کیفی خشکسالی در محدوده مطالعاتی، شاخص بارش استاندارد در مقیاس فصلی و سالانه محاسبه شده است. فراوانی وقوع و درصد رخداد هر طبقه خشکسالی در طول سری زمانی مشخص شده است. وضعیت SPI در فصل های بهار، تابستان، پاییز، زمستان و مقیاس سالانه به صورت سری زمانی در دو ایستگاه نمایش داده شده است. کمترین مقدار SPI، تاریخ وقوع اوج خشکسالی و شدت خشکسالی در طول سری زمانی در دو ایستگاه مشخص شده است. با هدف مشخص شدن وابستگی زمانی SPI در مقیاس فصلی و سالانه در طول سری زمانی، تحلیل خودهمبستگی جزئی انجام شده است. نتایج تحلیل خودهمبستگی جزئی به عنوان ورودی الگوریتم جنگل تصادفی استفاده شده است و SPI پیش بینی شده است. ۷۰ درصد نمونه ها برای آموزش دیدن الگوریتم و ۳۰ درصد نمونه ها در فرایند آزمایش استفاده شده اند. پس از اجرا نمودن الگوریتم RF، نتایج به دست آمده بررسی و تحلیل شده اند. ضریب تبیین بین مقادیر مشاهداتی و پیش بینی شده SPI معین شده است و همچنین مقادیر مشاهداتی و پیش بینی شده SPI به صورت سری زمانی به تفکیک فصل نمایش داده شده است. دقت و اعتبار الگوریتم RF با به کارگیری پارامتر های آماری بین SPI پیش بینی شده و SPI محاسباتی تعیین شده است. اثرات زیست محیطی خشکسالی در مناطق مورد مطالعه به صورت کیفی ذکر شده است.

۲-۴- بررسی همگن بودن داده های بارش

در جدول ۱-۴ نتایج آزمون ران تست در ایستگاه زابل بیان شده است که ماه های ۱ تا ۱۲ به ترتیب بیانگر ماه های ژانویه تا دسامبر می باشد. لازم به ذکر است داده های ماه ۷ و ماه ۸ ایستگاه زابل در کل سری زمانی صفر بوده است.

جدول ۱-۴ آزمون ران تست در ایستگاه زابل

پارامترهای آماری	ماه ۱	ماه ۲	ماه ۳	ماه ۴	ماه ۵	ماه ۶	ماه ۷	ماه ۸	ماه ۹	ماه ۱۰	ماه ۱۱	ماه ۱۲
Test Valuea	12.4670	9.8957	14.7860	4.6373	1.1760	0.0673	0.0000b	0.0000b	0.0003	0.3283	3.1990	6.9267
Cases < Test Value	17	19	21	21	22	29	0	0	29	26	19	22
Cases >= Test Value	13	11	9	9	8	1	30	30	1	4	11	8
Total Cases	30	30	30	30	30	30	30	30	30	30	30	30
Number of Runs	13	17	10	12	10	3	1c	1c	3	5	16	10
Z	-0.846	0.628	-1.381	-0.490	-1.072	0.000			0.000	-2.043	0.227	-1.072
Asymp. Sig. (2-tailed)	0.398	0.530	0.167	0.624	0.284	1.000			1.000	0.041	0.820	0.284

a. Mean

b. All values are greater than or less than the cutoff. Runs Test cannot be performed.

c. Only one run occurs. Runs Test cannot be performed.

داده های ایستگاه زابل در سطح معنی داری ۱ درصد همگن می باشند زیرا سطح معنی داری آزمون یا Asymp.

Sig. (2-tailed) در هر ماه بزرگ تر از ۰/۰۱ می باشد. در جدول ۲-۴ نتایج آزمون ران تست در ایستگاه زاهدان

بیان شده است که ماه های ۱ تا ۱۲ به ترتیب بیانگر ماه های ژانویه تا دسامبر می باشد.

جدول ۴-۲ آزمون ران تست در ایستگاه زاهدان

پارامترهای آماری	ماه ۱	ماه ۲	ماه ۳	ماه ۴	ماه ۵	ماه ۶	ماه ۷	ماه ۸	ماه ۹	ماه ۱۰	ماه ۱۱	ماه ۱۲
Test Value ^a	16.4967	12.4027	13.9223	8.6270	3.6700	1.2473	0.9587	0.1600	0.1030	3.9770	3.8917	12.9183
Cases < Test Value	20	20	19	19	21	26	27	28	26	24	20	20
Cases >= Test Value	10	10	11	11	9	4	3	2	4	6	10	10
Total Cases	30	30	30	30	30	30	30	30	30	30	30	30
Number of Runs	17	16	18	10	17	7	7	5	7	13	17	15
Z	0.910	0.490	1.030	-1.778	1.292	-0.364	0.110	0.000	-0.364	1.126	0.910	0.070
Asymp. Sig. (2-tailed)	0.363	0.624	0.303	0.075	0.196	0.716	0.912	1.000	0.716	0.260	0.363	0.944

a. Mean

داده های ایستگاه زاهدان در سطح معنی داری ۵ درصد همگن می باشند زیرا سطح معنی داری آزمون یا Asymp. Sig. (2-tailed) در هر ماه بزرگ تر از ۰/۰۵ می باشد.

۴-۳ بررسی وضعیت روند داده های بارش

نتایج آزمون تاوکندال در ایستگاه زابل و زاهدان در جدول ۴-۳ و ۴-۴ بیان شده است.

جدول ۴-۳ آزمون تاوکندال در ایستگاه زابل

Kendall's tau_b Correlation Coefficient	Sig. (2-tailed)
-0.005	0.906

جدول ۴-۴ آزمون تاوکندال در ایستگاه زاهدان

Kendall's tau_b Correlation Coefficient	Sig. (2-tailed)
0.021	0.578

نتایج نشان داد داده های هر دو ایستگاه در سطح معنی داری ۵ درصد بدون روند و تصادفی می باشند زیرا سطح معنی داری آزمون یا Asymp. Sig. (2-tailed) در هر ایستگاه بزرگ تر از ۰/۰۵ می باشد.

۴-۴-نتایج برآورد پارامترهای تابع گاما

در نرم افزار Easy Fit تابع گاما بر داده های بارش، برازش داده شده است و پارامترهای تابع گاما معین شده اند. پارامترهای تابع گاما در هر مقیاس زمانی SPI در مورد ایستگاه زابل و زاهدان در جدول ۴-۵، ۴-۶، ۴-۷ و ۴-۸ نمایش داده شده اند.

جدول ۴-۵ پارامترهای تابع گاما در ایستگاه زابل و مقیاس ۳ ماهه

α	β
0.4698592	28.45728

جدول ۴-۶ پارامترهای تابع گاما در ایستگاه زابل و مقیاس ۱۲ ماهه

α	β
2.056777	26.00363

جدول ۴-۷ پارامترهای تابع گاما در ایستگاه زاهدان و مقیاس ۳ ماهه

α	β
0.571543	34.28205

جدول ۴-۸ پارامترهای تابع گاما در ایستگاه زاهدان و مقیاس ۱۲ ماهه

α	β
2.597239	30.17615

۴-۵- نتایج SPI در مقیاس ۳ و ۱۲ ماهه

مقدار SPI در مقیاس ۳ ماهه و ۱۲ ماهه در مورد ایستگاه زابل و زاهدان در جدول ۴-۹ و ۴-۱۰ ذکر شده است.

جدول ۴-۹ SPI در ایستگاه زابل و زاهدان در مقیاس ۳ ماهه

ایستگاه زاهدان	ایستگاه زابل	نوع فصل	سری زمانی
2.141713872	1.935244369	زمستان	1991
1.749701883	1.798592228	بهار	1991
-0.967421566	-0.548522283	تابستان	1991
-0.151572291	-0.29094264	پاییز	1991
0.558290052	1.220904156	زمستان	1992
0.997281192	1.629901346	بهار	1992
-0.967421566	-0.548522283	تابستان	1992
1.103194504	-0.548522283	پاییز	1992
0.962636609	1.93162546	زمستان	1993
-0.087148009	0.769568548	بهار	1993
-0.253444226	-0.548522283	تابستان	1993
-0.967421566	-0.548522283	پاییز	1993
0.977591268	1.672041264	زمستان	1994
1.021507922	1.988759072	بهار	1994
-0.967421566	-0.548522283	تابستان	1994
0.281188912	0.190492819	پاییز	1994
0.636068188	0.886723342	زمستان	1995
1.045324922	1.328294196	بهار	1995
0.723999992	-0.548522283	تابستان	1995
-0.742704055	-0.548522283	پاییز	1995
2.66023859	2.02539956	زمستان	1196
1.678242502	1.821736693	بهار	1196
-0.3927652	-0.548522283	تابستان	1196
-0.3927652	-0.25589781	پاییز	1196
0.714891709	0.727210597	زمستان	1997
1.229384865	1.614904671	بهار	1997
-0.177610376	-0.548522283	تابستان	1997
1.414753541	0.744381152	پاییز	1997
1.248074315	2.064801356	زمستان	1998

ایستگاه زاهدان	ایستگاه زابل	نوع فصل	سری زمانی
0.208018222	0.110805217	بهار	1998
-0.841926999	-0.548522283	تابستان	1998
-0.967421566	-0.548522283	پاییز	1998
1.171937581	2.03747018	زمستان	1999
0.012112501	0.631351205	بهار	1999
-0.329888199	-0.548522283	تابستان	1999
-0.00796312	0.709718377	پاییز	1999
0.771435506	0.930131493	زمستان	2000
0.365866461	0.234764586	بهار	2000
-0.967421566	-0.548522283	تابستان	2000
-0.706745321	0.152221974	پاییز	2000
0.044346629	0.142189457	زمستان	2001
-0.00120639	0.336109379	بهار	2001
-0.967421566	-0.548522283	تابستان	2001
-0.967421566	-0.548522283	پاییز	2001
0.267048342	1.101108589	زمستان	2002
-0.263701402	0.371064121	بهار	2002
-0.967421566	-0.548522283	تابستان	2002
0.831021393	0.410712734	پاییز	2002
0.687133871	1.336028072	زمستان	2003
0.00548494	-0.075404616	بهار	2003
-0.274160277	-0.548522283	تابستان	2003
-0.967421566	0.243210142	پاییز	2003
0.511690203	0.740117771	زمستان	2004
-0.621510584	-0.075404616	بهار	2004
-0.967421566	-0.548522283	تابستان	2004
-0.967421566	-0.548522283	پاییز	2004
2.221159327	2.463268248	زمستان	2005
1.210470834	2.154539955	بهار	2005
-0.516766388	-0.548522283	تابستان	2005
-0.967421566	0.040898931	پاییز	2005
0.523717985	0.757057658	زمستان	2006
-0.182064407	0.089807919	بهار	2006
-0.967421566	-0.548522283	تابستان	2006
-0.138967327	0.377836574	پاییز	2006

ایستگاه زاهدان	ایستگاه زابل	نوع فصل	سری زمانی
1.133741552	1.36454262	زمستان	2007
1.116748688	0.430333637	بهار	2007
0.815128426	0.042158983	تابستان	2007
-0.093299305	-0.493813005	پاییز	2007
1.507464221	0.680105952	زمستان	2008
-0.687452453	-0.109446334	بهار	2008
-0.967421566	-0.548522283	تابستان	2008
-0.51315459	-0.493813005	پاییز	2008
1.473670885	0.968810297	زمستان	2009
1.017177915	1.104168195	بهار	2009
-0.967421566	-0.548522283	تابستان	2009
-0.967421566	-0.493813005	پاییز	2009
1.760280856	0.932257063	زمستان	2010
0.345210437	0.261362365	بهار	2010
-0.967421566	-0.548522283	تابستان	2010
-0.38608376	-0.473156215	پاییز	2010
1.664773765	1.12447657	زمستان	2011
1.277325795	0.540248778	بهار	2011
-0.915633464	-0.548522283	تابستان	2011
-0.139799527	-0.109446334	پاییز	2011
0.571185069	1.131672214	زمستان	2012
0.749137259	0.33897651	بهار	2012
-0.828979149	-0.493813005	تابستان	2012
-0.614210823	0.583238729	پاییز	2012
0.84315411	0.413915022	زمستان	2013
0.948276044	0.344672455	بهار	2013
-0.141467361	-0.548522283	تابستان	2013
0.006815502	0.830614299	پاییز	2013
0.947515992	1.258628748	زمستان	2014
0.971380005	0.953290964	بهار	2014
-0.902533799	-0.548522283	تابستان	2014
-0.724042272	0.153214303	پاییز	2014
0.933761241	1.108134475	زمستان	2015
0.917015058	1.278582398	بهار	2015
-0.700151664	-0.548522283	تابستان	2015

ایستگاه زاهدان	ایستگاه زابل	نوع فصل	سری زمانی
1.234806776	0.577088512	پاییز	2015
0.973370672	0.818463887	زمستان	2016
0.817931266	1.313484161	بهار	2016
-0.915633464	-0.548522283	تابستان	2016
-0.932294445	-0.39121419	پاییز	2016
1.154057249	0.856010892	زمستان	2017
0.512789446	0.308951094	بهار	2017
-0.932294445	-0.548522283	تابستان	2017
-0.932294445	-0.548522283	پاییز	2017
0.164038924	-0.05551638	زمستان	2018
0.089559554	0.612829501	بهار	2018
-0.932294445	-0.548522283	تابستان	2018
-0.706745321	-0.39121419	پاییز	2018
0.26894714	0.545081588	زمستان	2019
1.677757301	1.508498727	بهار	2019
-0.967421566	-0.548522283	تابستان	2019
0.270841775	0.740117771	پاییز	2019
2.048671023	1.636500862	زمستان	2020
0.887918831	1.299574075	بهار	2020
-0.967421566	-0.548522283	تابستان	2020
0.470925773	0.35031833	پاییز	2020

جدول ۴-۱۰ SPI در ایستگاه زابل و زاهدان در مقیاس ۱۲ ماهه

ایستگاه زاهدان	ایستگاه زابل	سری زمانی
1.719408956	1.424067109	1991
0.430491751	0.648860854	1992
-0.744225779	0.690343518	1993
0.212966168	1.420778661	1994
0.237220633	0.063982883	1995
2.163530512	1.527412366	1996
1.00215322	0.59540936	1997
-0.33300462	0.62005766	1998
-0.291856878	0.993590264	1999
-0.799606411	-0.744821834	2000
-1.986640659	-1.811754349	2001
-0.715571127	-0.333625806	2002
-1.051661342	-0.277752118	2003
-1.700163124	-1.328649412	2004
1.379675808	2.216768151	2005
-1.293054322	-0.896984395	2006
0.773428776	-0.127593637	2007
-0.200173396	-1.439128064	2008
0.447727285	-0.106464519	2009
0.405261582	-0.867662376	2010
0.911243528	-0.395346472	2011
-0.639429626	-0.210577366	2012
0.02946228	-0.652670418	2013
-0.096625314	0.127387398	2014
0.711321058	0.43090367	2015
-0.219523817	-0.002077253	2016
-0.26579145	-0.942863965	2017
-1.729186863	-1.518144217	2018
0.446988375	0.390865713	2019
1.213084696	0.831442929	2020

بر مبنای مقدار هر شاخص، طبقه خشکسالی در طی سری زمانی مشخص شده است. در جدول ۱۱-۴ و ۱۲-۴ به ترتیب فراوانی و درصد رخداد طبقات مختلف خشکسالی در مقیاس ۳ و ۱۲ ماهه هر دو ایستگاه ذکر شده است.

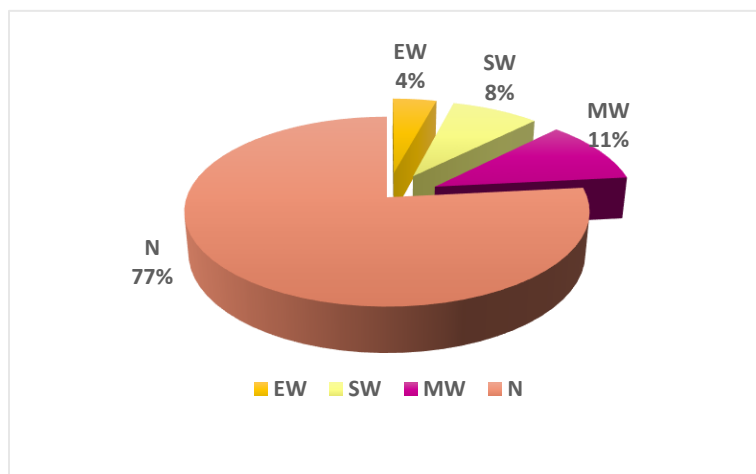
جدول ۱۱-۴ فراوانی رخداد طبقات مختلف خشکسالی در ایستگاه زابل و زاهدان

ایستگاه	شاخص	EW	SW	MW	N	MD	SD	ED
زابل	SPI3	5	10	13	92	0	0	0
زابل	SPI12	1	1	2	22	2	2	0
زاهدان	SPI3	4	6	16	94	0	0	0
زاهدان	SPI12	1	1	3	20	2	3	0

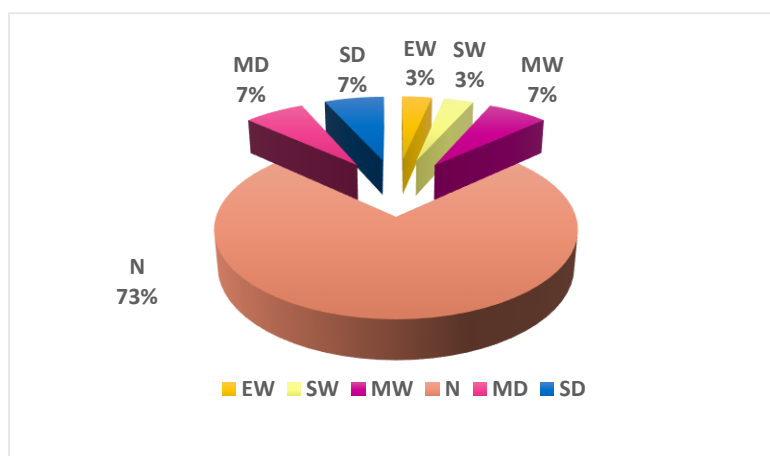
جدول ۱۲-۴ درصد رخداد طبقات مختلف خشکسالی در ایستگاه زابل و زاهدان

ایستگاه	شاخص	EW	SW	MW	N	MD	SD	ED
زابل	SPI3	4	8	11	77	0	0	0
زابل	SPI12	3	3	7	73	7	7	0
زاهدان	SPI3	3	5	14	78	0	0	0
زاهدان	SPI12	3	3	10	67	7	10	0

درصد رخداد طبقات گوناگون خشکسالی در شکل های ۱-۴، ۲-۴، ۳-۴ و ۴-۴ نشان داده شده است.

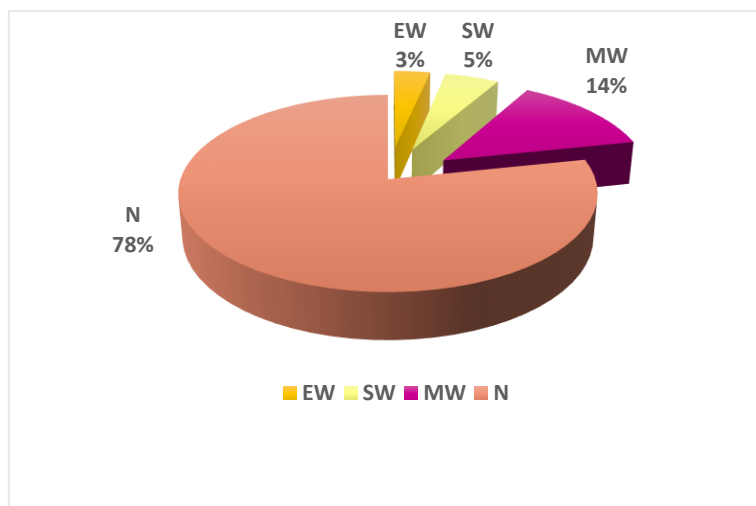


شکل ۴-۱ نمودار درصد رخداد طبقات مختلف خشکسالی SPI3 در ایستگاه زابل

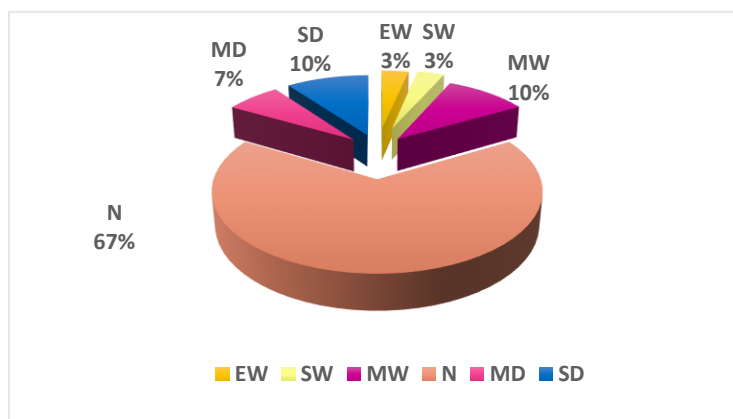


شکل ۴-۲ نمودار درصد رخداد طبقات مختلف خشکسالی SPI12 در ایستگاه زابل

طبق شکل های ۴-۱، ۴-۲ در بازه ۳۰ ساله مورد بررسی در ایستگاه زابل، طبقه خشکسالی تقریباً نرمال (N) بیشترین درصد وقوع را در دو مقیاس شاخص دارد طوری که سهم طبقه خشکسالی تقریباً نرمال (N) در مقیاس ۳ ماهه و ۱۲ ماهه به ترتیب برابر ۷۷ درصد و ۷۳ درصد می باشد.



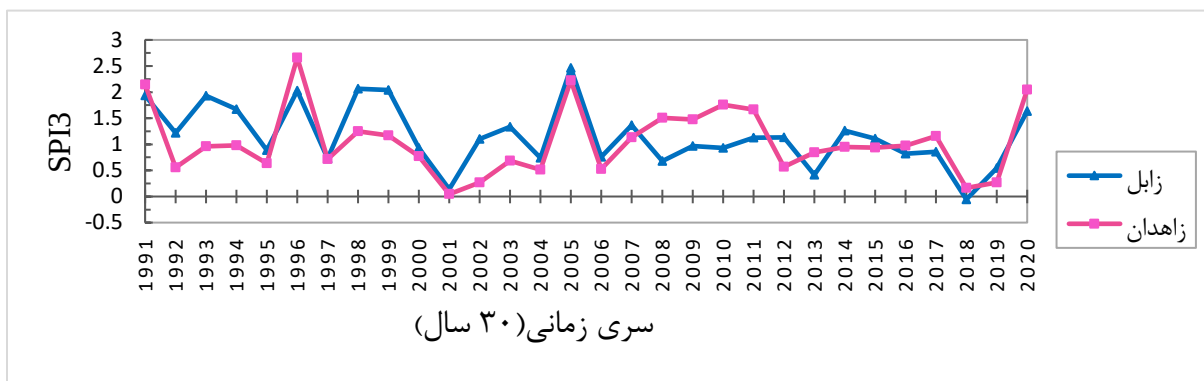
شکل ۳-۴ نمودار درصد رخداد طبقات خشکسالی SPI3 در ایستگاه زاهدان



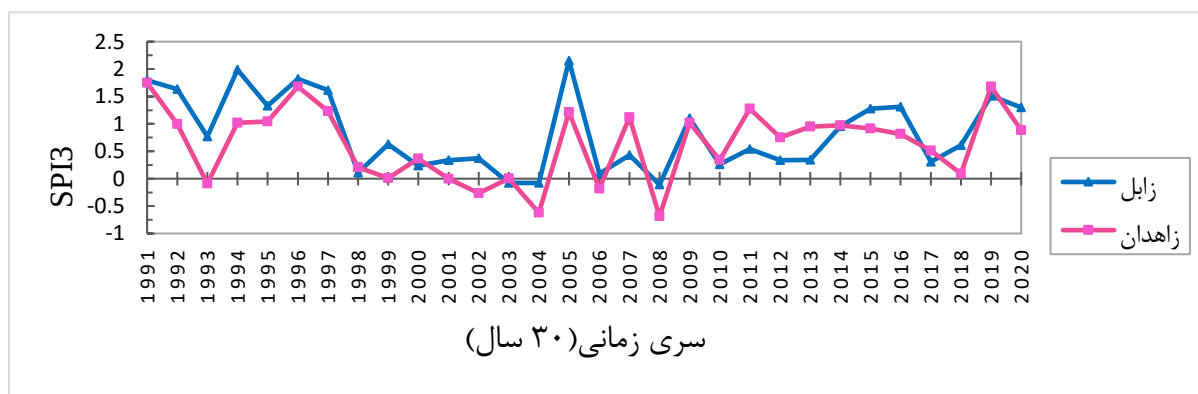
شکل ۴-۴ نمودار درصد رخداد طبقات خشکسالی SPI12 در ایستگاه زاهدان

بر اساس شکل های ۳-۴، ۴-۴ در طول سری زمانی در ایستگاه زاهدان، طبقه خشکسالی تقریباً نرمال (N) بیشترین درصد وقوع را در دو مقیاس شاخص دارد طوری که طبقه خشکسالی تقریباً نرمال (N) در مقیاس ۳ ماهه و ۱۲ ماهه به ترتیب ۷۸ درصد و ۶۷ درصد را به خود اختصاص داده است. سایر طبقات مختلف خشکسالی ایستگاه ها در دو مقیاس نیز در اشکال ۱-۴ تا ۴-۴ قابل مشاهده است.

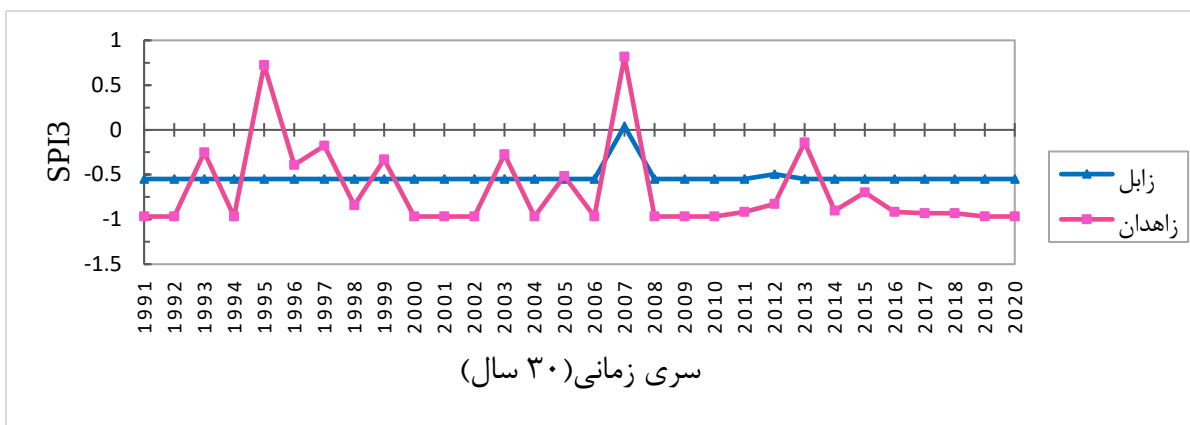
SPI فصل زمستان، بهار، پاییز، تابستان و سالانه به ترتیب در اشکال ۴-۵، ۴-۶، ۴-۷، ۴-۸ و ۴-۹ در مورد ایستگاه زابل و زاهدان نمایش داده شده است.



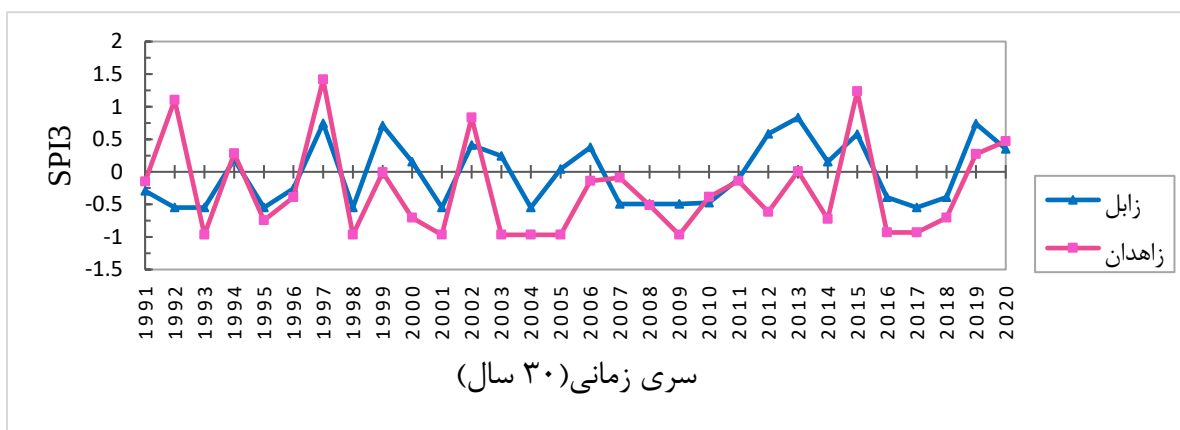
شکل ۴-۵ نمودار SPI ایستگاه زابل و زاهدان در فصل زمستان



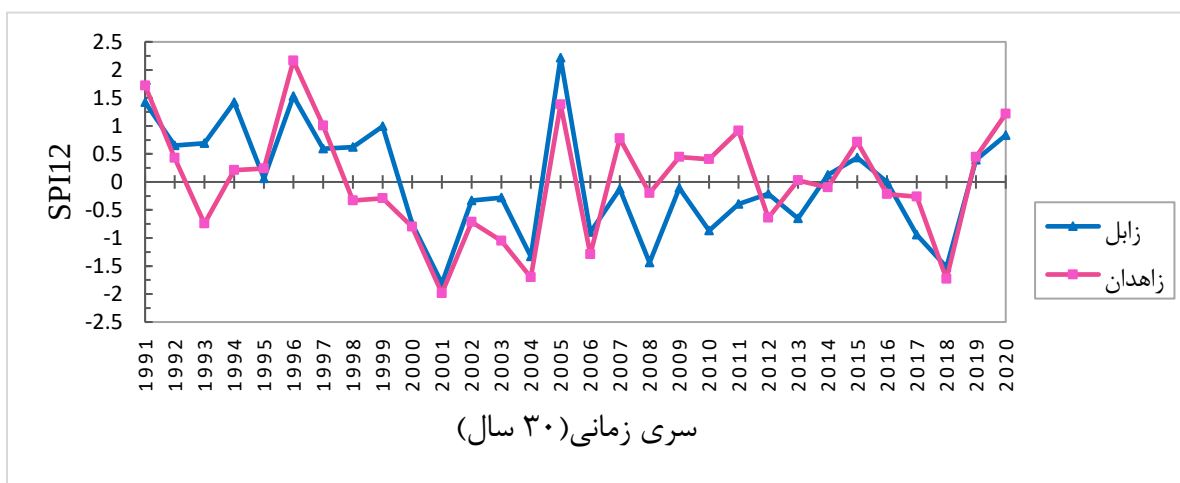
شکل ۴-۶ نمودار SPI ایستگاه زابل و زاهدان در فصل بهار



شکل ۴-۷ نمودار SPI ایستگاه زابل و زاهدان در فصل تابستان



شکل ۴-۸ نمودار SPI ایستگاه زابل و زاهدان در فصل پاییز



شکل ۴-۹ نمودار SPI ایستگاه زابل و زاهدان در مقیاس سالانه

در جدول ۴-۱۳ کمترین مقدار SPI، تاریخ وقوع اوج خشکسالی و شدت خشکسالی در طول سری زمانی در ایستگاه زابل و زاهدان تعیین شده است. شدت خشکسالی از تقسیم نمودن مجموع مقادیر متوالی منفی SPI در طول سری زمانی بر مدت خشکسالی محاسبه شده است [۳].

جدول ۴-۱۳ کمترین مقدار SPI، تاریخ وقوع اوج خشکسالی و شدت خشکسالی

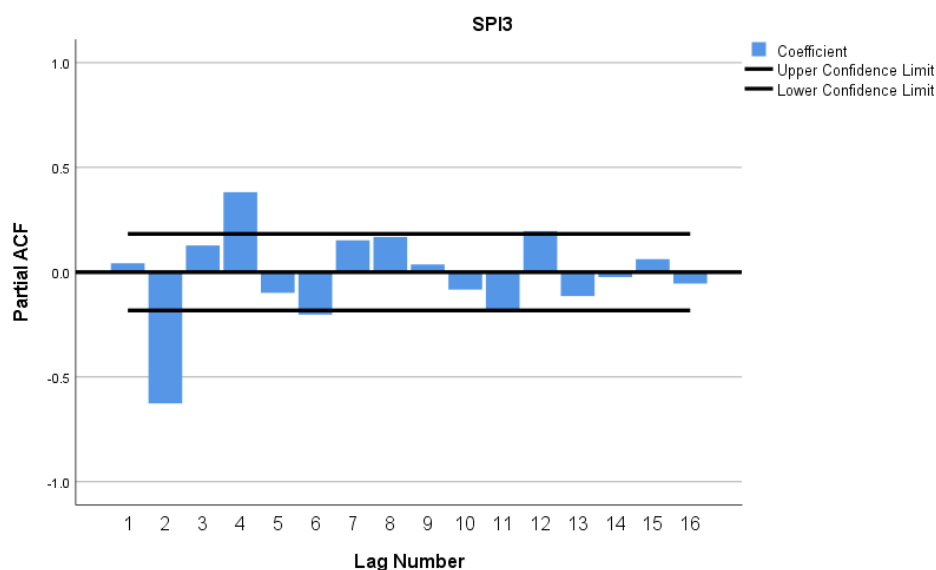
نام ایستگاه	مقیاس مورد بررسی	کمترین مقدار SPI در سری زمانی	تاریخ وقوع اوج خشکسالی	شدت خشکسالی در طول سری زمانی
زابل	فصلی	-0.548522283	فصل های فاقد بارش	-14.90606903
زابل	سالانه	-1.811754349	سال ۲۰۰۱	-2.307402588
زاهدان	فصلی	-0.967421566	فصل های فاقد بارش	-20.03449593
زاهدان	سالانه	-1.986640659	سال ۲۰۰۱	-4.694319265

شدت خشکسالی در ایستگاه زاهدان در مقیاس فصلی و سالانه نسبت به ایستگاه زابل زیادتر می باشد و اوج خشکسالی سالانه هر دو ایستگاه در سال ۲۰۰۱ رخ داده است، همچنین اوج خشکسالی هر دو ایستگاه در مقیاس فصلی در فصولی رخ داده است که فاقد بارش بوده اند.

۴-۶- نتایج تحلیل PACF در مقیاس ۳ و ۱۲ ماهه

جدول ۴-۱۴ تحلیل PACF در ایستگاه زابل برای SPI3

Lag	Partial Autocorrelation	Std.Error
1	0.043	0.091
2	-0.626	0.091
3	0.127	0.091
4	0.382	0.091
5	-0.099	0.091
6	-0.203	0.091
7	0.152	0.091
8	0.168	0.091
9	0.037	0.091
10	-0.083	0.091
11	-0.177	0.091
12	0.196	0.091
13	-0.114	0.091
14	-0.023	0.091
15	0.062	0.091
16	-0.054	0.091

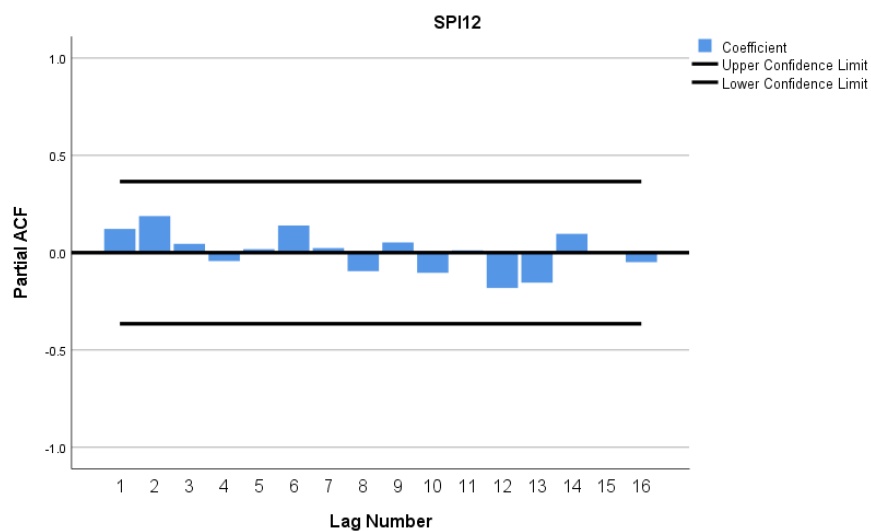


شکل ۴-۱۰ نمودار تحلیل PACF در ایستگاه زابل برای SPI3

طبق جدول ۴-۱۴ و شکل ۴-۱۰ تاخیر زمانی یا Lag های ۲ و ۴ و ۶ و ۱۲ بالا و پایین خطوط سیاه رنگ یعنی سطوح معناداری تحلیل هستند و لذا قابل قبول اند و در سایر تحلیل ها و محاسبات ریاضی قابل استفاده می باشند.

جدول ۴-۱۵ تحلیل PACF در ایستگاه زابل برای SPI12

Lag	Partial Autocorrelation	Std.Error
1	0.122	0.183
2	0.188	0.183
3	0.045	0.183
4	-0.043	0.183
5	0.018	0.183
6	0.139	0.183
7	0.024	0.183
8	-0.095	0.183
9	0.053	0.183
10	-0.104	0.183
11	0.011	0.183
12	-0.182	0.183
13	-0.154	0.183
14	0.097	0.183
15	0.001	0.183
16	-0.049	0.183

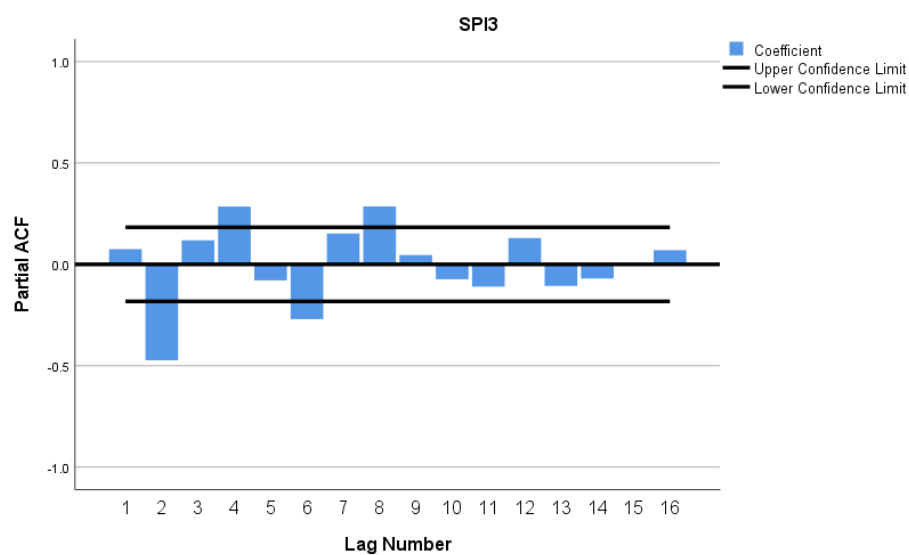


شکل ۴-۱۱ نمودار تحلیل PACF در ایستگاه زابل برای SPI12

طبق جدول ۴-۱۵ و شکل ۴-۱۱ به دلیل تغییرات اقلیمی سالانه در طی سری زمانی، تاخیر زمانی یا Lag معنادار وجود ندارد.

جدول ۴-۱۶ تحلیل PACF در ایستگاه زاهدان برای SPI3

Lag	Partial Autocorrelation	Std.Error
1	0.075	0.091
2	-0.473	0.091
3	0.119	0.091
4	0.285	0.091
5	-0.079	0.091
6	-0.270	0.091
7	0.152	0.091
8	0.285	0.091
9	0.046	0.091
10	-0.074	0.091
11	-0.110	0.091
12	0.129	0.091
13	-0.107	0.091
14	-0.070	0.091
15	-0.007	0.091
16	0.070	0.091

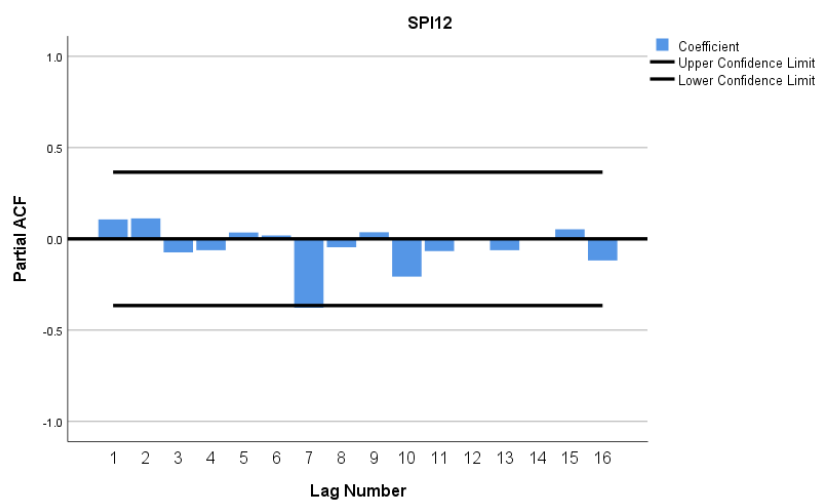


شکل ۴-۱۲ نمودار تحلیل PACF در ایستگاه زاهدان برای SPI3

طبق جدول ۴-۱۶ و شکل ۴-۱۲ تاخیر زمانی یا Lag های ۲ و ۴ و ۶ و ۸ بالا و پایین خطوط سیاه رنگ یعنی سطوح معناداری تحلیل هستند و لذا قابل قبول اند و در سایر تحلیل ها و محاسبات ریاضی قابل استفاده می باشند.

جدول ۴-۱۷ تحلیل PACF در ایستگاه زاهدان برای SPI12

Lag	Partial Autocorrelation	Std.Error
1	0.106	0.183
2	0.112	0.183
3	-0.074	0.183
4	-0.062	0.183
5	0.035	0.183
6	0.019	0.183
7	-0.378	0.183
8	-0.046	0.183
9	0.036	0.183
10	-0.207	0.183
11	-0.067	0.183
12	0.004	0.183
13	-0.062	0.183
14	0.003	0.183
15	0.052	0.183
16	-0.118	0.183



شکل ۴-۱۳ نمودار تحلیل PACF در ایستگاه زاهدان برای SPI12

طبق جدول ۴-۱۷ و شکل ۴-۱۳ به دلیل تغییرات اقلیمی سالانه در طی سری زمانی، تاخیر زمانی یا Lag معنادار وجود ندارد.

۷-۴- نتایج پیش بینی SPI به کمک الگوریتم RF

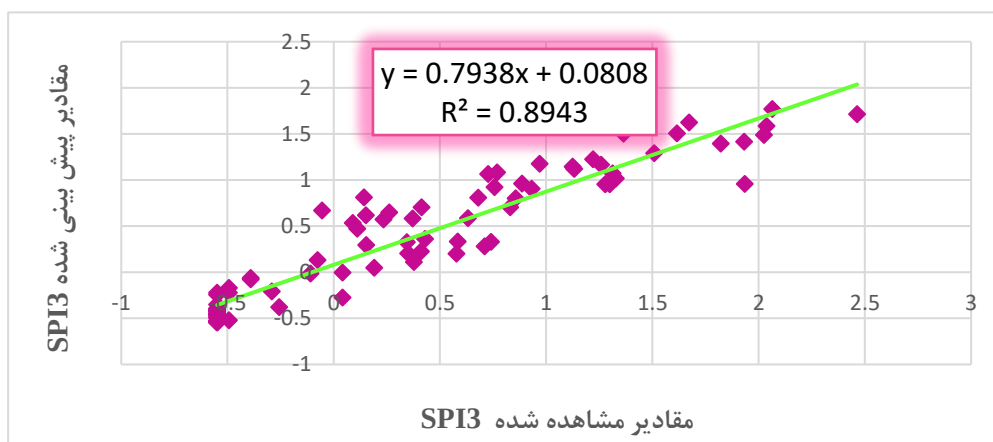
تاخیر زمانی هایی که در تحلیل PACF معنادار بودند به عنوان ورودی الگوریتم جنگل تصادفی به کار گرفته شده است. طبق شکل ۴-۱۰ که نمودار تحلیل PACF در ایستگاه زابل برای SPI3 می باشد، تاخیر زمانی های ۶،۴،۲ و ۱۲ معنادار هستند به همین دلیل به منظور مقایسه نمودن نتایج، یک بار تاخیر زمانی های ۴،۲ و ۶ و بار دیگر تاخیر زمانی های ۶،۴،۲ و ۱۲ به عنوان ورودی الگوریتم جنگل تصادفی استفاده شده اند. همچنین طبق شکل ۴-۱۲ که نمودار تحلیل PACF در ایستگاه زاهدان برای SPI3 می باشد، تاخیر زمانی های ۶،۴،۲ و ۸ معنادار هستند به همین دلیل به منظور مقایسه نمودن نتایج، یک بار تاخیر زمانی های ۴،۲ و ۶ و بار دیگر تاخیر زمانی های ۶،۴،۲ و ۸ به عنوان ورودی الگوریتم جنگل تصادفی به کار گرفته شده اند.

در جدول ۴-۱۸ پارامترهای الگوریتم جنگل تصادفی شامل ورودی های الگوریتم، تعداد درختان جنگل (n_{tree})، تعداد متغیرهای تصادفی در هر گره (m_{try})، درصد نمونه های آموزش و درصد نمونه های آزمایش به کار بسته شده در نرم افزار Rstudio نشان داده شده است. درصد نمونه های آموزش و آزمایش به ترتیب ۷۰ درصد و ۳۰ درصد در نظر گرفته شده است. به منظور همگرا شدن خطا و افزایش دقت پیش بینی SPI در الگوریتم جنگل تصادفی، تعداد درختان جنگل (n_{tree})، ۵۰۰ درخت در نظر گرفته شده است و متناسب با فرایند رگرسیون و تعداد ورودی ها، تعداد متغیرهای تصادفی در هر گره (m_{try}) در نرم افزار Rstudio معین شده است. با توجه به فرایند رگرسیون در الگوریتم جنگل تصادفی، پارامتر m_{try} ، یک سوم تعداد ورودی ها در نظر گرفته شده است. با نوشتن کد نرم افزاری در محیط نرم افزار Rstudio، SPI فصلی در الگوریتم جنگل تصادفی پیش بینی شده است.

جدول ۱۸-۴ تنظیمات الگوریتم جنگل تصادفی

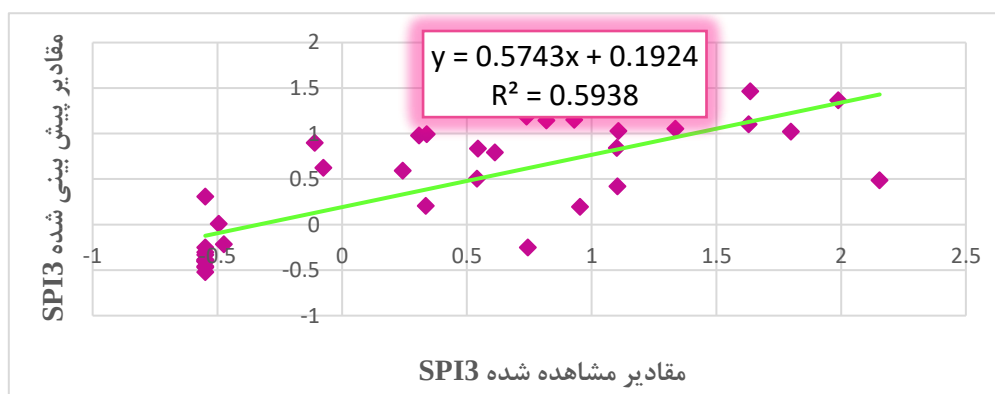
نام	مقیاس	ورودی های الگوریتم جنگل	درصد نمونه های	درصد نمونه	mtry	ntree
ایستگاه	SPI	تصادفی	آزمایش	های آموزش		
		تاخیر زمانی 12و6و4و2	0.3	0.7	2	500
زابل	فصلی	تاخیر زمانی 6و4و2	0.3	0.7	1	500
		تاخیر زمانی 8و6و4و2	0.3	0.7	2	500
زاهدان	فصلی	تاخیر زمانی 6و4و2	0.3	0.7	1	500

ضریب تبیین و معادله بهترین خط برازش بین مقادیر پیش بینی شده SPI و مشاهداتی SPI فصلی هر دو ایستگاه، با دو نوع ورودی متفاوت از تاخیر زمانی در شکل های ۱۴-۴ تا ۲۱-۴ در مجموعه آموزش و آزمایش نمایش داده شده است.



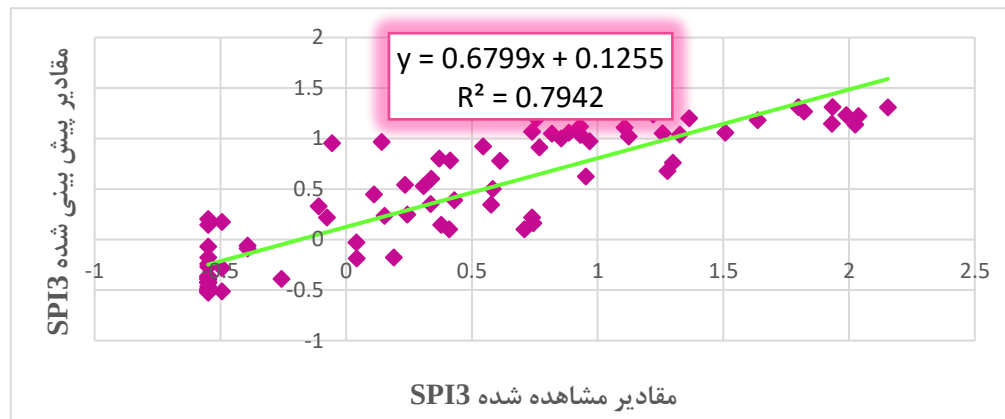
شکل ۴-۱۴ ضریب تبیین SPI فصلی زابل با تاخیر زمانی ۲ و ۴ و ۱۲ در مجموعه آموزش

در ایستگاه زابل به ازای تاخیر زمانی ۲ و ۴ و ۱۲ به عنوان ورودی الگوریتم جنگل تصادفی در مجموعه آموزش ضریب تبیین ۰/۸۹ حاصل شده است که نشان دهنده دقت مناسب الگوریتم جنگل تصادفی در پیش بینی SPI می باشد.



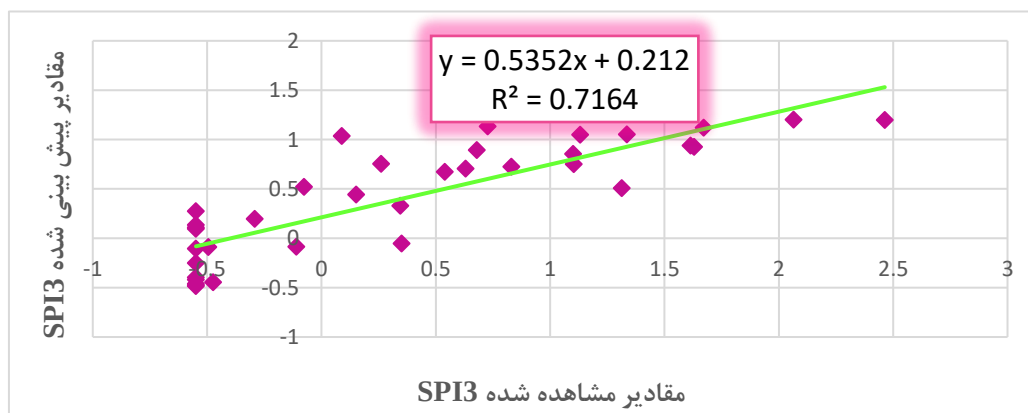
شکل ۴-۱۵ ضریب تبیین SPI فصلی زابل با تاخیر زمانی ۲ و ۴ و ۱۲ در مجموعه آزمایش

در ایستگاه زابل به ازای تاخیر زمانی ۲ و ۴ و ۱۲ به عنوان ورودی الگوریتم جنگل تصادفی در مجموعه آزمایش ضریب تبیین ۰/۵۹ حاصل شده است که نشان دهنده دقت مناسب الگوریتم جنگل تصادفی در پیش بینی SPI می باشد.



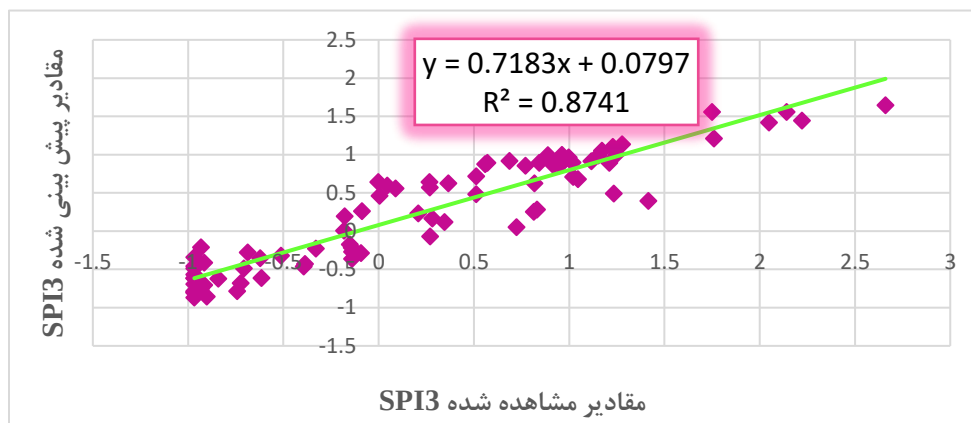
شکل ۴-۱۶ ضریب تبیین SPI فصلی زابل با تاخیر زمانی ۲ و ۴ در مجموعه آموزش

در ایستگاه زابل به ازای تاخیر زمانی ۲ و ۴ به عنوان ورودی الگوریتم جنگل تصادفی در مجموعه آموزش ضریب تبیین ۰/۷۹ حاصل شده است که نشان دهنده دقت بالا الگوریتم جنگل تصادفی در پیش بینی SPI می باشد.



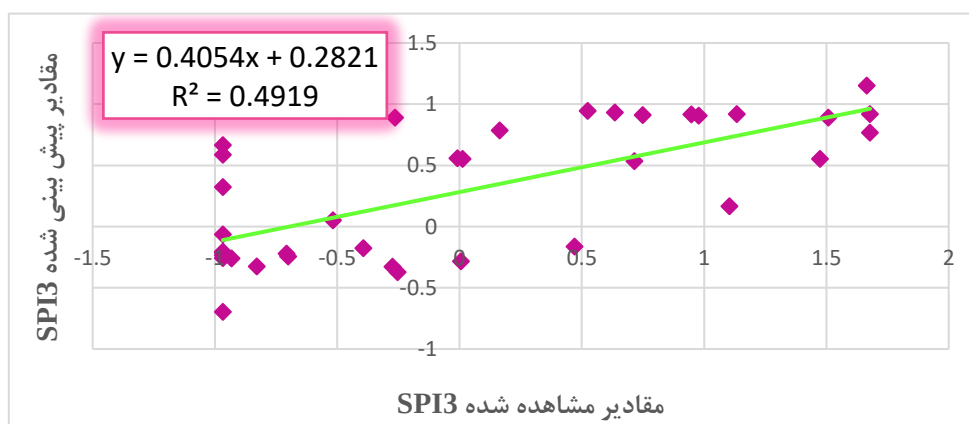
شکل ۴-۱۷ ضریب تبیین SPI فصلی زابل با تاخیر زمانی ۲ و ۴ در مجموعه آزمایش

در ایستگاه زابل به ازای تاخیر زمانی ۲ و ۴ به عنوان ورودی الگوریتم جنگل تصادفی در مجموعه آزمایش ضریب تبیین ۰/۷۲ حاصل شده است که نشان دهنده دقت بالا الگوریتم جنگل تصادفی در پیش بینی SPI می باشد.



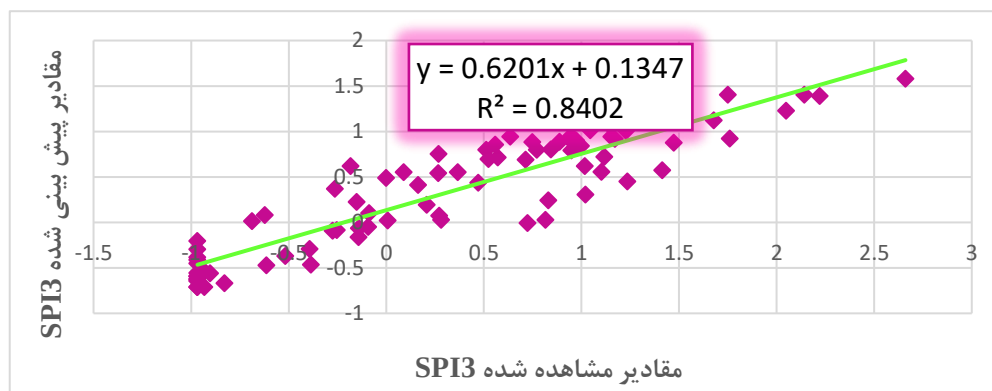
شکل ۴-۱۸ ضریب تبیین SPI فصلی زاهدان با تاخیر زمانی ۲ و ۴ و ۸ در مجموعه آموزش

در ایستگاه زاهدان به ازای تاخیر زمانی ۲ و ۴ و ۸ به عنوان ورودی الگوریتم جنگل تصادفی در مجموعه آموزش ضریب تبیین ۰/۸۷ حاصل شده است که نشان دهنده دقت مطلوب الگوریتم جنگل تصادفی در پیش بینی SPI می باشد.



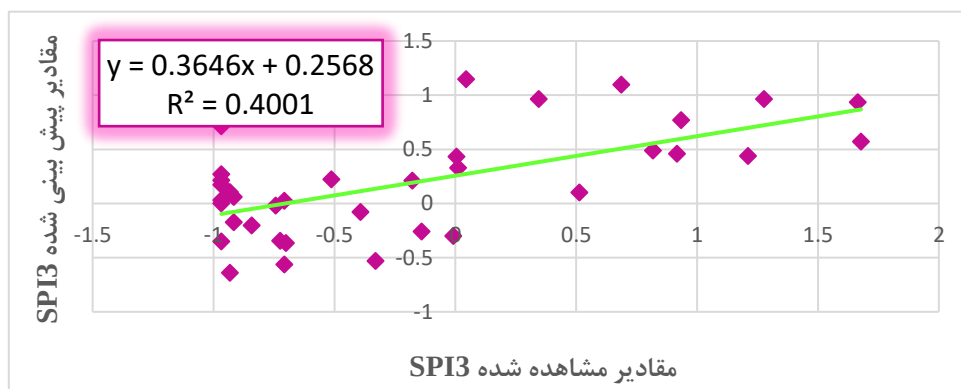
شکل ۴-۱۹ ضریب تبیین SPI فصلی زاهدان با تاخیر زمانی ۲ و ۴ و ۸ در مجموعه آزمایش

در ایستگاه زاهدان به ازای تاخیر زمانی ۲ و ۴ و ۸ به عنوان ورودی الگوریتم جنگل تصادفی در مجموعه آزمایش ضریب تبیین ۰/۴۹ حاصل شده است که نشان دهنده دقت مطلوب الگوریتم جنگل تصادفی در پیش بینی SPI می باشد.



شکل ۴-۲۰ ضریب تبیین SPI فصلی زاهدان با تاخیر زمانی ۲ و ۴ در مجموعه آموزش

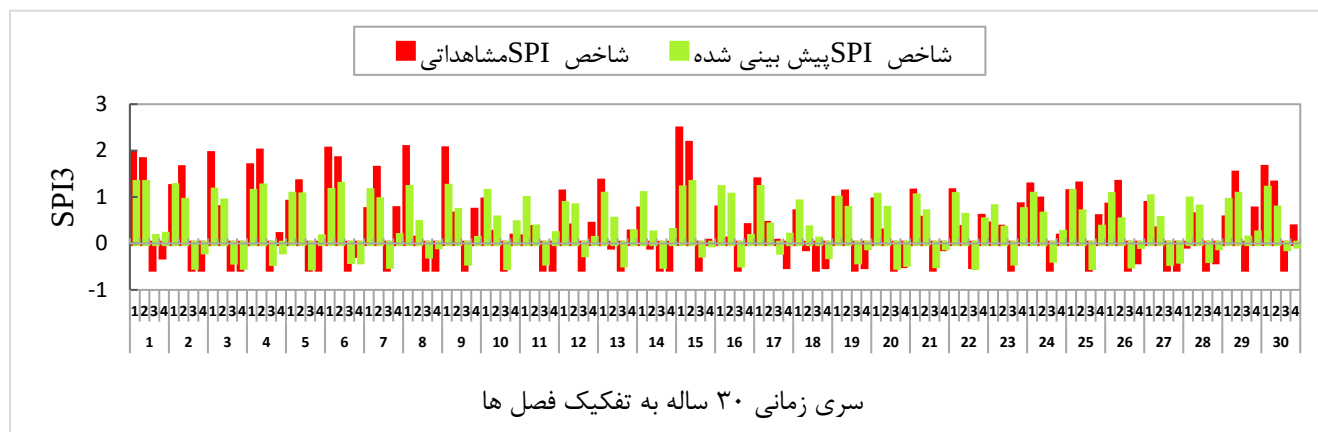
در ایستگاه زاهدان به ازای تاخیر زمانی ۲ و ۴ به عنوان ورودی الگوریتم جنگل تصادفی در مجموعه آموزش ضریب تبیین ۰/۸۴ حاصل شده است که نشان دهنده کارایی مناسب الگوریتم جنگل تصادفی در پیش بینی SPI می باشد.



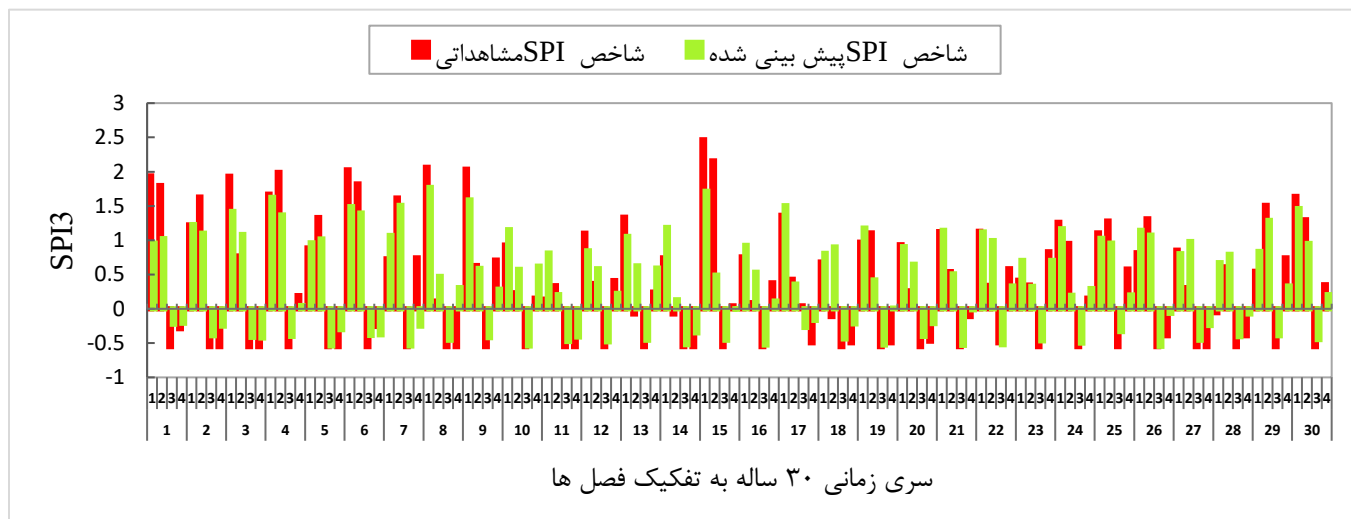
شکل ۴-۲۱ ضریب تبیین SPI فصلی زاهدان با تاخیر زمانی ۲ و ۴ در مجموعه آزمایش

در ایستگاه زاهدان به ازای تاخیر زمانی ۲ و ۴ به عنوان ورودی الگوریتم جنگل تصادفی در مجموعه آزمایش ضریب تبیین ۰/۴ حاصل شده است که نشان دهنده کارایی مناسب الگوریتم جنگل تصادفی در پیش بینی SPI می باشد.

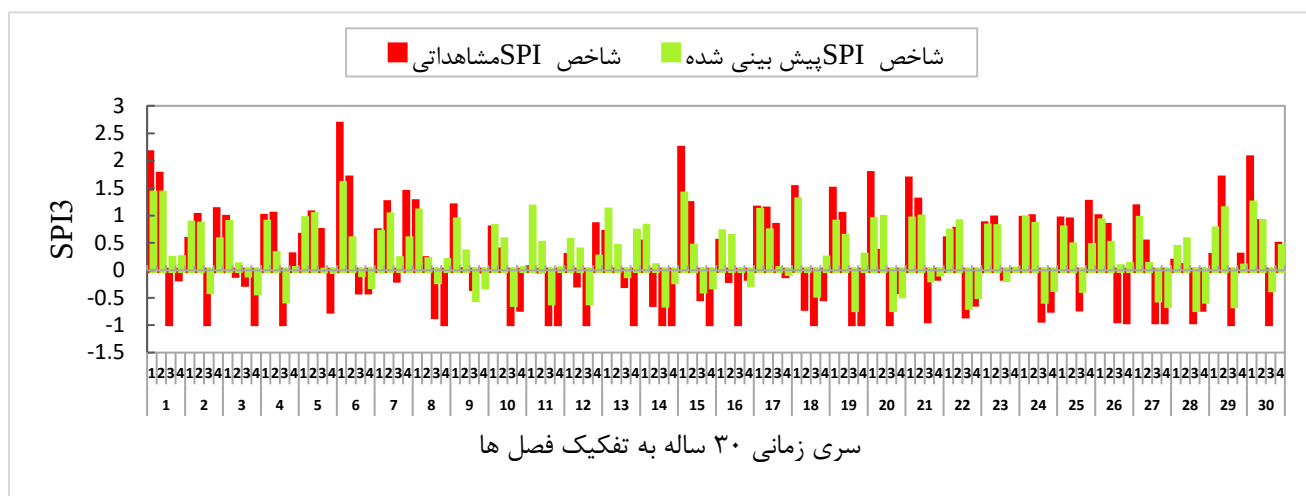
مقادیر مشاهداتی و پیش بینی شده SPI فصلی به صورت سری زمانی در قالب شکل های ۴-۲۲ تا ۴-۲۵ نمایش داده شده است که فصل های زمستان، بهار، تابستان و پاییز در نمودار ها با اعداد ۱، ۲، ۳ و ۴ مشخص شده اند.



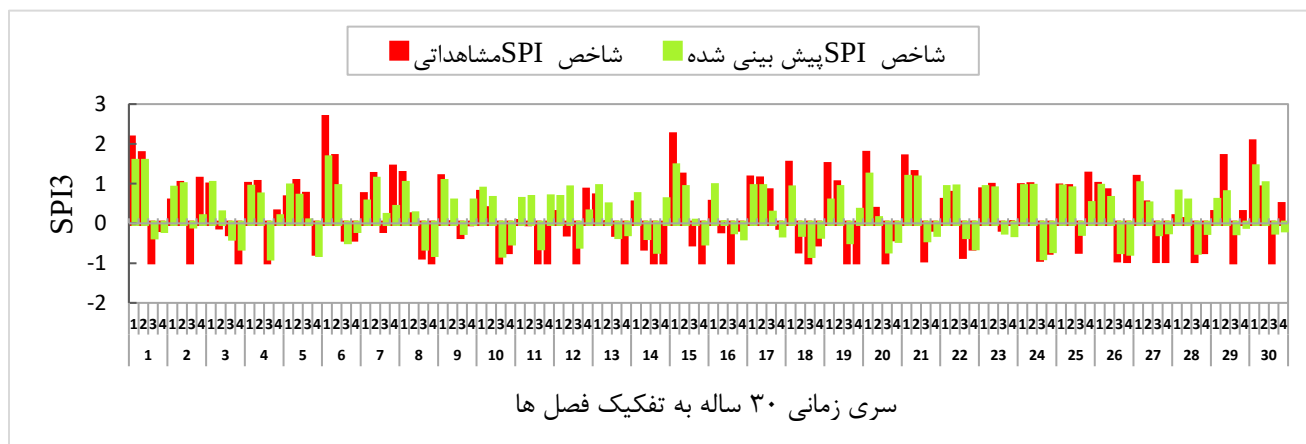
شکل ۴-۲۲ نمودار مقادیر مشاهداتی و پیش بینی شده SPI فصلی زابل با تاخیر زمانی ۲ و ۴و۶



شکل ۴-۲۳ نمودار مقادیر مشاهداتی و پیش بینی شده SPI فصلی زابل با تاخیر زمانی ۲ و ۴و۶ و ۱۲



شکل ۴-۲۴ نمودار مقادیر مشاهده‌ای و پیش‌بینی شده SPI فصلی زاهدان با تاخیر زمانی ۲ و ۴ و ۶



شکل ۴-۲۵ نمودار مقادیر مشاهده‌ای و پیش‌بینی شده SPI فصلی زاهدان با تاخیر زمانی ۲ و ۴ و ۸ و ۶

۸-۴- اعتبار سنجی الگوریتم RF

به منظور بررسی اعتبار و دقت الگوریتم جنگل تصادفی در پیش بینی SPI فصلی، پارامترهای آماری R^2 ، MSE، RMSE، NRMSE، MAE و IOA در مورد ایستگاه زابل و زاهدان به تفکیک مجموعه آموزشی و آزمایشی میان مقادیر مشاهداتی SPI و پیش بینی شده SPI به کار گرفته شده است که نتایج آن در جدول ۱۹-۵ و ۲۰-۵ نمایش داده شده است.

جدول ۱۹-۴ اعتبارسنجی پیش بینی SPI فصلی ایستگاه زابل

ایستگاه	ورودی	مجموعه مورد	پارامتر آماری	مقدار پارامتر
زابل	تاخیر زمانی 12و6و4و2	آموزش	بررسی	آماري
			R^2	0.8943
			MSE	0.0828
			RMSE	0.2878
			NRMSE	9.5569
			MAE	0.2164
			IOA	0.9648
زابل	تاخیر زمانی 12و6و4و2	آزمایش	R^2	0.5938
			MSE	0.2949
			RMSE	0.5431
			NRMSE	20.0915
			MAE	0.4175
			IOA	0.8556
زابل	تاخیر زمانی 6و4و2	آموزش	R^2	0.7942
			MSE	0.1526
			RMSE	0.3907
			NRMSE	14.4526
			MAE	0.3028
			IOA	0.9256
زابل	تاخیر زمانی 6و4و2	آزمایش	R^2	0.7164
			MSE	0.2571
			RMSE	0.507
			NRMSE	16.8347
			MAE	0.4033
			IOA	0.8673

بیشتر بودن پارامتر های R^2 و IOA و هم زمان کم تر بودن پارامتر های MSE، RMSE، NRMSE، MAE بیانگر دقت و کارایی مطلوب الگوریتم جنگل تصادفی در پیش بینی SPI می باشد. در مورد هر دو ایستگاه و به ازای مقادیر ورودی متفاوت از تاخیر زمانی به عنوان ورودی الگوریتم جنگل تصادفی و در نظرگیری مجموعه آموزش و آزمایش به طور مجزا از هم، نتایج اعتبارسنجی الگوریتم جنگل تصادفی بیان شده است.

طبق جدول ۵-۱۹ پارامتر های آماری بررسی شده به ازای تاخیر زمانی ۲،۴،۶،۱۲ به عنوان ورودی الگوریتم در مجموعه آموزش دقت زیادتری را داشته است در حالی که به ازای تاخیر زمانی ۲،۴،۶ به عنوان ورودی الگوریتم در مجموعه آزمایش دقت مطلوب تری داشته است. در ایستگاه زابل پارامتر آماری IOA برای مجموعه آزمایش به ترتیب مقدار ۰/۸۶ و ۰/۸۷ را داراست که به علت نزدیک بودن به عدد یک، نشانگر دقت مطلوب الگوریتم جنگل تصادفی در پیش بینی SPI می باشد. بر اساس سایر پارامتر های آماری بررسی شده در مطالعه یعنی R^2 ، MSE، RMSE، NRMSE و MAE دقت و کارایی الگوریتم جنگل تصادفی در پیش بینی SPI بالا می باشد.

جدول ۴-۲۰ اعتبارسنجی پیش بینی SPI فصلی ایستگاه زاهدان

ایستگاه	ورودی	مجموعه مورد بررسی	پارامتر آماری	مقدار پارامتر آماري
زاهدان	تاخیر زمانی 8و64و2	آموزش	R ²	0.8741
			MSE	0.1317
			RMSE	0.363
			NRMSE	10.0057
			MAE	0.282
			IOA	0.9495
زاهدان	تاخیر زمانی 8و64و2	آزمایش	R ²	0.4919
			MSE	0.5096
			RMSE	0.7139
			NRMSE	26.9824
			MAE	0.5982
			IOA	0.7408
زاهدان	تاخیر زمانی 6و4و2	آموزش	R ²	0.8402
			MSE	0.1871
			RMSE	0.4325
			NRMSE	11.9226
			MAE	0.3433
			IOA	0.9205
زاهدان	تاخیر زمانی 6و4و2	آزمایش	R ²	0.4001
			MSE	0.5484
			RMSE	0.7406
			NRMSE	27.9918
			MAE	0.6402
			IOA	0.6303

پارامتر های آماری بررسی شده در جدول ۵-۲۰ در مجموعه آموزش و آزمایش به ازای تاخیر زمانی ۲،۴،۶،۸ به عنوان ورودی الگوریتم نسبت به تاخیر زمانی ۲،۴،۶ به عنوان ورودی الگوریتم از دقت بیشتری برخوردار است. در ایستگاه زاهدان پارامتر آماری IOA برای مجموعه آزمایش به ترتیب مقدار ۰/۷۴ و ۰/۶۳ را دارد که حاکی از دقت مناسب پیش بینی الگوریتم جنگل تصادفی می باشد. بر اساس سایر پارامتر های آماری بررسی شده در مطالعه یعنی R^2 ، MSE، RMSE، NRMSE و MAE دقت و کارایی الگوریتم جنگل تصادفی در پیش بینی SPI بالا می باشد.

۹-۴- ارزیابی اثرات زیست محیطی خشکسالی در منطقه مورد مطالعه

رطوبت نسبی کم، کم بودن بارندگی سالانه نسبت به بارندگی متوسط کشور و تغییرات دمایی بالا از ویژگی های مناطق زابل و زاهدان می باشد [۷]. منطقه زابل و زاهدان در معرض باد های ۱۲۰ روزه قرار دارند، این باد ها معمولا از اواخر بهار تا اوایل پاییز ادامه دارند. سرعت غالب باد های ۱۲۰ روزه نزدیک به ۱۰ متر بر ثانیه می باشد. که موجب افزایش نرخ تبخیر، افزایش فرسایش خاک، کاهش پوشش گیاهی می گردد. در دهه گذشته دریاچه هامون که از رودخانه هیرمند ارتفاعات کشور افغانستان سرچشمه می گیرد به دلیل سدسازی های این کشور و کاهش بارندگی خشک شده است [۸]. رسوبات مناطق مورد مطالعه عمدتا ریزدانه بوده که این ذرات با وزش ملایم باد حرکت می کنند [۶۱]. در شکل ۴-۲۶ نمایی از طوفان گرد و غبار در منطقه سیستان نمایش داده شده است.



شکل ۴-۲۶ تصویر ماهواره ای از طوفان های گرد و غبار در تاریخ ۱۳۸۸/۶/۱۴ [۶۲].

اقلیم خشک، کاهش بارندگی، وقوع خشکسالی، خشک شدن دریاچه هامون، رسوبات ریزدانه مناطق و وجود باد های ۱۲۰ روزه شرایط مناسبی برای وقوع طوفان گرد و غبار فراهم می‌آورد.

طوفان های گرد و غبار در اثر حرکت شن های روان، تاسیسات زیر بنایی از جمله جاده ها و راه آهن را مسدود می‌کنند در نتیجه انسداد راه ها موجب اختلال در حمل و نقل و ایجاد ترافیک می‌گردد. در اثر طوفان های گرد و غبار حجم زیادی ماسه وارد کانال های آبرسانی می‌شود که باید لایروبی شوند. همچنین مخازن آبی چاه نیمه تحت خطر قرار می‌گیرند و تامین آب شرب دچار اختلال خواهد شد. وقوع طوفان های گرد و غبار باعث کاهش دید افقی و کاهش میدان دید خواهد شد و منجر به لغو شدن پرواز هواپیما ها می‌گردد. در بعضی از مناطق روستایی حجم رسوبات ماسه به گونه ای است که باعث دفن شدن واحد های مسکونی و تخلیه روستا می‌شوند [۶۱]. طوفان های گرد و غبار موجب به خطر افتادن سلامتی انسان، ایجاد بیماری چشمی و ریوی، آثار نامطلوبی در زمینه بهداشت محیط و سوانح رانندگی می‌شود [۶۲]. در ادامه اثرات خشکسالی در جنبه های کیفیت هوا و عوامل اقلیمی، خاک، زمین شناسی و ژئومورفولوژی، آب و بوم شناسی بیان شده است.

۴-۹-۱- کیفیت هوا و عوامل اقلیمی

خشکسالی موجب افزایش نرخ تبخیر می‌شود همچنین آئروسول های معدنی ایجاد شده از طوفان های گرد و غبار بر روی تشکیل ابر و کمیت نزولات جوی اثر می‌گذارد. طوفان های گرد و غبار باعث تغییر در شرایط اقلیمی می‌شوند [۹]. طوفان های گرد و غبار موجب کاهش میزان فتولیز (تجزیه شیمیایی بر اثر نیرو تابشی) می‌گردد [۱۰]. وقوع طوفان های گرد و غبار باعث کاهش دید افقی و کاهش میدان دید خواهد شد [۶۱]. در زمان وقوع طوفان گرد و غبار، غلظت فلزات سنگین و سمی به مقدار زیادی افزایش می‌یابد و آلودگی هوا ایجاد می‌شود [۶۲].

۴-۹-۲- خاک، زمین شناسی و ژئومورفولوژی

در شرایط خشکسالی رطوبت لایه سطحی خاک کم شده و مواد آلی و مغذی خاک از بین می‌رود [۱۰]. همچنین فرسایش خاک افزایش یافته و پوشش گیاهی کاهش می‌یابد [۸]. طوفان های گرد و غبار، شرایط زمین شناسی را تغییر می‌دهند [۹].

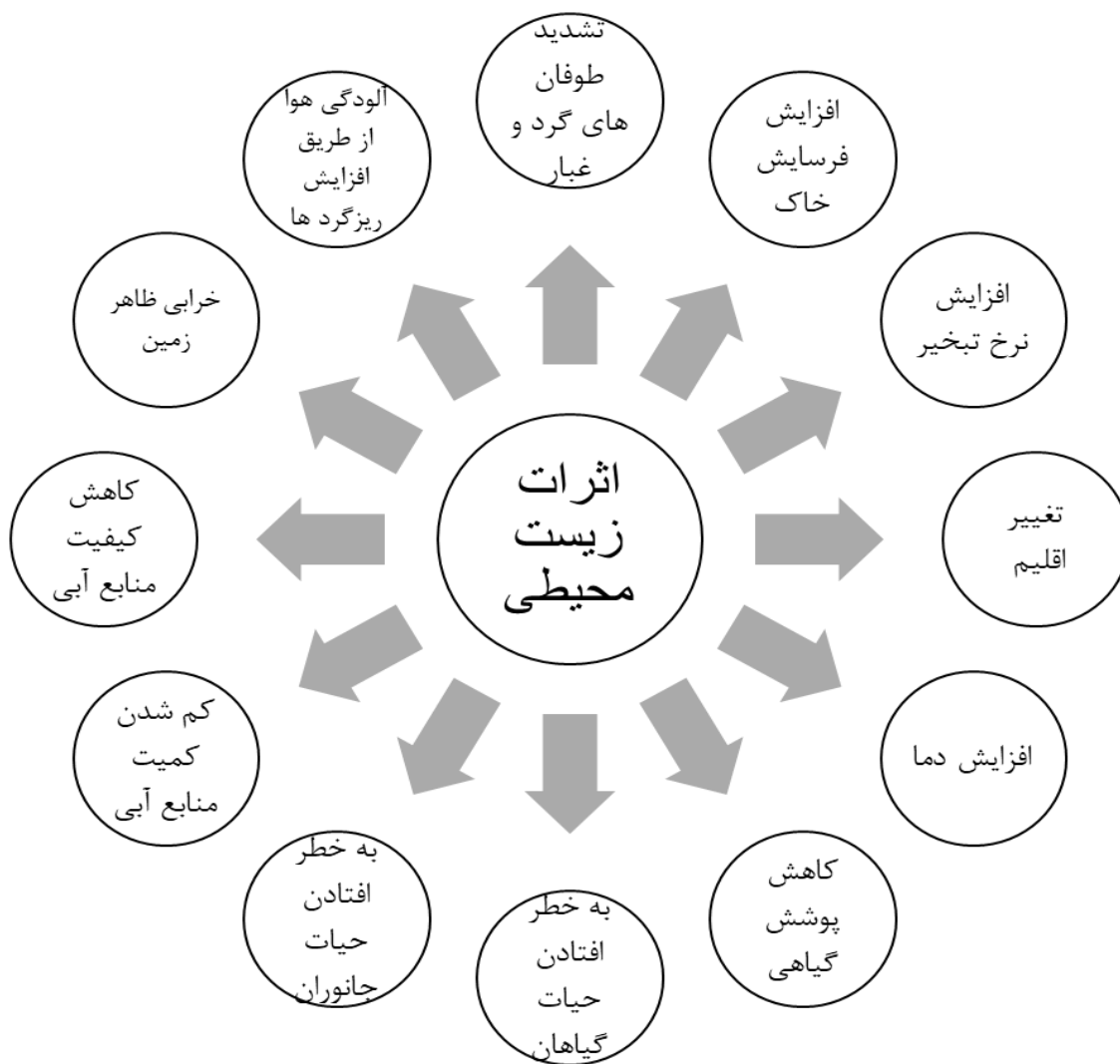
۴-۹-۳- آب

یکی از مولفه های خشکسالی کاهش رواناب است [۴]. کاهش رواناب اثرات نامطلوبی روی کمیت و کیفیت منابع آب سطحی و زیر زمینی بر جای می‌گذارد [۵].

۴-۹-۴- بوم شناسی

در شرایط خشکسالی حیات گیاهان و جانوران و نیز تنوع زیستی آن ها دچار خطر می‌شود [۵]. طوفان های گرد و غبار منجر به ایجاد تغییر در چرخه بیولوژیکی می‌شوند [۹]. در شکل ۴-۲۷ اثرات زیست محیطی خشکسالی در محدوده مطالعاتی بیان شده است.

به منظور تحلیل دقیق طوفان های گرد و غبار لازم است پارامتر هایی چون درجه حرارت، رطوبت نسبی، بارندگی، سرعت و جهت حداکثر باد، میزان دید و تعداد روز های گرد و غبار در نظر گرفته شوند تا نتایج حقیقی آن به دست آید [۶۳]. در مورد سایر جنبه های محیط زیست شامل خاک و زمین شناسی، آب و بوم شناسی نیاز است پارامتر های مختلفی بررسی شوند لذا در این تحقیق به ارزیابی کیفی اثرات زیست محیطی خشکسالی در منطقه مورد مطالعه بسنده شده است.



شکل ۴-۲۷ اثرات زیست محیطی خشکسالی در محدوده مورد مطالعه

فصل پنجم: نتیجه‌گیری و پیشنهادات

۵-۱- مقدمه

خشکسالی اثرات نامطلوب جبران ناپذیری را در جنبه های اجتماعی، اقتصادی و محیط زیستی ایجاد می کند. علت بررسی ایستگاه زابل و زاهدان ویژگی هایی همچون رطوبت نسبی کم، کم بودن بارندگی سالانه نسبت به بارندگی متوسط کشور و تغییرات دمایی بالا، وجود باد های ۱۲۰ روزه و اثرات نامطلوب این باد ها روی نرخ تبخیر، فرسایش خاک، پوشش گیاهی، طوفان های گرد و غبار و ایجاد مشکلات تنفسی برای افراد می باشد. از طرفی با پایش خشکسالی و تحلیل آن، می توان پیامد های نامطلوب خشکسالی را کاهش داد.

۵-۲- نتیجه گیری

طبقات خشکسالی، شدت خشکسالی و شدیدترین خشکسالی در طول سری زمانی ۳۰ ساله در ایستگاه زابل و زاهدان با محاسبه SPI در مقیاس فصلی و سالانه، تعیین گشته است. طبقه خشکسالی تقریباً نرمال در هر دو ایستگاه و هر دو مقیاس فصلی و سالانه در طی سری زمانی بیشترین سهم را به خود اختصاص داده است. شدت خشکسالی در ایستگاه زاهدان در مقیاس فصلی و سالانه نسبت به ایستگاه زابل زیادتر می باشد و اوج خشکسالی مربوط به SPI سالانه هر دو ایستگاه در سال ۲۰۰۱ رخ داده است، همچنین اوج خشکسالی مربوط به SPI فصلی هر دو ایستگاه در فصول فاقد بارش رخ داده است. بر اساس مطالب بیان شده، SPI پتانسیل خوبی در تعیین طبقات خشکسالی، بررسی روند تغییرات SPI و معین نمودن شدت خشکسالی دارد. ترکیبی از تاخیر زمانی های معنادار حاصل شده از تحلیل PACF به عنوان ورودی های الگوریتم جنگل تصادفی استفاده شده است و SPI در این الگوریتم پیش بینی شده است و مقادیر مشاهداتی SPI و مقادیر پیش بینی شده SPI در دو ایستگاه به صورت سری زمانی نمایش داده شدند. در مرحله اعتبارسنجی دقت الگوریتم جنگل تصادفی، پارامتر های آماری R^2 ، MSE، RMSE، NRMSE، MAE و IOA در مورد مجموعه آموزشی و آزمایشی دقت بالا و مناسب الگوریتم جنگل تصادفی را در پیش بینی SPI نشان می دهند.

طبق نتایج صحت سنجی الگوریتم جنگل تصادفی در ایستگاه زابل در مورد مجموعه آزمایش، تاخیر زمانی های ۲،۴،۶ به عنوان ورودی الگوریتم، بالاترین میزان دقت را داشته است ($R^2 = 0.72$ ، $MSE = 0.26$ ، $RMSE = 0.51$ ، $NRMSE = 16/83$ ، $MAE = 0.40$ و $IOA = 0.87$) و همچنین در ایستگاه زاهدان با در نظرگیری مجموعه آزمایش، تاخیر زمانی های ۲،۴،۶،۸ به عنوان ورودی الگوریتم، دقت بیشتری داشته است ($R^2 = 0.49$ ، $MSE = 0.51$ ، $RMSE = 0.71$ ، $NRMSE = 26/98$ و $MAE = 0.60$ ، $IOA = 0.74$). در نتیجه می توان بیان داشت به کارگیری تحلیل PACF در راستای پیش بینی خشکسالی می تواند مفید واقع شود.

۵-۳- مقایسه نتایج تحقیق با تحقیقات مشابه پیشین

طبری و همکاران در سال ۲۰۱۲ طی دوره آماری سال ۱۹۶۶ تا ۲۰۰۵ شاخص بارش استاندارد را در مقیاس ۱۲ ماهه، در مورد ۱۰ ایستگاه سینوپتیک واقع در شرق ایران شامل ایستگاه های زابل و زاهدان، بررسی نمودند که بر اساس مطالعه آن ها شدیدترین خشکسالی در ایستگاه های زابل و زاهدان در سال ۲۰۰۱ رخ داده است [۶۴]. طبق نتیجه تحقیق حاضر شدیدترین خشکسالی سالانه ایستگاه زابل و زاهدان نیز در سال ۲۰۰۱ اتفاق افتاده است که با نتیجه تحقیق طبری و همکاران همخوانی دارد. بهرامی و همکاران در سال ۲۰۱۹ داده های اقلیمی ۳۸ ایستگاه سینوپتیک ایران شامل ایستگاه های زابل و زاهدان را در بازه زمانی سال های ۱۹۶۷ تا ۲۰۱۴ با به کارگیری شاخص بارش استاندارد فصلی استفاده نمودند، که مطابق جدول سوم تحقیقشان، طبقه خشکسالی تقریباً نرمال در ایستگاه زابل با ۸۳/۸۵ درصد بالاترین درصد رخداد را داشته است و نیز طبقه خشکسالی تقریباً نرمال در ایستگاه زاهدان با ۸۲/۲۹ درصد بالاترین درصد رخداد را داشته است. همچنین دو ایستگاه زابل و زاهدان فاقد طبقات خشکسالی متوسط، خشکسالی شدید و خشکسالی بسیار شدید هستند [۶۵]. بر اساس نتیجه تحقیق حاضر و با توجه به جدول ۴-۱۲ حالت غالب خشکسالی ایستگاه های زابل و زاهدان در مقیاس فصلی تقریباً نرمال می باشد و به ترتیب دارای ۷۷ درصد و ۷۸ درصد می باشد از طرفی هر دو ایستگاه در مقیاس فصلی فاقد طبقات خشکسالی متوسط، خشکسالی شدید و خشکسالی بسیار شدید هستند که منطبق بر نتایج تحقیق بهرامی و همکاران می باشد.

۵-۴- پیشنهادات جهت مطالعات آتی

پیشنهادات جهت مطالعات آتی به شرح زیر می‌باشند:

- ❖ استفاده از دیگر توزیع‌های احتمال پارامتری ریاضی جهت محاسبه SPI
- ❖ به کارگیری سایر شاخص‌های خشکسالی جهت ارزیابی خشکسالی
- ❖ بهره‌گیری از سایر روش‌های یادگیری ماشین به منظور تخمین SPI
- ❖ استفاده از خروجی مدل‌های گردش عمومی جو جهت بررسی عملکردشان در تخمین شاخص‌های خشکسالی
- ❖ به کارگیری سنجنده‌ها در سنجش از دور

[۱] Chen, S. T., Kuo, C. C., & Yu, P. S. (2009). " Historical trends and variability of meteorological droughts in Taiwan". **Hydrological sciences journal**, 54(3), 430-441.

[۲] علیزاده، ا. (۱۳۸۹). "اصول هیدرولوژی کاربردی". دانشگاه امام رضا(ع).

[۳] بذرافشان، ج. ، & حجابی، س. (۱۳۹۵). " خشکسالی : روش های پایش ". دانشگاه تهران.

[۴] Şen, Z. (2015). "**Applied drought modeling, prediction, and mitigation**": Elsevier

[۵] Morris, P., & Therivel, R. (2009). "Methods of invironmental impact assessment".

[۶] Dashtpagerdi, M. M., Kousari, M. R., Vagharfard, H., Ghonchepour, D., Hosseini, M. E., & Ahani, H. (2015). "An investigation of drought magnitude trend during 1975–2005 in arid and semi-arid regions of Iran". **Environmental earth sciences**, 73(3), 1231-1244.

[۷] Jozaghi, A., Alizadeh, B., Hatami, M., Flood, I., Khorrami, M., Khodaei, N., & Ghasemi Tousi, E. (2018). "A comparative study of the AHP and TOPSIS techniques for dam site selection using GIS: a case study of Sistan and Baluchestan Province, Iran". **Geosciences**, 8(12), 494.

[۸] Moghbeli, A., Delbari, M., & Amiri, M. (2020). "Application of a standardized precipitation index for mapping drought severity in an arid climate region, southeastern Iran". **Arabian Journal of Geosciences**, 13(5), 1-16.

[۹] Wang, Y., Zhang, X., Arimoto, R., Cao, J., & Shen, Z. (2005). "Characteristics of carbonate content and carbon and oxygen isotopic composition of northern China soil and dust aerosol and its application to tracing dust sources". **Atmospheric Environment**, 39(14), 2631-2642.

- [١٠] Xuan, J., Sokolik, I. N., Hao, J., Guo, F., Mao, H., & Yang, G. (2004). "Identification and characterization of sources of atmospheric mineral dust in East Asia". **Atmospheric Environment**, 38(36), 6239-6252.
- [١١] Wilhite, D. A., & Glantz, M. H. (1985). "Understanding: the drought phenomenon: the role of definitions". **Water international**, 10(3), 111-120.
- [١٢] Shin, J. Y., Chen, S., Lee, J.-H., & Kim, T.-W. (2018). "Investigation of drought propagation in South Korea using drought index and conditional probability". **Terrestrial, Atmospheric & Oceanic Sciences**, 29(2).
- [١٣] Sharma, A., Dadhwal, V., Jeganathan, C., & Tolpekin, V. (2009). "Drought monitoring using standardized precipitation index: a case study for the state of Karnataka, India". **Geospatial Application Papers: Environment-Forest**.
- [١٤] Heim Jr, R. R. (2002). "A review of twentieth-century drought indices used in the United States". **Bulletin of the American Meteorological Society**, 83(8), 1149-1166.
- [١٥] Jain, V. K., Pandey, R. P., Jain, M. K., & Byun, H.-R. (2015). "Comparison of drought indices for appraisal of drought characteristics in the Ken River Basin". **Weather and Climate Extremes**, 8, 1-11.
- [١٦] Choi, M., Jacobs, J. M., Anderson, M. C., & Bosch, D. D. (2013). "Evaluation of drought indices via remotely sensed data with hydrological variables". **Journal of Hydrology**, 476, 265-273.
- [١٧] Sobral, B. S., de Oliveira-Júnior, J. F., de Gois, G., Pereira-Júnior, E. R., de Bodas Terassi, P. M., Muniz-Júnior, J. G. R., . . . Zeri, M. (2019). "Drought characterization for the state of Rio de Janeiro based on the annual SPI index: trends, statistical tests and its relation with ENSO". **Atmospheric research**, 220, 141-154.

- [١٨] Byakatonda, J., Parida, B., Moalafhi, D., & Kenabatho, P. K. (2018). " Analysis of long term drought severity characteristics and trends across semiarid Botswana using two drought indices". **Atmospheric research**, 213, 492-508.
- [١٩] Lowe, B., & Kulkarni, A. (2015). " Multispectral image analysis using random forest".
- [٢٠] Belgiu, M., & Drăguț, L. (2016). "Random forest in remote sensing: A review of applications and future directions". **ISPRS journal of photogrammetry and remote sensing**, 114, 24-31.
- [٢١] Everingham, Y., Sexton, J., Skocaj, D., & Inman-Bamber, G. (2016). "Accurate prediction of sugarcane yield using a random forest algorithm". **Agronomy for sustainable development**, 36(2), 27.
- [٢٢] Rahmati, O., Pourghasemi, H. R., & Melesse, A. M. (2016). " Application of GIS-based data driven random forest and maximum entropy models for groundwater potential mapping: a case study at Mehran Region, Iran". **Catena**, 137, 360-372.
- [٢٣] Kulkarni, A. D., & Lowe, B. (2016). "Random forest algorithm for land cover classification".
- [٢٤] Chen, W., Xie, X., Wang, J., Pradhan, B., Hong, H., Bui, D. T., . . . Ma, J. (2017). " A comparative study of logistic model tree, random forest, and classification and regression tree models for spatial prediction of landslide susceptibility". **Catena**, 151, 147-160.
- [٢٥] Sylvester, E. V., Bentzen, P., Bradbury, I. R., Clément, M., Pearce, J., Horne, J., & Beiko, R. G. (2018). "Applications of random forest feature selection for fine-scale genetic population assignment". **Evolutionary applications**, 11(2), 153-165.

- [۲۶] Ilia, I., Loupasakis, C., & Tsangaratos, P. (2018). "Land subsidence phenomena investigated by spatiotemporal analysis of groundwater resources, remote sensing techniques, and random forest method: the case of Western Thessaly, Greece". **Environmental monitoring and assessment**, 190(11), 1-19.
- [۲۷] Norouzi, H., & Moghaddam, A. A. (2020). "Groundwater quality assessment using random forest method based on groundwater quality indices (case study: Miandoab plain aquifer, NW of Iran)". **Arabian Journal of Geosciences**, 13(18), 1-13.
- [۲۸] Saha, S., Saha, M., Mukherjee, K., Arabameri, A., Ngo, P. T. T., & Paul, G. C. (2020). "Predicting the deforestation probability using the binary logistic regression, random forest, ensemble rotational forest, REPTree: A case study at the Gumani River Basin, India". **Science of The Total Environment**, 730, 139197.
- [۲۹] Moradi, H., Rajabi, M., & Faragzadeh, M. (2011). "Investigation of meteorological drought characteristics in Fars province, Iran". **Catena**, 84(1-2), 35-46.
- [۳۰] Kousari, M. R., Dastorani, M. T., Niazi, Y., Soheili, E., Hayatzadeh, M., & Chezgi, J. (2014). "Trend detection of drought in arid and semi-arid regions of Iran based on implementation of reconnaissance drought index (RDI) and application of non-parametrical statistical method". **Water resources management**, 28(7), 1857-1872.
- [۳۱] Huang, S., Huang, Q., Chang, J., Zhu, Y., Leng, G., & Xing, L. (2015). "Drought structure based on a nonparametric multivariate standardized drought index across the Yellow River basin, China". **Journal of Hydrology**, 530, 127-136.
- [۳۲] Pramudya, Y., & Onishi, T. (2018). "Assessment of the standardized precipitation index (SPI) in Tegal City, Central Java, Indonesia". **Paper presented at the IOP Conference Series: Earth and Environmental Science**.

- [33] Kamruzzaman, M., Cho, J., Jang, M.-W., & Hwang, S. (2019). "Comparative evaluation of standardized precipitation index (SPI) and effective drought index (EDI) for meteorological drought detection over Bangladesh". **한국농공학회논문집**, 61(1), 145.
- [34] Li, L., She, D., Zheng, H., Lin, P., & Yang, Z.-L. (2020). "Elucidating diverse drought characteristics from two meteorological drought indices (SPI and SPEI) in China". **Journal of Hydrometeorology**, 21(7), 1513-1530.
- [35] Chen, J., Li, M., & Wang, W. (2012). "Statistical uncertainty estimation using random forests and its application to drought forecast". **Mathematical Problems in Engineering**, 2012.
- [36] Park, S., Im, J., Jang, E., & Rhee, J. (2016). "Drought assessment and monitoring through blending of multi-sensor indices using machine learning approaches for different climate regions". **Agricultural and forest meteorology**, 216, 157-169.
- [37] Rhee, J., & Im, J. (2017). "Meteorological drought forecasting for ungauged areas based on machine learning: Using long-range climate forecast and remote sensing data". **Agricultural and forest meteorology**, 237, 105-122.
- [38] Park, H., & Kim, K. (2019). "Prediction of severe drought area based on random forest: Using satellite image and topography data". **Water**, 11(4), 705.
- [39] Dikshit, A., Pradhan, B., & Alamri, A. M. (2020). "Short-term spatio-temporal drought forecasting using random forests model at New South Wales, Australia". **Applied Sciences**, 10(12), 4254.
- [40] Mehr, A. D., Tur, R., Çalışkan, C., & Tas, E. (2020). "A Novel Fuzzy Random Forest Model for Meteorological Drought Classification and Prediction in Ungauged Catchments". **Pure and Applied Geophysics**, 177(12), 5993-6006.

- [۴۱] Rahmati, O., Falah, F., Dayal, K. S., Deo, R. C., Mohammadi, F., Biggs, T., . . . Bui, D. T. (2020). "Machine learning approaches for spatial modeling of agricultural droughts in the south-east region of Queensland Australia". **Science of The Total Environment**, 699, 134230.
- [۴۲] Khan, N., Sachindra, D., Shahid, S., Ahmed, K., Shiru, M. S., & Nawaz, N. (2020). "Prediction of droughts over Pakistan using machine learning algorithms". **Advances in Water Resources**, 139, 103562.
- [۴۳] Shamshirband, S., Hashemi, S., Salimi, H., Samadianfard, S., Asadi, E., Shadkani, S., . . . Chau, K.-W. (2020). "Predicting standardized streamflow index for hydrological drought using machine learning models". **Engineering Applications of Computational Fluid Mechanics**, 14(1), 339-350.
- [۴۴] Xu, D., Zhang, Q., Ding, Y., & Huang, H. (2020). "Application of a Hybrid ARIMA–SVR Model Based on the SPI for the Forecast of Drought—A Case Study in Henan Province, China". **Journal of Applied Meteorology and Climatology**, 59(7), 1239-1259.
- [۴۵] Arsanjani, J. J., Tayyebi, A., & Vaz, E. (2016). "GlobeLand30 as an alternative fine-scale global land cover map: Challenges, possibilities, and implications for developing countries". **Habitat International**, 55, 25-31.

[۴۶] فلاح قالهري، غ.، موسوی بايگی، س.، حبیبی نوخندان، م. (۲۰۱۶). "پیش بینی بارش فصلی بر اساس الگوهای سینوپتیکی سطوح فوقانی جو با استفاده از مدل های آماری".

[۴۷] جودی، ر.، ستاری، ت. (۲۰۱۶). "ارزیابی عملکرد روش های مختلف در بازسازی داده های بارش ماهانه". نشریه تحقیقات کاربردی علوم جغرافیایی، ۱۶ (۴۲)، ۱۵۵.

- [۴۸] McKee, T. B., Doesken, N. J., & Kleist, J. (1993). "The relationship of drought frequency and duration to time scales". **Paper presented at the Proceedings of the 8th Conference on Applied Climatology.**
- [۴۹] Hayes, M. J., Svoboda, M. D., Wihite, D. A., & Vanyarkho, O. V. (1999). "Monitoring the 1996 drought using the standardized precipitation index". **Bulletin of the American Meteorological Society, 80(3), 429-438.**
- [۵۰] Cancelliere, A., Di Mauro, G., Bonaccorso, B., & Rossi, G. (2007). "Drought forecasting using the standardized precipitation index". **Water resources management, 21(5), 801-819.**
- [۵۱] Siqueira, H., Boccato, L., Luna, I., Attux, R., & Lyra, C. (2018). "Performance analysis of unorganized machines in streamflow forecasting of Brazilian plants". **Applied Soft Computing, 68, 494-506.**
- [۵۲] Tinungki, G. (2019). "The analysis of partial autocorrelation function in predicting maximum wind speed". **Paper presented at the IOP Conference Series: Earth and Environmental Science.**
- [۵۳] Barboza, F., Kimura, H., & Altman, E. (2017). "Machine learning models and bankruptcy prediction". **Expert Systems with Applications, 83, 405-417.**
- [۵۴] Mosavi, A., Ozturk, P., & Chau, K.-w. (2018). "Flood prediction using machine learning models: Literature review". **Water, 10(11), 1536.**
- [۵۵] Liakos, K. G., Busato, P., Moshou, D., Pearson, S., & Bochtis, D. (2018). "Machine learning in agriculture: A review". **Sensors, 18(8), 2674.**

[۵۶] بندری، ف. (۱۳۹۴). پایان نامه کارشناسی ارشد " مطالعات QSAR مشتقات کوئینولیزیدین با استفاده از الگوریتم جنگل تصادفی". دانشکده شیمی، دانشگاه دامغان.

[۵۷] Breiman, L. (2001), "Random forests", **Machine Learning**, 45(1), 5–32.

[۵۸] Breiman, L. (2003), "Random Forests Manual v4.0", **Technical report**, UC Berkeley, [ftp://ftp.stat.berkeley.edu/pub/users/breiman/Using random forests v4.0.pdf](ftp://ftp.stat.berkeley.edu/pub/users/breiman/Using_random_forests_v4.0.pdf).

[۵۹] Cutler, D.R., Edwards, T.C., Beard, K.H., Cutler, A., Hess, K.T., Gibson, J., Lawler, J.J., (2007). "Random forests for classification in ecology". **Ecology** 88 (11), 2783 –2792

[۶۰] Iqbal, M. M., Shoaib, M., Farid, H. U., & Lee, J. L. (2018). "Assessment of water quality profile using numerical modeling approach in major climate classes of Asia". **International journal of environmental research and public health**, 15(10), 2258.

[۶۱] کریمی، م.، یزدانی، م. نادری، ا. (۱۳۹۰). "تاثیر باد های ۱۲۰ روزه بر امنیت منطقه سیستان". **جغرافیا و برنامه ریزی محیطی**

[۶۲] توکلی، م.، رضایی کیخا، ع. (۲۰۱۶). "اثرات باد های ۱۲۰ روزه بر محیط زیست سیستان".

[۶۳] مهربانی، ش.، سلطانی، س.، جعفری، ر. (۱۳۹۳). "بررسی رابطه بین پارامترهای اقلیمی و وقوع ریزگردها (مطالعه موردی: استان خوزستان)". **مجله علوم و فنون کشاورزی و منابع طبیعی، علوم آب و خاک**، شماره ۷۱، صفحات ۶۹–۸۱

[۶۴] Tabari, H., Abghari, H., & Hosseinzadeh Talaei, P. (2012). "Temporal trends and spatial characteristics of drought and rainfall in arid and semiarid regions of Iran". **Hydrological Processes**, 26(22), 3351-3361.

[۶۵] Bahrami, M., Bazrkar, S., & Zarei, A. R. (2019). "Modeling, prediction and trend assessment of drought in Iran using standardized precipitation index". **Journal of Water and Climate Change**, 10(1), 181-196.

Abstract

Drought intensifies the water crisis and causes irreparable damage to communities. In recent years, machine learning methods have been considered by researchers for drought assessment. The goal of this study is drought assessment in Zabol and Zahedan cities during (1990-2020). In this regard, standardized precipitation index (SPI) at seasonal and annual time scales, partial autocorrelation analysis (PACF), and random forest algorithm (RF) were employed. After SPI calculation, the results of PACF analysis for SPI are used as inputs of the model. Training and testing data were verified with two different inputs. According to the results on a seasonal and annual scale of SPI, the normal drought category (N) has almost the highest frequency in both stations and drought intensity in Zahedan station was higher on seasonal and annual scale than Zabol station and the peak of annual drought in both stations occurred in 2001. Also, the peak of drought in both stations on seasonal scale occurred in seasons without rainfall. According to the PACF analysis, the study areas have undergone climate change over 30 years. The results of the model were evaluated by several statistical parameters. The index of agreement (IOA) of Zabol station for training data by considering four time-lags (2, 4, 6, and 12 months) and three time-lags (2, 4, and 6 months) as input, were 0.9648 and 0.9256, respectively and for testing data was approximately 0.8556 and 0.8673, respectively. IOA at Zahedan station for training data with four time-lags (2, 4, 6, and 8 months) and three time-lags (2, 4, and 6 months) were 0.9495 and 0.9205, respectively, and for testing data were 0.7408 and 0.6303, respectively. Other statistical parameters also indicate the permissible accuracy of the RF model in SPI estimation.

Keywords: Drought, Standardized precipitation index, Random forest algorithm, Partial autocorrelation analysis



Shahrood University of Technology

Faculty of Civil Engineering

M.Sc. Thesis in Civil and Environmental engineering

Application of random forest algorithm and standardized precipitation index for drought forecasting and its environmental impact assessment

By: Zeinab Torabi

Supervisor:

Dr. Ramazan Vagheei

Advisor:

Alireza Ghaemi

December 2021