

## ارزیابی شاخص‌های سنجش از دوری پایش خشکسالی با استفاده از رگرسیون بردار پشتیبان و مدل‌های جنگل تصادفی (مطالعه موردی: شهرستان مریوان)

جمال سیدی قلدره<sup>۱</sup>، سلمان احمدی<sup>۲\*</sup>، مهدی غلام‌نیا<sup>۳</sup>

۱- دانشجوی کارشناسی ارشد سنجش از دور، دانشکده مهندسی، دانشگاه کردستان

۲- استادیار گروه مهندسی عمران، دانشکده مهندسی، دانشگاه کردستان

۳- استادیار گروه مهندسی عمران، دانشکده مهندسی، دانشگاه آزاد واحد سنندج

تاریخ دریافت مقاله: ۱۴۰۱/۰۶/۱۶ تاریخ پذیرش مقاله: ۱۴۰۱/۱۰/۲۷

### چکیده

خشکسالی به عنوان پدیده‌ای طبیعی و اقلیمی همه ساله در مناطق وسیعی در سراسر جهان رخ می‌دهد و وقوع آن ناشی از کمبود بارندگی و افزایش تبخیر و تعرق در دمای بالا است. هدف پژوهش حاضر ارزیابی سنجش از دور در پایش خشکسالی برای شهرستان مریوان و تجزیه و تحلیل توزیع مکانی - زمانی شرایط خشکسالی و شناسایی شدت خشکسالی است. در این مطالعه با استفاده از شاخص‌های مختلف خشکسالی تولید شده از داده‌های ماهواره‌ای مادیس و *TRMM* که از بسترگوگل ارث انجین استخراج شده‌اند، شرایط خشکسالی در شهرستان مریوان از ماه‌های فوریه تا نوامبر برای سال‌های ۲۰۰۱ تا ۲۰۱۷ مورد تجزیه و تحلیل قرار گرفته است. همچنین شاخص‌های سنجش از دوری مانند شاخص تفاضل نرمال شده پوشش گیاهی، شاخص وضعیت پوشش گیاهی، شاخص وضعیت دما، شاخص پوشش گیاهی بهبود یافته، شاخص تبخیر و تعرق و شاخص وضعیت بارندگی به عنوان متغیرهای مستقل انتخاب گردیدند. بعلاوه شاخص استاندارد بارندگی حاصل از داده‌های هواشناسی نیز به عنوان متغیر وابسته برای ارزیابی شرایط خشکسالی مورد استفاده قرار گرفت. در ادامه جهت مدلسازی شاخص خشکسالی از دو روش مرسوم یادگیری ماشین شامل جنگل تصادفی و رگرسیون بردار پشتیبان استفاده شده است. نتایج مدل‌سازی با استفاده از الگوریتم رگرسیون بردار پشتیبان نشان‌دهنده ضریب تبیین ۰/۸۸ و مجذور میانگین مربعات خطا ۰/۳۱۳ بین شاخص‌های خشکسالی حاصل از داده‌های ماهواره‌ای و نیز اطلاعات ایستگاه‌های زمینی، می‌باشد. همچنین نتایج مدل جنگل تصادفی با مقادیر ضریب تبیین ۰/۹۰۹ و مجذور میانگین مربعات خطا ۰/۲۵۹ بیانگر کارایی بالای این مدل است. در نهایت نتایج تحقیق نشان می‌دهد که شاخص‌های سنجش از دوری *PCI*، *ET*، *EVI* و *NDVI* به ترتیب بیش‌ترین همبستگی را با شاخص هواشناسی داشته‌اند.

کلیدواژه‌ها: خشکسالی، تصاویر سنجش از دور، جنگل تصادفی، رگرسیون بردار پشتیبان.

\* نویسنده مکاتبه کننده: گروه عمران، دانشکده مهندسی، دانشگاه کردستان، سنندج، ایران

تلفن: ۰۹۱۲۶۷۲۲۱۷۶۹

## ۱- مقدمه

خشکسالی به عنوان یک پدیده پیچیده وابسته به متغیرهای محیطی متنوعی است و لذا نیاز به بررسی بیشتر در نواحی جغرافیایی مختلف دارد. خشکسالی در اثر کاهش میزان بارندگی در یک دوره زمانی مشخص نسبت به دوره‌های زمانی مشابه می‌باشد که شامل خشکسالی کشاورزی، هیدرولوژیکی و هواشناسی است. در این مطالعه به بحث خشکسالی کشاورزی با استفاده از داده‌های سنجش از دوری و مقایسه نتایج آن با نتایج خشکسالی هواشناسی مستخرج از ایستگاه هواشناسی مریوان پرداخته شده است. از آنجایی که پایش خشکسالی برای کاهش و هشدار زودرس آن بسیار مهم است، ابزارهای مختلفی از جمله شاخص‌های خشکسالی برای نظارت بر شرایط خشکسالی توسعه داده شده‌اند [۱]. با توسعه فناوری سنجش از دور، بررسی مداوم وضعیت رشد گیاهان به عنوان یکی از پارامترهای مهم در پایش خشکسالی مورد استفاده قرار می‌گیرد که این امر با محاسبه شاخص‌های مختلف پوشش گیاهی و نیز برآورد دمای تاج برگ توسط حسگرهای ماهواره‌ای انجام می‌شود [۲]. از آنجایی که مورفولوژی و سبزی پوشش گیاهی و همچنین دمای تاج پوشش برگ بسیار حساس به خشکی هستند، شاخص‌های متعدد خشکسالی مبتنی بر سنجش از دور بر اساس شاخص‌های پوشش گیاهی و دمای سطح زمین پیشنهاد شده است. هر کدام از شاخصها نیازهای مختلفی دارد و از روش‌های منحصر به فرد برای اندازه‌گیری خشکسالی استفاده می‌کند. بنابراین محصولات مبتنی بر داده‌های سنجش از دور ماهواره‌ای در حال حاضر تقریباً تمام جنبه‌های انتشار خشکسالی را نشان می‌دهد و در درک ما از پدیده‌ها سطح زمین نقش اساسی ایفا می‌کند. این امر به ویژه در مناطق با کمبود داده و ایستگاه‌های نظارت در محل بسیار مهم است [۳]. مکی و همکاران (۱۹۹۳) شاخص استاندارد

بارندگی ( $SPI^1$ ) را به عنوان شاخص خشکسالی برای پایش و تجزیه و تحلیل خشکسالی هواشناسی پیشنهاد کردند که توسط سازمان هواشناسی جهانی به عنوان یک شاخص استاندارد نظارت بر خشکسالی توصیه شده است [۴]. هیز و همکاران (۱۹۹۹) از شاخص  $SPI$  حاصل از داده‌های سنجش از دور برای بررسی خشکسالی شدید در دشت‌های جنوبی و جنوب غربی ایالات متحده استفاده نمودند که شاخص  $SPI$  روشی مفید برای تشخیص ظهور خشکسالی و نظارت بر پیشرفت آن بوده است [۵]. در مطالعه‌ای میرموسوی و کریمی (۱۳۸۸) به منظور بررسی اثر خشکسالی بر پوشش گیاهی در استان کردستان، از داده‌های هواشناسی و سنجش از دور استفاده کردند. آنها نخست شاخص بارندگی را برای ۶ ایستگاه هواشناسی استان طی سال‌های ۲۰۰۰-۲۰۰۹ محاسبه و براساس ۴۰ تصویر ۱۶ روزه برگرفته از سنجنده مادیس<sup>۲</sup> در ماه‌های آگوست، سپتامبر، اکتبر و نوامبر، شاخص تفاضل نرمال شده پوشش گیاهی ( $NDVI^3$ ) را به دست آوردند. نتایج این مطالعه نشان داد بین میانگین شاخص‌های  $SPI$  و شاخص  $NDVI$  همبستگی زیادی (۰/۷۷) در سطح معناداری (۰/۱) وجود دارد [۶]. غلام‌نیا و همکاران (۲۰۱۹) با استفاده از داده‌های ماهواره‌ای گروه خطرات اقلیمی بارش مادون قرمز با داده‌های ایستگاهی ( $CHRIPS^4$ ) و پارامترهای فنولوژی استخراج شده از شاخص  $NDVI$  در منطقه کردستان به بررسی اثر خشکسالی بر میزان سبزی‌نگی پرداخته‌اند که نتایج تحقیق آنها بیانگر اثر قابل ملاحظه خشکسالی در سال

<sup>۱</sup> Standardized Precipitation Index

<sup>۲</sup> MODIS

<sup>۳</sup> Normalized Difference Vegetation Index

<sup>۴</sup> Climate Hazards Group Infrared Precipitation with Station Data

هواشناسی بارش جهانی ( $GPCP^y$ )، بارندگی حاصل از آنالیز داده های اقلیمی ادغام شده ( $CMAP^h$ ) و  $TRMM$  از الگوریتم‌های پرسپترون چند لایه، استنباط فازی عصبی قابل تنظیم و رگرسیون بردار پشتیبان ( $SVR^g$ ) استفاده نمودند و در نهایت با شاخص  $SPI$  به دست آمده از داده‌های هواشناسی ارزیابی شد که نتایج مشابهی را کسب کردند [۱۰]. همچنین پویو فنگ و همکاران (۲۰۱۹) برای نظارت بر خشکسالی کشاورزی در جنوب شرقی استرالیا از فاکتورهای خشکسالی داده‌های سنجش از دور مانند سنجنده  $TRMM$  و مادیس استفاده نمودند. آنها برای تولید شاخص  $ETI$  پیش‌بینی شده بارش از سه روش یادگیری ماشین از جمله جنگل تصادفی، ماشین بردار پشتیبان ( $SVM^{10}$ ) و شبکه عصبی پرسپترون چند لایه به عنوان مدل رگرسیون طی سال‌های ۲۰۰۱ تا ۲۰۱۷ استفاده کردند. نتایج آنها نشان داد که مدل جنگل تصادفی از دو مدل دیگر برای پیش‌بینی شاخص  $ETI$  پیش‌بینی شده بارش بهتر عمل کرده است [۱۱]. در ادامه سان و همکاران (۲۰۲۰) در یک پژوهش دیگر سازگاری شاخص‌های خشکسالی سنجش از دور را در کشور قزاقستان تحلیل نمودند و اثربخشی این شاخص‌ها را در پایش خشکسالی کشاورزی با توجه به داده‌های عملکرد محصولات کشاورزی در یک بازه ۱۵ ساله (۲۰۰۴-۲۰۱۸) ارزیابی نمودند. نتایج نشان داد که در این منطقه در مقایسه با مدل خشکسالی سنجش از دور تک عاملی، شاخص‌های مبتنی بر شاخص پوشش گیاهی تفاضل نرمال شده و دمای سطح زمین ( $LST^{11}$ )

۲۰۰۸ بر پارامترهای اصلی فنولوژی سالانه گیاه بوده است [۷]. سونیونگ پارک و همکاران (۲۰۱۶) با استفاده از داده‌های ماهواره مادیس و داده‌های ماهواره-ای ماموریت اندازه گیری بارندگی گرمسیری ( $TRMM^1$ ) و نیز شاخص تبخیر و تعرق ( $ETI^2$ ) پیش‌بینی شده بارش از سال ۲۰۰۰ تا ۲۰۱۲ در ایالات متحده آمریکا با روش‌های یادگیری ماشین مانند جنگل‌های تصادفی<sup>۳</sup>، درختان رگرسیون بهبود یافته و الگوریتم کوبیست، اقدام به مدل‌سازی تبخیر و تعرق کردند که روش جنگل تصادفی نسبت به دو روش دیگر بهتر عمل کرده است [۸]. در ادامه پارک و همکاران (۲۰۱۷) از روش جنگل‌های تصادفی با در نظر گرفتن شاخص بارش از داده‌های ماهواره  $TRMM$ ، شاخص وضعیت پوشش گیاهی ( $VCI^4$ )، شاخص وضعیت دما ( $TCI^5$ ) از داده‌های ماهواره‌ای مادیس و سایر شاخص‌های خشکسالی سنجش از دور برای ساختن یک مدل جامع نظارت بر خشکسالی در شبه جزیره کره استفاده کردند. نتایج نشان داد که این مدل، خشکسالی هواشناسی را به ویژه در فصل خشک به خوبی نشان می‌دهد. با این حال این روش در نظارت بر خشکسالی هیدرولوژیکی با شاخص خشکسالی رطوبت خاک با قدرت تفکیک بالا ( $HSMDI^6$ ) همبستگی ضعیفی داشت [۹]. بعلاوه علیزاده و همکاران (۲۰۱۸) برای برآورد خشکسالی هواشناسی با استفاده از داده‌های سنجش از دور ماهواره‌های بارندگی شامل پروژه آب و

<sup>۱</sup> The Tropical Rainfall Measuring Mission

<sup>۲</sup> Evapotranspiration Index

<sup>۳</sup> Random Forests

<sup>۴</sup> Vegetation Condition Index

<sup>۵</sup> Temperature Condition Index

<sup>۶</sup> High Resolution Soil Moisture Drought Index

<sup>y</sup> Global Precipitation Climatology Project

<sup>h</sup> Climate Marged Analysis of Precipitation

<sup>g</sup> Support Vector Regression

<sup>10</sup> Support Vector Machine

<sup>11</sup> Land Surface Temperature

خشکسالی برای اولین بار در منطقه مریوان انجام گرفته است.

## ۲- مواد و روش‌ها

در این پژوهش با استفاده از مشاهدات سنجش از دور ماهواره‌ای و از دیدگاه‌های اقلیمی و اکوسیستم، رویکرد پایش خشکسالی بررسی شده است. مشاهدات ماهواره‌ای که در حال حاضر برای نظارت بر خشکسالی استفاده می‌شوند، فرصت‌هایی را برای توسعه مدل‌های خشکسالی ترکیبی و چند شاخص فراهم می‌کنند. بنابراین نیاز به مدل‌ها و الگوریتم‌هایی است که چندین مجموعه داده را با هم ترکیب کرده و یا مشاهدات ماهواره‌ای را در شبیه‌سازی مدل با داده‌های زمینی دخیل نموده و در نهایت بتواند خشکسالی را در یک بازه زمانی طولانی مدت پایش کند. از جمله فعالیت‌هایی که می‌توان با استفاده از تصاویر سنجش از دوری انجام داد، مطالعاتی در مورد پوشش گیاهی است که می‌توان از این طریق رشد، بیماری، رطوبت، خشکی و شادابی گیاه را مطالعه نمود. علت این توانایی این است که پوشش‌های گیاهی مختلف، حال چه زنده و چه غیره زنده در مواجهه با امواج الکترو مغناطیسی واکنش متفاوتی دارند، که همین امر سبب اختلاف پوشش گیاهی مختلف می‌شود. طی دهه‌های گذشته شاخص‌های خشکسالی سنجش از دوری متعددی پیشنهاد شده است که برخی از آنها به عنوان شاخص ضروری در سیستم‌های نظارت بر خشکسالی به طور گسترده در سراسر جهان مورد استفاده قرار گرفته است [۱۲]. در این پژوهش برای تجزیه و تحلیل و همبستگی بین شاخص‌های خشکسالی سنجش از دور و شاخص داده‌های زمینی از طریق الگوریتم‌های یادگیری ماشین مقایسه‌ای انجام می‌گیرد. منطقه مورد مطالعه این پژوهش، طبق شکل (۱)، کل سطح شهرستان مریوان است. شهرستان مریوان با طول و عرض جغرافیایی بین ۳۵ درجه و ۴۸ دقیقه تا ۳۵ درجه و ۲ دقیقه عرض

برای پایش خشکسالی کشاورزی با توجه به ویژگی‌های اقلیمی و پوشش گیاهی در مناطق خشک مؤثرتر هستند [۱۲]. همچنین مهرآور و همکاران (۲۰۲۱) در مطالعه خود یک شاخص جدید خشکسالی مبتنی بر سنجش از دور با نام شاخص خشکسالی دما-گیاه-رطوبت خاک-بارندگی ( $TVMPDI^1$ ) را پیشنهاد نمودند که شاخص ترکیبی آنها با داده‌های بارش ماهانه و دمای خاک در محل به ترتیب در مقادیر ۰٫۷۶ و ۰٫۷۵ همبستگی نشان داد [۱۳]. همچنین پرودهان و همکاران (۲۰۲۱) مطالعه‌ای با هدف پایش خشکسالی با استفاده از رویکرد یادگیری عمیق با داده‌های سنجش از دور در جنوب آسیا از سال ۲۰۰۱ تا ۲۰۱۶ انجام دادند. آنها شاخص‌های بارندگی، پوشش گیاهی و عوامل خاک را به عنوان پارامترهای ورودی شبکه عصبی پیشرو عمیق ( $DFNN^2$ ) در نظر گرفتند و خشکسالی کشاورزی را با استفاده از شاخص کمبود رطوبت خاک ( $SMDI^3$ ) به عنوان یک متغیر پاسخ طی سه مرحله فنولوژی محصول ارزیابی نمودند [۱۴].

طیف وسیعی از شاخص‌های سنجش از دوری توسط محققین مختلف توسعه داده شده است. در این پژوهش چند شاخص مشهور مستخرج از تصاویر ماهواره‌ای برای ارزیابی خشکسالی مورد مطالعه قرار گرفته است. ارزیابی و مطالعه تطبیقی شاخص‌ها براساس روش‌های مختلف مدل‌سازی انجام گردیده و در زمینه توسعه این مدل‌ها مطالعه و بحث شده است. در ادامه دو الگوریتم مرسوم جنگل‌های تصادفی و  $SVR$  به منظور مدل‌سازی شدت خشکسالی براساس داده‌های ماهواره‌ای توسعه داده شده است. همچنین استفاده از پارامترهای مختلف مستخرج از داده‌های ماهواره‌ای و ارتباط آن با شاخص

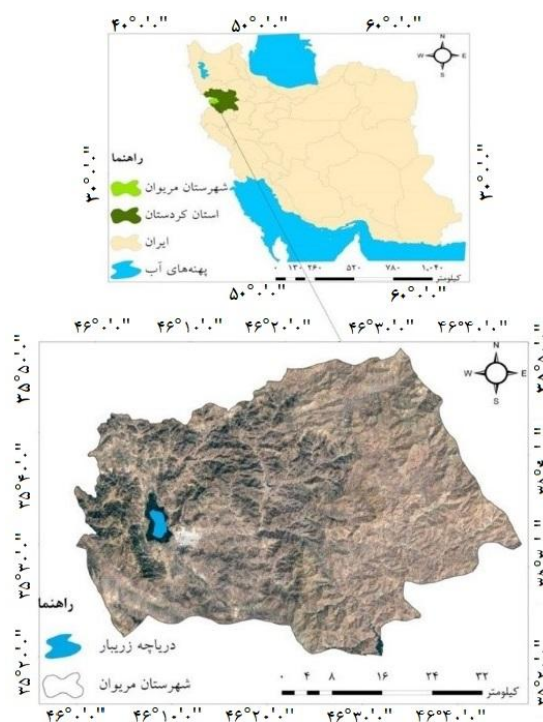
<sup>۱</sup> Temperature-Vegetation-Soil Moisture-Precipitation

<sup>۲</sup> Deep Feed Neural Network

<sup>۳</sup> Soil Moisture Deficit Index

در این مطالعه تصاویر شاخص پوشش گیاهی از محصولات مختلف این سنجنده شامل داده‌های MOD13A با قدرت تفکیک مکانی ۱ کیلومتر و تفکیک زمانی ۱۶ روزه، تصاویر درجه حرارت سطح از داده‌های MOD11A2 با قدرت تفکیک مکانی ۱ کیلومتر و تفکیک زمانی ۸ روزه، تصاویر شاخص ETI از داده‌های MOD16A2 با قدرت تفکیک مکانی ۱ کیلومتر و تفکیک زمانی ۸ روزه و در نهایت مجموعه داده بارش متوسط شبکه ماهانه از داده‌های TRMM 3B4  $\times 0.25$  (میلی‌متر در ماه) تبدیل شد، به عنوان اصلی‌ترین منبع سنجش از دور انتخاب شدند.

شمالی و ۴۶ درجه و ۴۵ دقیقه تا ۴۵ درجه و ۵۸ دقیقه طول شرقی نصف‌النهار گرینویچ در کشور ایران و استان کردستان واقع شده است [۱۵]. شهرستان مریوان دارای ویژگی‌های اقلیمی معتدل میانی و اقلیم‌های مختلف است و دارای بارندگی سالانه مابین ۹۰۰ - ۵۰۰ میلی‌متر است. در این تحقیق میانگین داده‌های بارندگی ماهیانه برای ایستگاه‌های سینوپتیک (قم چیان، ساوجی، گلچیدر، اسلام دشت، شی عطار) واقع در شهرستان مریوان برحسب میلی‌متر از سازمان هواشناسی کشور تهیه شده است. همچنین از داده‌های سنجنده مادیس و ماهواره TRMM، به عنوان تصاویر سنجش از دور برای تجزیه و تحلیل استفاده شده است. سنجنده مادیس که بر روی دو ماهواره ترا (Terra) و آکوا (Aqua) قرار دارد، دارای ۳۶ باند طیفی است که



شکل ۱: نقشه منطقه مورد مطالعه

خشکسالی و ترسالی یک سیستم طبقه‌بندی معین بنام شاخص SPI تعیین و معرفی نمودند که بیانگر شدت-

۱-۲- شاخص SPI  
مک کی و همکاران در سال ۱۹۹۳ برای تعیین

کاهش بارندگی در زیستگاه‌های مختلف همچنین نظارت بر خشکسالی و علائم هشداردهنده اولیه استفاده کرد. بر این اساس پس از محاسبه مقدار شاخص  $SPI$  در بازه زمانی دلخواه وضعیت رطوبت و شدت خشکسالی به کلاس‌های مختلف که در جدول (۱) ذکر شده است، تقسیم بندی می‌شود [۴].

های متفاوتی از خشکسالی در یک منطقه می‌باشد. این شاخص بر اساس بهترین تناسب بارش با توزیع گاما و برازش داده‌های بارشی محاسبه می‌شود [۱۶]. شاخص  $SPI$  در مقیاس‌های زمانی مختلف به صورت دوره آماری برای بازه‌های زمانی ۱، ۳، ۶، ۱۲، ۲۴ و ۴۸ ماهه محاسبه می‌شود که می‌توان برای نظارت بر تأثیرات

جدول ۱: درجه‌بندی مقادیر شاخص  $SPI$  (مک کی و همکاران، ۱۹۹۲)

مقدار شاخص	درجه خشکسالی یا ترسالی
$SPI \leq -2$	خشکسالی حاد
$-2 \leq SPI \leq -1$	خشکسالی شدید
$-1.5 \leq SPI \leq -1$	خشکسالی متوسط
$-1 \leq SPI \leq 1$	نرمال
$1 \leq SPI \leq 1.5$	ترسالی متوسط
$1.5 \leq SPI \leq 2$	ترسالی شدید
$SPI \geq 2$	ترسالی حاد

$$NDVI = \frac{(NIR - RED)}{(NIR + RED)} \quad \text{رابطه (۱)}$$

ب) شاخص پوشش گیاهی بهبود یافته ( $EVI$ )  
 شاخص  $EVI$  برای اولین بار توسط هیوت و لو در سال ۱۹۹۴ پیشنهاد شد. این شاخص بیش‌تر به تغییرات ساختاری تاج پوشش حساس است و اثرات جوی و تفاوت در بازتاب‌های آبی و قرمز را به حداقل می‌رساند [۱۹]. حساسیت شاخص  $EVI$  به خاک، اثرات اتمسفر و پخش آبروسل نسبت به شاخص  $NDVI$  کم‌تر است زیرا دارای طول موج‌های طیفی آبی است [۲۰]. شاخص  $EVI$  می‌تواند با تنش و تغییرات پوشش گیاهی مرتبط با حوادث خشکسالی در ارتباط باشد. باتوجه به رابطه (۲)  $G$ ، ضریب افزایش،  $L$ ، ضریب تعدیل کننده خاک،  $C_1$  و  $C_2$  ضرایب استفاده شده جهت تصحیح

## ۲-۲- شاخص‌های سنجش از دور

### الف) شاخص $NDVI$

شاخص  $NDVI$  معیار شناخته شده‌ای است که به متغیرهای بیوفیزیکی گیاهان مرتبط است و به طور گسترده‌ای برای نظارت بر پوشش گیاهی مورد استفاده قرار گرفته است. این شاخص از تقسیم اختلاف مقادیر باندهای طیفی قرمز و مادون قرمز نزدیک بر مجموع آنها محاسبه می‌شود. مقادیر این شاخص بین -۱ تا +۱ است. مقادیر منفی شاخص  $NDVI$  نشان‌دهنده آب و محیط مرطوب و مقادیر بین صفر و ۰/۱ به معنای زمین خالی، سنگ، شن یا برف، مقادیر بین ۰/۱ و ۰/۵ پوشش گیاهی تنک و مقادیر بیش‌تر از ۰/۵ نشان‌دهنده پوشش گیاهی متراکم است. شاخص  $NDVI$  از نسبت اختلاف مقادیر باندهای طیفی قرمز ( $RED$ ) و مادون قرمز نزدیک ( $NIR$ )، بر مجموع آنها محاسبه می‌شود [۱۷]. با استفاده از رابطه (۱) مقادیر شاخص  $NDVI$  از داده‌های سنجنده مادیس تولید می‌شود [۱۸].

<sup>۱</sup> Enhanced Vegetation Index

۱۹۹۵ معرفی گردید [۲۵]. این شاخص براساس این اصل تعریف می‌شود که دمای سطح سایبان یا خاک با افزایش تنش آب تحت شرایط دمایی بالا که باعث کمبود پوشش گیاهی یا آب در خاک می‌شود، افزایش می‌یابد. چون وقوع و توسعه خشکسالی ارتباط نزدیکی با دمای سطح زمین دارد، لذا زیاد شدن دمای سطح زمین در طول فصل رویش گیاهان، نشانگر شرایط نامساعد یا خشکسالی است در حالی که دمای پایین سطح زمین نشانگر شرایط نرمال خاک و گیاه از نظر رطوبت می‌باشد [۲۶]. در این راستا شاخص  $TCI$  حاصل از داده‌های مادون قرمز حرارتی نیز در این تحقیق برای نظارت بر خشکسالی مورد استفاده قرار گرفته است. باتوجه به رابطه (۴)، شاخص  $TCI$  در یک ماه خاص،  $LST_i$  مقدار دمای سطح زمین در آن زمان و همچنین  $LST_{max}$  و  $LST_{min}$  حداکثر و حداقل مقادیر  $LST$  برای ماه مربوطه در دوره مطالعه هستند. مقادیر این شاخص نیز در بازه صفر تا ۱۰۰+ بر حسب درصد قرار دارد بطوریکه مقدار صفر نشان‌دهنده خشکسالی شدید و مقدار نزدیک به ۱۰۰+ به معنای شرایط نرمال می‌باشد [۲۴]. رابطه (۴) شاخص  $TCI$  را بیان می‌کند.

$$TCI = \frac{LST_{max} - LST_i}{LST_{max} - LST_{min}} \times 100 \quad \text{رابطه (۴)}$$

#### ه) شاخص $ETI$

رطوبت خاک یک شاخص ساده و حساس به پدیده خشکسالی است و به طور گسترده برای نظارت بر کمبود آب مورد استفاده قرار می‌گیرد. این پارامتر همچنین یکی از مؤلفه‌های مهم هیدرولوژیکی است زیرا منعکس‌کننده مبادله جرم و انرژی بین سیستم خاک، آب و گیاه و جو می‌باشد. شرایط آب و هوایی غالب مانند متغیرهایی تابش، دما، باد و رطوبت نسبی بر روی شاخص  $ETI$  تأثیر می‌گذارند. مناطقی با پوشش گیاهی ضخیم و سالم می‌توانند شاخص  $ETI$  را برای فواصل طولانی‌تری پس از وقوع باران پایین نگه دارند

پراکندگی آئروسول در باند قرمز با استفاده از باند آبی و نیز  $P_{BLUE}$ ،  $P_{RED}$ ،  $P_{NIR}$  به ترتیب انعکاس در طول موج‌های آبی، قرمز و مادون قرمز نزدیک است. به طور معمول مقادیر ضرایب فوق برابر  $G = 2.5$ ،  $C_2 = 7.5$ ،  $C_1 = 6$  و  $L = 1$  در نظر گرفته می‌شوند. مقادیر این شاخص نیز در بازه ۱- تا ۱+ قرار دارد. تفاوت شاخص  $EVI$  با شاخص  $NDVI$  لحاظ کردن فاکتور تنظیم خاک ( $L$ ) و دو ضریب  $C_1$  و  $C_2$  می‌باشد [۲۱]. شاخص  $EVI$  در رابطه (۲) بیان شده است.

$$EVI = G \times \frac{P_{NIR} - P_{RED}}{P_{NIR} + C_1 \times P_{RED} - C_2 P_{BLUE} + L}$$

#### ج) شاخص $VCI$

هنگامی که پوشش گیاهی تحت تنش خشکی است، مقدار  $NDVI$  کاهش می‌یابد. با این حال، یک تصویر منفرد  $NDVI$  فقط نشان‌دهنده سلامت نسبی رشد گیاهان در یک زمان خاص است. بنابراین در این مطالعه شاخص  $VCI$  که می‌تواند رشد گیاهان را در سری زمانی منعکس کند، انتخاب گردید [۲۲]. این شاخص تأثیر نوبت تغییرات فصلی را از طریق یک نسبت کاهش می‌دهد. باتوجه به رابطه (۳)،  $VCI_i$  شاخص  $VCI$  یک ماه خاص و  $NDVI_i$  مقدار شاخص  $NDVI$  آن و  $NDVI_{max}$  و  $NDVI_{min}$  نیز حداقل و حداکثر مقادیر  $NDVI$  در ماه مربوطه در دوره مطالعه می‌باشند [۲۳]. مقادیر این شاخص برحسب درصد بیان شده و برخلاف دو شاخص قبلی در بازه صفر تا ۱۰۰+ قرار دارد [۲۳]. مقدار شاخص  $VCI$  صفر نشان‌دهنده وضعیت بدتر رشد گیاهان و خشکسالی شدید و همچنین مقدار نزدیک به ۱+ به معنای شرایط مطلوب گیاه می‌باشد [۲۴]. شاخص  $VCI$  در رابطه (۳) تعریف شده است.

$$VCI = \frac{NDVI_i - NDVI_{min}}{NDVI_{max} - NDVI_{min}} \times 100 \quad \text{رابطه (۳)}$$

#### د) شاخص $TCI$

شاخص  $TCI$  برای اولین بار توسط کوگان در سال

تئوری دو مدل مرسوم SVM و جنگل‌های تصادفی که در مدل‌سازی مورد استفاده قرار گرفته اند، بیان خواهد شد.

### الف) الگوریتم SVM

الگوریتم SVM در سال ۱۹۹۵ میلادی توسط بوزر و همکاران طراحی شد. این مدل را می‌توان جزو «روش‌های یادگیری نظارت شده» دسته‌بندی کرد که بر اساس تئوری یادگیری آماری و اصل به حداقل رساندن ریسک ساختاری عمل می‌کند [۳۳]. یک مدل داده محور که مبنای تئوری آن جداسازی خطی می‌باشد و در ساده‌ترین حالت طبقه‌بندی داده‌ها به صورت خطی و در فضایی دو بعدی است که دارای کم‌ترین پیچیدگی می‌باشد. داده‌ها در این فضا با استفاده از یک ابر صفحه به صورت خطی به دو طبقه کاملاً مجزا تفکیک می‌شوند. کورتس و همکاران در ادامه مطالعات تئوری ابر صفحه بهینه را به عنوان طبقه‌بندی‌کننده ارائه داده و طبقه‌بندی‌کننده‌های غیرخطی را به کمک توابع کرنل معرفی کردند [۳۴]. روند آموزش یک تابع تصمیم‌گیری SVM شناسایی یک ابر صفحه قابل تکرار است که فاصله (یعنی "حاشیه") بین بردارهای پشتیبانی هر دو برجسب کلاس را به حداکثر می‌رساند. بنابراین ابر صفحه مطلوب صفحه‌ای است که حاشیه بین طبقات را به حداکثر برساند. این یادگیری‌ها اساساً با نگاشت ورودی‌ها در فضای مشخصه‌ای با ابعاد بالا، یک مسئله غیرخطی را به یک مسئله خطی تبدیل کرده و یک مدل خطی را وارد یک فضای ویژگی می‌کنند. مدل SVM در اصل برای حل مشکلات طبقه‌بندی استفاده می‌شد تا اینکه واپنیک و همکاران در سال ۱۹۹۶ برای برطرف کردن مشکلات رگرسیون، از مدل SVR استفاده نمودند. رگرسیون ماشین بردار از دو دسته داده یکی داده‌ها جهت آموزش مدل و دیگری داده‌های آزمون برای پیش بینی و تست اطلاعات استفاده می‌کند [۲۱]. دقت این مدل رگرسیون وابسته

زیرا ریشه‌ها دسترسی به رطوبت را در خاک فراهم می‌کنند [۲۷]. در این پژوهش میزان شاخص ETI از محصولات سنجنده مادیس و براساس الگوریتم پنمن مانیتس استخراج شده است [۲۸].

### و- شاخص وضعیت بارش (PCI)

داده‌های TRMM می‌تواند تخمین بارش ماهانه را ارائه داده و از آن در استخراج اطلاعات خشکسالی هواشناسی استفاده نمود [۲۹]. شاخص PCI با استفاده از داده‌های TRMM 3B43 و الگوریتم مشابه VCI محاسبه شده و برای تشخیص کمبودهای بارش از سیگنال آب و هوا تعریف می‌شود [۳۰]. شاخص PCI با توجه به میزان بارندگی در محدوده صفر تا ۱ تعریف شده است. در صورت خشکسالی هواشناسی که بارندگی آن بسیار کم است، PCI نزدیک یا برابر با صفر و در شرایط طغیان PCI نزدیک به ۱ است [۳۱]. در این شاخص PCI مطابق رابطه ۵ تعریف شده است که در آن  $TRMM_{max}$ ،  $TRMM_{min}$  و  $TRMM$  به ترتیب برابر مقادیر پیکسل بارندگی و نیز حداکثر و حداقل آن در دوره مطالعه می‌باشد [۳۲].

$$PCI = \frac{TRMM - TRMM_{min}}{TRMM_{max} - TRMM_{min}} \quad (۵) \text{ رابطه}$$

### ۳-۲- الگوریتم‌های یادگیری ماشین

یادگیری ماشین شاخه‌ای از علوم کامپیوتر و زیرمجموعه‌ای از هوش مصنوعی است که بر اساس فرآیند یادگیری بیولوژیکی ساخته می‌شود و هدف آن یادگیری الگوهای نهفته در داده‌ها برای بهبود عملکرد در کارهای مختلف است. رویکرد یادگیری ماشین با طراحی الگوریتم‌ها برای یادگیری از داده‌های قابل خواندن توسط ماشین سر و کار دارد و دامنه‌های اصلی مانند داده کاوی، برنامه نویسی‌های پیچیده و برنامه‌های نرم افزاری را پوشش می‌دهد [۳۳]. در ادامه مبانی

<sup>۱</sup> Precipitation Condition Index



( $\gamma$ )، اپسیلون ( $\epsilon$ ) و تابع کرنل است [۳۵]. به دلیل وابسته بودن پارامترهای مدل  $SVM$  انتخاب پارامتر بهینه پیچیده است. نرم‌افزارهای موجود  $SVR$ ، معمولاً پارامترهای  $SVM$  را به عنوان ورودی‌های مدل از کاربر دریافت می‌کند. پارامتر  $C$  رابطه جایگزینی بین پیچیدگی مدل و درجه برای این انحرافات بزرگتر از  $\epsilon$  را تعیین می‌کند. پس می‌توان گفت الگوریتم  $SVM$  داده‌ها را با توجه به دسته‌های از پیش تعیین شده آنها به یک فضای جدید می‌برد به گونه‌ای که داده‌ها به صورت خطی (یا ابر صفحه) قابل تفکیک و دسته‌بندی باشند. سپس با یافتن خطوط پشتیبان (صفحات پشتیبان در فضای چند بعدی)، سعی در یافتن معادله خطی دارد که بیش‌ترین فاصله را بین دو دسته ایجاد کند. نتایج حاصل از این روش علاوه بر دقت بالا دارای ثبات خوبی نیز هستند [۳۶]. در بعضی از مسائل رگرسیون به دلیل حضور داده‌های پراکنده و پرت استفاده از رگرسیون خطی به طور کامل امکان‌پذیر نیست. در نتیجه برای حل مسائل غیر خطی از توابع کرنل استفاده می‌شود که براساس ضرب داخلی داده‌ها تعریف می‌شوند [۳۷].

### ب) مدل جنگل‌های تصادفی

الگوریتم رگرسیون جنگل‌های تصادفی یک الگوریتم یادگیری جمعی است که مجموعه بزرگی از درخت‌های رگرسیون را ترکیب می‌کند. یک درخت رگرسیون مجموعه‌ای از شرایط یا محدودیت‌هایی را نشان می‌دهد که بصورت سلسله مراتبی سازماندهی شده‌اند و به ترتیب از ریشه تا یک برگ درخت استفاده می‌شود. این روش یادگیری گروه به عنوان "کیسه" شناخته می‌شود زیرا درختانی را رشد می‌دهد که درختان متوالی به درختان قبلی بستگی ندارند. دو ویژگی مهم در ساخت جنگل تصادفی روش بگینگ و انتخاب تصادفی در هر گره است [۳۷]. پارامترهای اصلی جنگل‌های تصادفی عبارتند از تعداد متغیرهای ورودی به طور تصادفی در

به تنظیم مناسب پارامترهای مدل‌سازی است [۲۱]. در رگرسیون  $SVR$  ورودی  $x$  در یک فضا با تعداد  $m$  عدد ویژگی مختلف به صورت غیرخطی است و سپس با استفاده از این ویژگی‌ها یک مدل خطی ساخته می‌شود. مدل خطی به صورت رابطه (۶) است.

$$f(x, w) = \sum_{i=1}^n w_i g_i(x) + b \quad \text{رابطه (۶)}$$

در رابطه (۶)،  $w$  بردار وزن در فضای ویژگی‌هاست،  $\xi_i$  و  $\xi_i^*$  متغیرهای کمبود کمکی نامیده می‌شوند و انحراف خروجی مدل تخمینی را از داده‌های هدف اندازه‌گیری می‌کنند. همچنین  $g_i(x)$  اشاره به یک مجموعه از انتقال‌های خطی دارد و  $b$  ترم بایاس است. کیفیت برآورد نیز به وسیله تابع  $L(y, f(x, w))$  اندازه‌گیری می‌شود. نکته این است که  $\epsilon$  منطبق بر کم‌ترین قدر مطلق است. فرمول این رگرسیون به صورت رابطه (۷) و (۸) است.

$$L(y, f(x, w)) = f(x) \quad \text{رابطه (۷)}$$

$$SVR = \left\{ \begin{array}{l} 0, |y - f(x, w)| \leq \epsilon \\ |y - f(x, w)| \leq \epsilon, \text{ otherwise} \end{array} \right\} \quad \text{رابطه (۸)}$$

در رابطه (۸)  $SVR$  یک رگرسیون خطی در ابعاد بالایی از فضای ویژگی‌ها با استفاده از  $\epsilon$  اجرا می‌کند و سعی می‌کند پیچیدگی مدل را با کم‌ترین  $\|w\|^2$  کاهش دهد. این مسئله می‌تواند به وسیله معرفی متغیر  $\xi_i, \xi_i^*$  و  $n = 1, 2, \dots$  برای اندازه‌گیری انحراف نمونه‌های آموزشی  $\epsilon$  شرح داده شود. پس  $SVR$  به عنوان کمینه کردن تابع به صورت رابطه (۹) می‌باشد.

$$\min \frac{1}{2} \|w\|^2 + c \sum_{i=1}^n (\xi_i + \xi_i^*) \quad \text{رابطه (۹)}$$

$$s.t. = \begin{cases} y_i - f(x_i, w) \leq \epsilon + \xi_i^* \\ f(x_i, w) - y_i \leq \epsilon + \xi_i \\ \xi_i^*, \xi_i \geq 0, i=1, 2, \dots, n \end{cases}$$

این مسئله بهینه می‌تواند تبدیل به مسئله دوگانه و انتخاب بردارهای حامل و تابع کرنل شود. عملکرد خوب تابع رگرسیون وابسته به انتخاب خوب پارامترهای گاما

## ۳- داده‌های مورد استفاده

در این پژوهش، شاخص  $SPI$  در ایستگاه‌های هواشناسی شهرستان مریوان برای یک سری زمانی ۱۷ ساله طی سال‌های ۲۰۰۱ تا ۲۰۱۷ در مقیاس ماهانه و سالانه محاسبه گردید. همچنین شاخص‌های سنجش از دور در محیط گوگل ارث انجین ( $GEE^2$ ) شامل شاخص  $ETI$  از ۷۸۱ تصویر، شاخص  $VCI$  ۳۹۱ تصویر شاخص  $NDVI$  ۳۹۱ تصویر، شاخص  $VCI$  ۳۹۱ تصویر و در نهایت شاخص  $TCI$  با ۷۸۱ تصویر از سنجنده مادیس و شاخص بارندگی از داده‌های ماهواره  $TRMM$  با ۲۰۴ تصویر برای شهرستان مریوان محاسبه و میانگین ۱ ماهه آنها استخراج گردید. برای همه شاخص‌های فوق-الذکر ماه‌های ژانویه و دسامبر به دلیل کاهش رشد گیاهان در ماتریس داده‌ها لحاظ نشده است. در ادامه شاخص  $SPI$  یک ماهه به عنوان متغیر وابسته و شش شاخص سنجش از دور مذکور دیگر به عنوان متغیر مستقل انتخاب و ماتریسی از داده‌ها با ۱۷۰ سطر داده بعنوان ورودی الگوریتم پیشنهادی تولید گردید. همچنین از دو روش  $SVR$  و جنگل تصادفی جهت مدل‌سازی سری زمانی شاخص‌های سنجش از دور و شاخص داده‌های هواشناسی و ارزیابی دقت آنها استفاده شد. برای ارزیابی دو مدل نیز از پارامترهای مجذور میانگین مربعات خطا ( $RMSE^3$ ) و ضریب تبیین ( $R^2$ ) استفاده شده است. روابط (۱۱) و (۱۲) نحوه محاسبه این دو پارامتر را بیان می‌کنند:

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y}_i)^2} \quad \text{رابطه (۱۱)}$$

هر تقسیم و درخت، همچنین تعداد درختان انتخاب شده است. جنگل تصادفی با بسیاری از نمونه‌های خود راه‌انداز<sup>۱</sup> که به طور تصادفی با جایگزینی داده‌های آموزشی اصلی ترسیم شده‌اند، آغاز می‌شود. یک درخت رگرسیون برای هر یک از نمونه‌های خود راه‌انداز متناسب می‌شود. برای هر گره در هر درخت، مجموعه کوچکی از متغیرهای ورودی انتخاب شده از مجموعه کلی به طور تصادفی برای تقسیم‌بندی باینری در نظر گرفته می‌شود. معیار تقسیم درخت رگرسیون بر اساس انتخاب متغیر ورودی با کم‌ترین شاخص جینی مطابق رابطه (۱۰) است.

$$I_G(t_{X(x_i)}) = 1 - \sum_{j=1}^m f(t_{X(x_i)}, j)^2 \quad \text{رابطه (۱۰)}$$

در رابطه (۱۰)،  $f(t_{X(x_i)}, j)$  نسبت نمونه‌ها با مقدار  $x_i$  متعلق به ترک  $j$  به عنوان گره  $t$  است. مقدار پیش‌بینی شده یک مشاهده با میانگین‌گذاری روی همه درختان محاسبه می‌شود. بنابراین دو پارامتر باید در جنگل تصادفی بهینه شود یکی تعداد درختان رگرسیون و دیگری تعداد گره ورودی متغیرهای ورودی است [۳۸]. جنگل‌های تصادفی برای استفاده در مسائل مربوط به رگرسیون به روش رشد درختان وابسته به یک بردار تصادفی  $\Theta$  ایجاد می‌شوند. هر درخت به بزرگ‌ترین حد خود توسعه می‌یابد به گونه‌ای که پیش‌بینی کننده درخت  $h(X, \theta)$  مقادیر عددی بر خلاف برچسب‌های کلاس به خود گیرد. مقادیر خروجی عددی است و فرض می‌شود که مجموعه داده آموزشی به طور مستقل از توزیع بردار تصادفی  $Y, X$  انتخاب شده است. خطای تعمیم مربع میانگین برای هر پیش‌بینی کننده، عددی مانند  $h(X)$  است که پیش‌بینی کننده جنگل تصادفی با گرفتن میانگین بیش از  $k$  درختان  $\{h(X, \theta_k)\}$  است [۳۹].

<sup>۲</sup> Google Earth Engine<sup>۳</sup> Root Mean Square Error<sup>۱</sup> Bootstrap

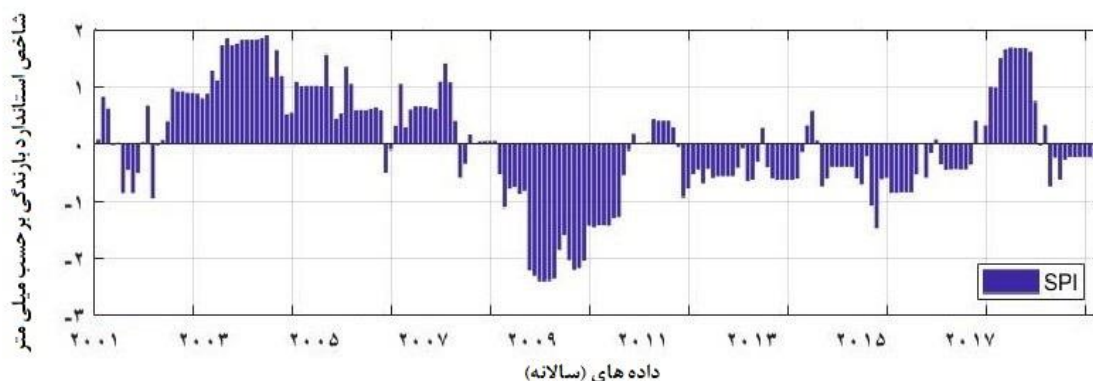
های اقلیمی، روند تغییرات شاخص *SPI* دوره آماری سالانه مربوط به شهرستان مریوان طبق کلاس‌های مختلف در جدول (۱) مورد بررسی قرار گرفت. نتایج نشان داد که مقدار شاخص *SPI* در دوره ۱۷ ساله کاهش یافته است (شکل (۳)). با توجه به طبقه‌بندی مک کی خشکسالی شدید در سال ۲۰۰۸ با ۱٫۸۳- *SPI*= و خشکسالی شدید در سال ۲۰۰۹ با ۱٫۳۹- *SPI*= اتفاق افتاده است. همچنین سال‌های ۲۰۰۳ و ۲۰۰۴ و ۲۰۱۶ به ترتیب با ترسالی متوسط ۶ و شاخص *SPI*= ۱٫۲۰، ۱٫۰۵، ۱٫۱۹ سال‌های مرطوب دوره را نشان می‌دهند.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad \text{رابطه (۱۲)}$$

در رابطه بالا  $y_i$  مقدار هدف،  $\hat{y}_i$  مقدار خروجی برآورد شده از مدل،  $\bar{y}$  میانگین مقادیر مشاهدات و  $n$  تعداد نمونه مشاهدات می‌باشد.

#### ۴- نتایج پیاده‌سازی و آنالیز

همان‌طور که در فصول قبل توضیح داده شد شاخص بارش استاندارد فقط بر اساس بارش استوار است و به عنوان یک شاخص نسبتاً ساده در طول سال طراحی شده است و برای شرایط تأمین آب اطلاعاتی را فراهم می‌کند. با توجه به نقش اساسی بارندگی در وقوع خشکسالی-



شکل ۳: روند تغییرات شاخص *SPI* میانگین ایستگاه شهرستان مریوان سال ۲۰۰۱ تا ۲۰۱۷ به صورت سالیانه

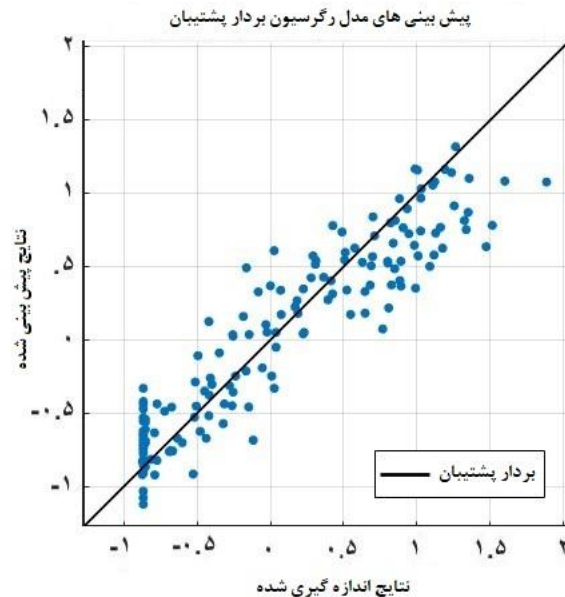
زمانی ۱ ماهه (میانگین) ارزیابی و محاسبه شد. در نهایت از شاخص *SPI* یک ماهه به عنوان متغیر وابسته و شش شاخص سنجش از دور دیگر به عنوان متغیرهای مستقل انتخاب گردید. بر اساس داده‌های این متغیرها ماتریسی از داده‌ها با ۱۷۰ سطر داده بعنوان ورودی مدل تشکیل گردید. از این تعداد داده ۸۰ درصد از داده‌ها (۲۰۰۱ تا ۲۰۱۴) در مرحله آموزش و ۲۰ درصد دیگر (۲۰۱۵ تا ۲۰۱۷) در مرحله آزمون مدل مورد استفاده قرار گرفت. لازم به توضیح است که از در مرحله آموزش، مدل بخشی از داده‌های آموزشی را برای اعتبارسنجی (*Validation*) مدل

#### ۴-۱- پیاده‌سازی و آنالیز الگوریتم *SVM*

الگوریتم *SVM* اساساً با استفاده از توابع هسته، مسئله رگرسیون غیر خطی را به یک مسئله خطی تبدیل می‌کند تا فضای ورودی اصلی را به یک فضای ویژگی جدید با ابعاد بالاتر ترسیم کند [۴۱]. در این مدل، شاخص *SPI* در ایستگاه‌های هواشناسی شهرستان مریوان و مقیاس ماهانه برای یک سری زمانی ۱۷ ساله طی سال‌های ۲۰۰۱ تا ۲۰۱۷ و شاخص‌های سنجش از دور با استفاده از محصولات سنجنده‌های مادیس و *TRMM* در محیط *GEE* به صورت ماهیانه برای سال ۲۰۰۱ تا ۲۰۱۷ از ماه فوریه تا نوامبر در مقیاس

زمینی و پیش بینی شده به وسیله مدل *SVM* دیده می شود، می توان دید که مدل بردار پشتیبان مدلی جامع برای خشکسالی است و کاربرد خوبی در پیش خشکسالی هواشناسی دارد (شکل (۴) و (۵)) و بیانگر همبستگی مثبت و معنی داری بین شاخص جامع خشکسالی خروجی مدل و شاخص خشکسالی هواشناسی در مقیاس ایستگاه اندازه گیری شد، که بین آنها همبستگی خوبی وجود دارد. این مطالعه روشی جدید برای ارزیابی جامع خشکسالی منطقه ارائه می دهد.

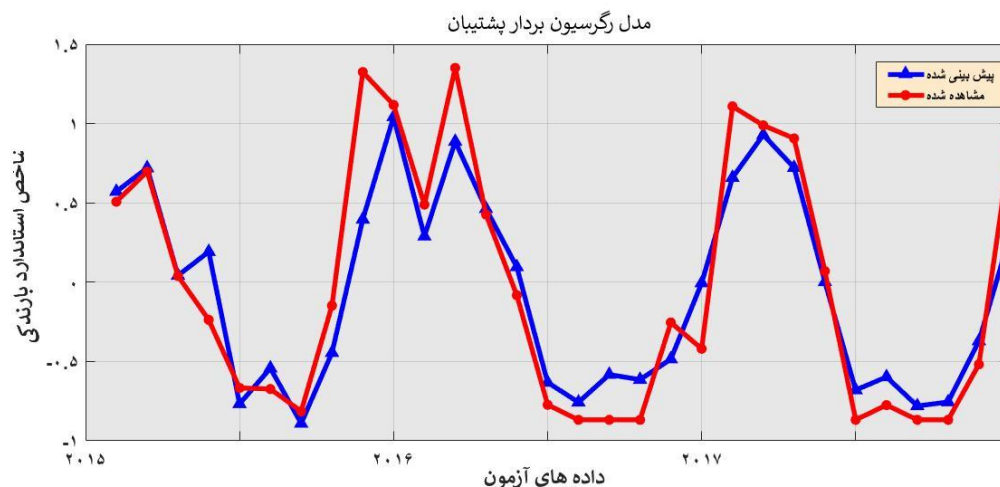
استفاده می نماید که با داده های آزمون مدل متفاوت است. مقدار پارامترهای بهینه مدل از جمله پارامتر تنظیم به طور معمول ( $C=1$ ) در نظر گرفته شد. همچنین از بین پارامتر کرنل های گاوسی، سیگموئید و پایه شعاعی، کرنل گاوسی بدلیل عملکرد خوب انتخاب گردید. سپس ارتباط بین شش فاکتور خشکسالی مبتنی بر سنجش از دور و داده های زمینی از طریق روش *SVR* مدل سازی شد. در ادامه نمودار پراکندگی بین مقادیر شاخص *SPI* و مقادیر پیش بینی شده در شهرستان مریوان ترسیم گردید (شکل (۴)). همانطور که در نمودار پراکنش مقادیر اندازه گیری شده



شکل ۴: نمودار پراکندگی بین مقادیر شاخص *SPI* و مقادیر پیش بینی شده شهرستان مریوان

سنجش از دور با شاخص *SPI* داده های زمینی به صورت رابطه خطی رگرسیونی مقایسه شد. در این راستا مقادیر *RMSE* و  $R^2$  بین مقادیر واقعی و پیش بینی شده، به ترتیب برابر  $0.274$  و  $0.88$  بدست آمد.

بعلاوه به منظور بررسی عملکرد مدل *SVR* و رابطه بین متغیرهای مستقل، نمودار مقادیر پیش بینی شده در مقابل پاسخ واقعی برای مدل ترسیم شد (شکل (۵)). همچنین نتایج پیش بینی مدل *SVR* برای داده های شهرستان مریوان و میزان همبستگی شاخص های



شکل ۵: مقیاسه نتایج پیش‌بینی شده و مشاهده شده در شهرستان مریوان برای سال ۲۰۱۵ تا ۲۰۱۷

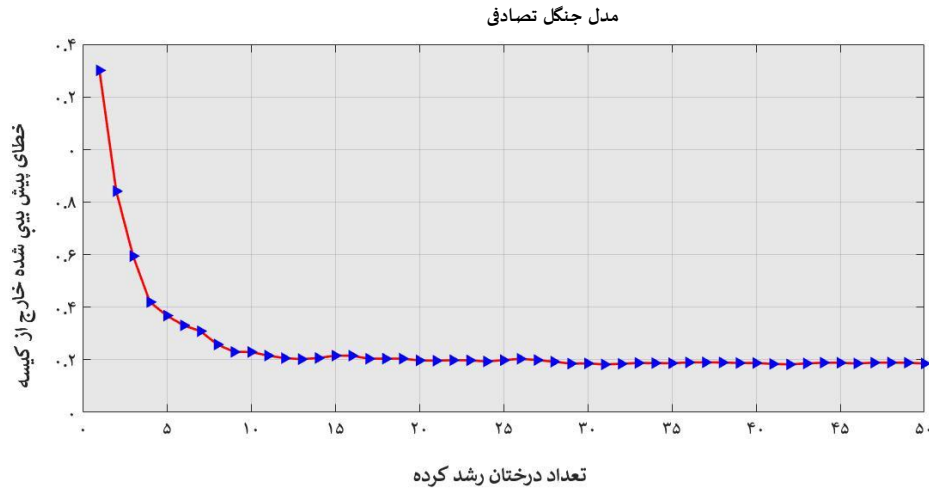
#### ۴-۲- پیاده‌سازی و آنالیز الگوریتم جنگل تصادفی

عملکرد الگوریتم جنگل تصادفی مشابه سایر مدل‌های یادگیری تحت نظارت است و همچنین یک تکنیک گروهی است که مشکل بیش‌برازش و عدم اطمینان را کاهش می‌دهد. همچنین این مدل از تکنیک‌های مبتنی بر یک درخت بسیار بهتر است و با ساختن جنگلی از درختان تصمیم‌گیرنده و استفاده از تکنیک خود راه‌انداز، که در آن هر درخت به‌طور مستقل و بر اساس زیرمجموعه‌ای تصادفی از متغیرهای پیش‌بینی‌کننده ایجاد شده است، شروع می‌شود. درختان بدون هرس به حداکثر اندازه می‌رسند و در نتیجه میانگین خروجی از تمام درختان چند تصمیم‌گیری نتیجه نهایی است که در دستیابی به عملکرد رگرسیون بسیار مؤثر است. هنگام اجرای مدل یک نکته اساسی وجود دارد و آن درک مجموعه‌ای صحیح از هایپر پارامترها (تعداد درختان تصمیم و تعداد ویژگی‌های زیر هر درخت) است که سایر پارامترها را تنظیم می‌کند و مدل را برای هر ترکیبی ارزیابی می‌کند. همانطور که از نام آن مشخص است، روش جستجوی تصادفی از ترکیبات

بر اساس نتایج تحقیق شاخص‌های ماهواره‌ای مدل با شاخص خشکسالی هواشناسی در طول دوره ارتباط معنی‌داری نشان داد که بیانگر اینست که این مدل از پتانسیل خوبی در پایش خشکسالی هواشناسی برخوردار است. پراکندگی و ضرایب همبستگی شاخص پوشش گیاهی با *SPI* در بیشتر ماه‌ها بسیار زیاد و مثبت ارزیابی گردید. در ماه اکتبر که بیش‌تر پوشش گیاهی علفزارهای این منطقه نزدیک به پیری کامل است، ارتباط معنی‌دار زیادی وجود ندارد. این بدان معناست که شاخص‌های سنجش از دور، شاخص خوبی برای وضعیت بارش است و می‌تواند منبع مهمی در هنگام استفاده برای شناسایی و پایش خشکسالی در این منطقه باشد. این نه تنها به دلیل پاسخ قابل درک به خشکسالی است، بلکه به این دلیل است که پوشش گیاهی یک ادغام‌کننده نسبتاً کامل متغیرهای فیزیکی است که تحت تأثیر شرایط آب و هوایی قرار دارند [۱۷]. قابل ذکر است که همبستگی زیاد فقط در اواسط فصل رشد اتفاق افتاده است اما در دوران سبز شدن و پیری شدن بسیار کم‌تر است. دلیل این امر حساسیت بیش‌تر گیاهان به آب در مرحله رشد و تولیدمثل است.

که از این تعداد ۸۰ درصد از داده‌ها (۲۰۰۱ تا ۲۰۱۴) در مرحله آموزش و ۲۰ درصد (۲۰۱۵ تا ۲۰۱۷) در مرحله آزمون استفاده شده است. در ادامه با تعیین پارامترهای بهینه مدل، نتایج پیش‌بینی از مدل استخراج شده و عملکرد آن براساس معیارهای آماری  $RMSE$  و  $R^2$  برای دوره آزمون ارزیابی گردید. در مدل پیشنهادی نحوه رسیدن به نتیجه نهایی و قابل قبول از طریق تعدادی درخت‌های تصمیم است که با هم یک جنگل را تولید می‌کنند که می‌تواند تصمیم‌های بهتری را اتخاذ نماید. به طور کلی تعداد بیش‌تر درخت‌ها، کارایی مدل را افزایش داده و پیش‌بینی‌ها را پایدار می‌سازند اما محاسبات را کندتر می‌کنند. در نتیجه این درختان تا زمانی افزایش می‌یابد که در جواب نهایی تغییرات کمی اتفاق بیفتد. در این تحقیق با تعداد ۵۰ درخت تصادفی انجام شد که روند رسیدن به جواب بهینه در شکل (۶) نمایش داده شده است.

تصادفی پارامترهای بیش از حد برای رسیدن به نتیجه‌ای بهتر استفاده می‌کند [۴۱]. استفاده از مدل جنگل تصادفی شامل درک اهمیت نسبی متغیرهای مورد استفاده برای پیش‌سری زمانی خشکسالی است. مطالعات زیادی در پیش‌بینی خشکسالی با استفاده از مدل‌های مختلف یادگیری ماشین در یادگیری بانظارت انجام شده است. با این حال، مدل جنگل تصادفی بخصوص در ایران نسبتاً کم‌تر مورد آزمایش قرار گرفته است. در این مدل از ۶ متغیر مستقل شامل شاخص‌های  $NDVI$ ، شاخص  $VCI$ ، شاخص پوشش گیاهی پیشرفته، شاخص  $TCI$  و شاخص  $ETI$  همگی حاصل از سنجنده مادیس و شاخص بارندگی مستخرج از ماهواره  $TRMM$  و متغیر وابسته شاخص  $SPI$  منتج از ایستگاه‌های هواشناسی برای شهرستان مریوان محاسبه گردید. در روند ساخت مدل جنگل تصادفی ماتریسی از داده‌ها با ۱۷۰ سطر داده در اختیار الگوریتم قرار گرفت



شکل ۶: تأثیر تعداد درختان و متغیرهای تقسیم تصادفی

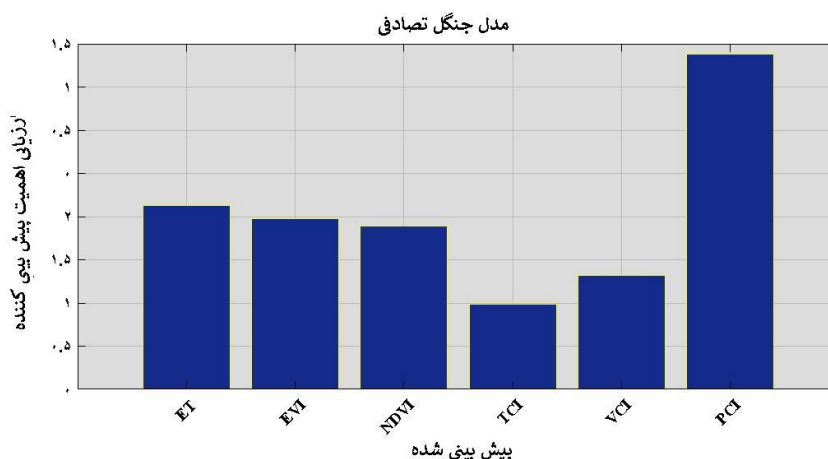
متغیرهای پیوسته مورد استفاده قرار می‌گیرد. مقدار ضریب همبستگی +۱ و -۱ در نوسان است، اگر متغیرها همسو باشند به آن همبستگی مستقیم یا مثبت می‌گویند و اگر مقدار ضریب همبستگی صفر باشد این بدان معناست که دو متغیر مورد نظر مستقل از هم

#### ۴-۳- آنالیز دقت و اهمیت شاخص‌ها

در تحقیق حاضر میزان همبستگی بین مقادیر شاخص‌های سنجنش از دور با شاخص  $SPI$  داده‌های زمینی در مدل‌سازی جنگل تصادفی شهرستان مریوان محاسبه گردید که به منظور بررسی رابطه بین

با اهمیت نسبی ۰/۸، ۰/۷۵، ۰/۷۳ و ۱/۵ بیش‌ترین ارتباط و سطح معنادار را در بین سایر متغیرها دارند (شکل ۷).

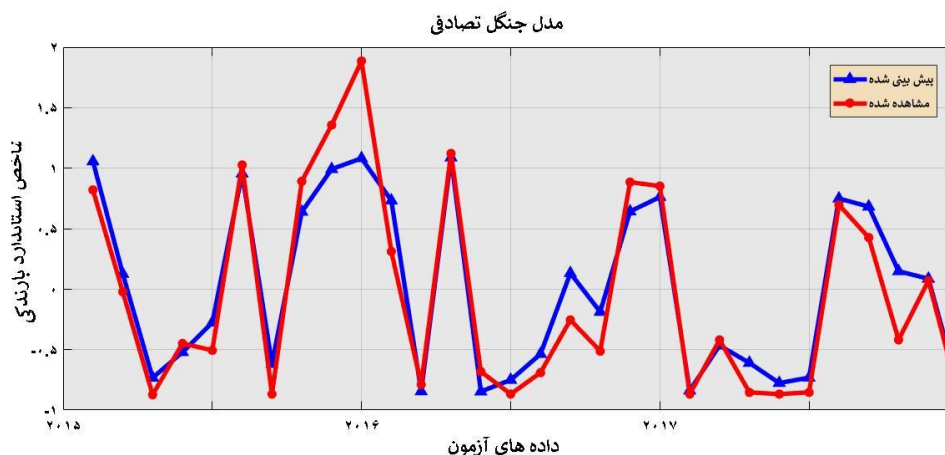
هستند و هیچگونه رابطه‌ای بین آنها وجود ندارد. نتایج میزان اهمیت شاخص‌های سنجش از دور با شاخص *SPI* داده‌های زمینی نشان می‌دهد که شاخص *ETI*، شاخص *EVI*، شاخص *NDVI* و شاخص *PCI* به ترتیب



شکل ۷: اهمیت نسبی متغیرهای پیش‌بینی برای شهرستان مریوان

۰/۹۰۹ برای  $R^2$  بدست آمد (شکل ۸). بنابراین نتایج مدل جنگل تصادفی نشان می‌دهد که در صورت فقدان داده‌های بارندگی می‌توان از شاخص‌های سنجش از دور برای پیش‌بینی و نظارت بر خشکسالی استفاده کرد.

همچنین با مقایسه مقادیر برآورده شده توسط مدل و مقادیر واقعی از طریق شاخص آماری *RMSE* و میانگین قدر مطلق خطا نشان داد که مدل از صحت بالایی برخوردار است. در این رابطه بر اساس مقادیر واقعی و پیش‌بینی شده، مقدار ۰/۲۵۹ برای *RMSE* و مقدار

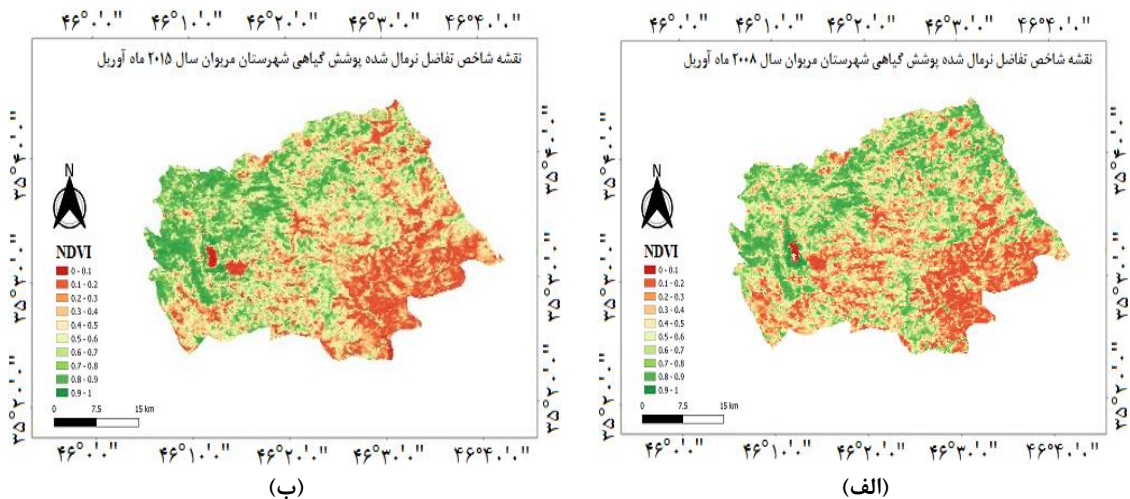


شکل ۸: مقایسه نتایج پیش‌بینی شده و مشاهده شده داده‌های آزمون در شهرستان مریوان برای سال ۲۰۱۷ تا ۲۰۱۵

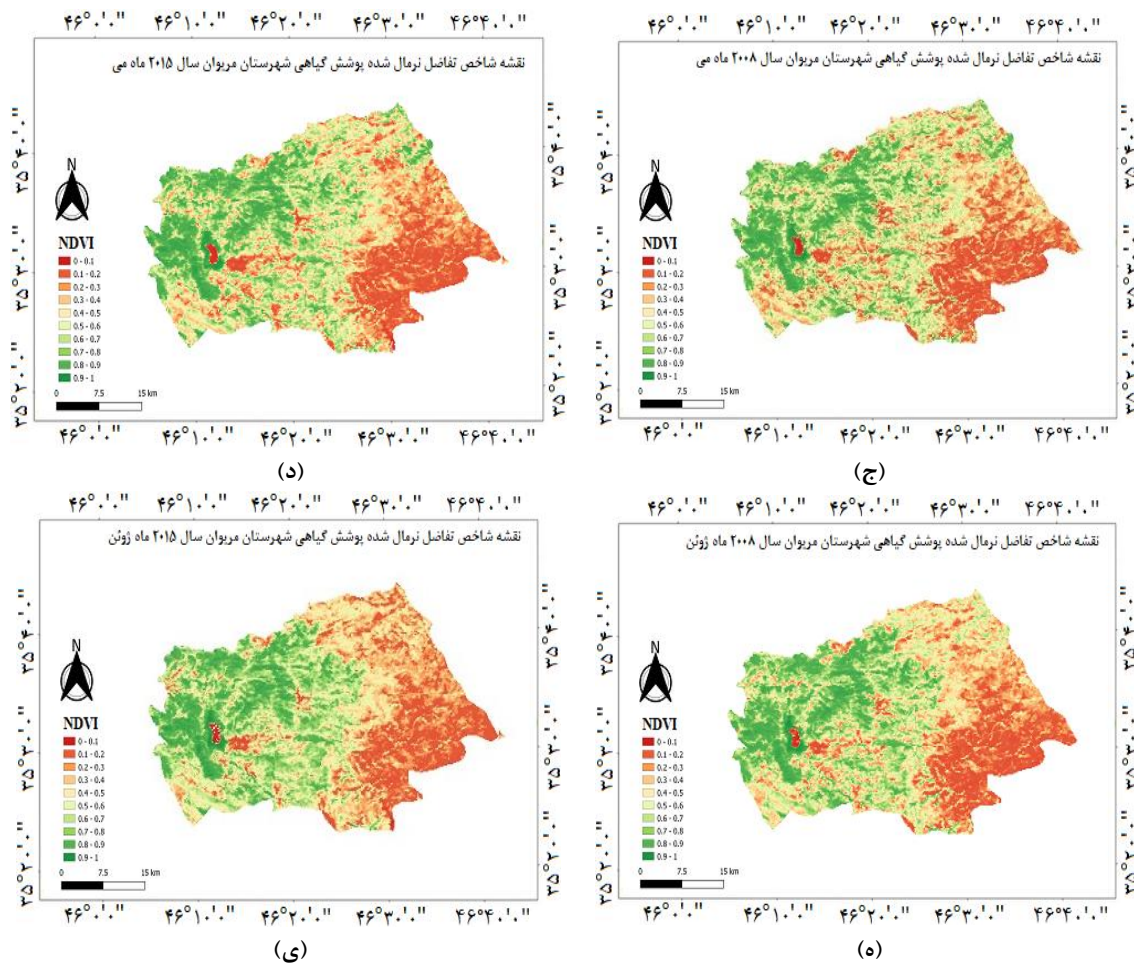
## ۴-۴- تأثیر عوامل مختلف بر شبیه‌سازی

به منظور ارزیابی دقیق از رخداد، گسترش و شدت خشکسای باید تصور صحیحی از توزیع مکانی و زمانی پارامترهای هواشناسی، هیدرولوژیکی و متغیرهای سطحی داشت. از طریق دریافت داده‌های سنجش از دور با قدرت تفکیک مکانی و زمانی مناسب در سطح وسیعی از زمین و به صورت نزدیک به زمان واقعی، امکان پایش روند خشکسالی‌ها امکان پذیر است. در این تحقیق دو روش یادگیری آماری محبوب یعنی جنگل تصادفی و SVR با هدف ارزیابی اهمیت متغیر و سپس برای انتخاب بهترین متغیر مقایسه گردید. بعلاوه به منظور ارائه یک شاخص معتبر، از شاخص‌های سنجش از دوری شامل *NDVI*, *EVI*, *ET*, *VCI*, *TCI*, *PCI* استفاده شد تا بدین وسیله بتوان در صورت لزوم این شاخص‌ها را جایگزین شاخص‌های اقلیمی کرد. شاخص‌های خشکسالی مبتنی بر سنجش از دور عمدتاً از طیف‌های نوری و مادون قرمز امواج الکترومغناطیس مشتق شده‌اند و کاربرد خوبی در نظارت بر خشکسالی بلند مدت به ویژه خشکسالی هواشناسی در مناطق

نیمه خشک ارائه می‌دهند. نتایج این مطالعه نشان داد شاخص *PCI*، شاخص *ETI* و شاخص *EVI* نسبت به دیگر شاخص‌های خشکسالی سنجش از دوری در نظارت بر خشکسالی عملکرد بهتری دارند و بالاترین همبستگی را با شاخص *SPI* یک ماهه نشان می‌دهند (شکل ۹). از آنجا که این مطالعه در جستجوی راهکاری برای جایگزینی شاخص‌های سنجش از دور به منظور تحلیل و پیش‌بینی خشکسالی است لذا نقشه شاخص *NDVI* به دست آمده از تصاویر مادیس برای ارزیابی خشکسالی منطقه مورد مطالعه در نرم‌افزار *QGIS* ترسیم شد و با نتایج شاخص *SPI* مقایسه گردید. بعنوان مثال شاخص *NDVI* در سال ۲۰۰۸ کاهش قابل توجهی داشته است و در همان زمان شاخص *SPI* نیز شرایط وقوع خشکسالی را نشان می‌دهد. این در حالی است که با افزایش میزان بارش و بهبود شرایط شاخص *SPI* در سال ۲۰۱۵، شاخص *NDVI* نیز در این سال در بهترین وضعیت کل دوره مطالعه می‌باشد.







شکل ۹: نقشه شاخص  $NDVI$  شهرستان مریوان برای (الف) سال ۲۰۰۸ ماه آوریل، (ب) سال ۲۰۱۵ ماه آوریل، (ج) سال ۲۰۰۸ ماه می، (د) سال ۲۰۱۵ ماه می، (ه) سال ۲۰۰۸ ماه ژوئن، (ی) سال ۲۰۱۵ ماه ژوئن.

##### ۵- نتیجه‌گیری و پیشنهادها

میزان شاخص  $SPI$  یک ماهه پیش‌بینی شده با روش جنگل تصادفی با میزان شاخص جامع به دست آمده از مدل یک ماهه تا حد زیادی مشابه بود. در حالی که میزان شاخص جامع یک ماهه پیش‌بینی شده از  $SVR$  محدودتر بوده است که بیانگر آن است روش جنگل تصادفی بهتر از روش دیگر آموزش دیده است. یکی از دلایل این امر این است که روش جنگل تصادفی در مقایسه با  $SVR$  از یک رویکرد گروهی برای جلوگیری از مشکل وجود بیش برآزش استفاده می‌کند. با تعیین پارامترهای بهینه مدل‌های پیشنهادی، نتایج

پیش‌بینی از مدل تولید شده و عملکرد آن براساس معیارهای  $RMSE$  و  $R^2$  بین شاخص  $SPI$  مشاهده شده و شاخص‌های سنجش از دور ارزیابی و همبستگی هرکدام از شاخص‌ها تعیین شد. در نهایت الگوریتم جنگل تصادفی برای مدل‌سازی شهرستان مریوان دارای مقادیر  $R^2$  برابر ۰/۹۲۱ و مقدار ضریب تبیین اصلاح شده معادلات رگرسیونی برابر ۰/۹۲۰ بود. همچنین مقادیر  $RMSE$  بین مقادیر واقعی و پیش‌بینی شده روش جنگل تصادفی ۰/۲۸۳ است که نشان می‌دهد این روش نسبت به الگوریتم  $SVR$  با مقادیر  $R^2$  برابر ۰/۹۰ و ضریب تبیین اصلاح شده معادلات رگرسیونی ۰/۸۹ و

خاص اقلیمی و پوشش گیاهی اقدام به انتخاب شاخص مناسب جهت پایش خشکسالی نمود. همچنین در مطالعه حاضر برای اولین بار در کشور برای موضوع پایش خشکسالی سری زمانی از پلتفرم *GEE* که ابزاری بسیار مفید و سریع برای دستیابی به تصاویر ماهواره‌ای مرتبط و انجام تجزیه و تحلیل سنجش از دور تاریخ بزرگ و طولانی مدت به طور موثر است، استفاده شد. این بستر نه تنها برای نظارت بر خشکسالی بلکه برای سایر برنامه‌های نظارت بر محیط زیست نیز می‌توان با موفقیت به داده‌های مختلف سنجنده‌های متنوع دسترسی و پردازش‌های لازم را روی آنها انجام داد [۴۲]. مدل ساخته شده در این تحقیق هنوز محدودیت‌هایی دارد که برای بهبود نیاز به مطالعه بیش‌تری در آینده است. موارد زیر محدودیت‌های اصلی و پیشرفت‌های آتی هستند:

در این تحقیق، از الگوریتم جنگل تصادفی و *SVR* برای استخراج اطلاعات خشکسالی از داده‌های یکپارچه سنجش از دور و توسعه رویکردهای بهینه شده پایش خشکسالی استفاده شده است. به منظور اثبات این که مدل جنگل تصادفی ساخته شده در این تحقیق برتر از ابزارهای پایش خشکسالی دیگر است، مقایسه بیش‌تر بین این مدل و دیگر مدل‌ها از قبیل یادگیری عمیق، شبکه عصبی یا درخت‌های تصمیم‌گیری پیشرفته در تحقیقات آینده مورد نیاز است.

ساخت مدل سری زمانی شاخص‌های مختلف خشکسالی از داده‌های ماهواره‌ای مادیس در بستر *GEE* جهت ارزیابی و توزیع مکانی-زمانی شرایط خشکسالی در دوره‌های بلند مدت مورد نیاز است.

*RMSE* ۰٫۳۱۳، دارای دقت بیش‌تری است. در مطالعه غلام نیا و همکاران (۲۰۱۹)، رفتار مکانی - زمانی پوشش گیاهی در استان کردستان ایران توسط الگوریتم *TIMESAT* مورد تجزیه و تحلیل قرار گرفت. نتایج الگوریتم *TIMESAT* نشان داد که کم‌ترین و بیش‌ترین دامنه‌های *NDVI* در کل دوره زمانی در سال‌های ۲۰۰۸ و ۲۰۱۰ اتفاق افتاده است که با نتایج به دست آمده مطالعه حاضر همخوانی دارد. در پایان عوامل پوشش گیاهی بر اساس اهمیت نسبی و دسته‌بندی آنها در مدل‌سازی برای توسعه شاخص‌های خشکسالی که نمایانگر خشکسالی در شهرستان مریوان هستند، انتخاب شدند و میزان خشکسالی مورد تجزیه و تحلیل قرار گرفت. با توجه به درصد کم انطباق بین *SPI* و *TCI* استفاده از آن برای بررسی تنش در پوشش گیاهی در سال‌های نرمال و مرطوب توصیه نمی‌شود. بنابراین می‌توان بیان کرد که مطالعات صورت گرفته با استفاده از شاخص‌های مختلف با توجه به محل اجرا و شرایط محیطی، نتایج متفاوتی را نشان می‌دهند و در نتیجه امکان مقایسه کلی بین شاخص‌های سنجش از دور و شاخص‌های درجا از نظر صحت داده‌ها و بررسی خشکسالی وجود ندارد. علاوه بر این، اغلب این مطالعات در مناطق با بارش نرمال و بالا صورت گرفته است. لذا با توجه به شرایط خاص آب و هوایی ایران و شرایط اقلیمی مناطق خشک و نیمه خشک نمی‌توان برای مطالعه خشکسالی از روش‌ها و شاخص‌های یکسانی استفاده نموده و نتایج کاملاً مشابهی را انتظار داشت. از این رو باید رابطه هرکدام از شاخص‌های ماهواره‌ای با خشکسالی در هر کدام از مناطق به طور جداگانه مورد مطالعه قرار گیرد و برای هر منطقه بسته به شرایط

## مراجع

[1] S.M. Quiring, "Monitoring Drought: An Evaluation of Meteorological Drought Indices", *Geography Compass*, Vol. 1, pp. 64-88, 2009.

[2] H. Sun, Y. Chen, and H. Sun, "Comparisons and classification system of typical remote sensing indexes for agricultural drought", *Transactions of the Chinese Society of Agricultural Engineering*, Vol. 28, pp. 147-

- 54, 2012.
- [3] H. West, N. Quinn, and M. Horswell, "Remote Sensing of Environment Remote sensing for drought monitoring & impact assessment: Progress, past challenges and future opportunities", *Remote Sensing of Environment*, Vol. 232, pp. 111291, 2019.
- [4] T.B. Mckee, N.J. Doesken, and J. Kleist, "The relationship of drought frequency and duration to time scales", *Proceedings of the 8th Conference on Applied Climatology*, Boston, USA, 1993.
- [5] M.J. Hayes, M.D. Svoboda, D.A. Wihite, and O.V. Vanyarkho, "Monitoring the 1996 Drought Using the Standardized Precipitation Index", *Bulletin of the American meteorological society*, Vol. 80(3), pp. 429-438, 1999.
- [6] H. Karimi, "Studying the Effect of Drought on Vegetation Using MODIS Data (case study: Kurdistan Province)", M.Sc. Thesis, Faculty of Human Sciences at University of Zanjan, 2009 (Persian)
- [7] M. Gholamnia, R. Khandan, S. Bonafoni, and A. Sadeghi, "Spatiotemporal analysis of MODIS NDVI in the semi-arid region of Kurdistan (Iran)", *Remote Sensing*, Vol. 11, pp. 8–12, 2011.
- [8] S. Park, J. Im, E. Jang, and J. Rhee, "Drought assessment and monitoring through blending of multi-sensor indices using machine learning approaches for different climate regions", *Agricultural and Forest Meteorology*, Vol. 216, pp. 157–69, 2016.
- [9] S.S. Park, J. Im, S.S. Park, and J. Rhee, "Drought monitoring using high resolution soil moisture through multi-sensor satellite data fusion over the Korean peninsula", *Agricultural and Forest Meteorology*, Vol. 237, pp. 257–69, 2017.
- [10] M.R. Alizadeh, and M.R. Nikoo, "A fusion-based methodology for meteorological drought estimation using remote sensing data", *Remote sensing of environment*, Vol. 211, pp. 229–47, 2018.
- [11] P. Feng, B. Wang, D.L. Liu, and Q. Yu, "Machine learning-based integration of remotely-sensed drought factors can improve the estimation of agricultural drought in South-Eastern Australia", *Agricultural Systems*, Vol. 173, pp. 303–16, 2019.
- [12] B. Sun, J. Qian, X. Chen, and Q. Zhou, "Comparison and Evaluation of Remote Sensing Indices for Agricultural Drought Monitoring over Kazakhstan", *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vol. 3, pp. 899–903, 2020.
- [13] S. Mehravar, M. Amani, A. Moghimi, F. Dadrass, and S.M. Mirmazloumi, "Temperature-Vegetation-soil Moisture-Precipitation Drought Index (TVMPDI); 21-Year Drought Monitoring in Iran using Satellite Imagery within Google Earth Engine", *Advances in Space Research*, Vol. 68 (11), pp. 4573-4593, 2021.
- [14] F.A. Proadhan, J. Zhang, F. Yao, L. Shi, T. Prasad, and P. Sharma, "Deep Learning for Monitoring Agricultural Drought in South Asia Using Remote Sensing Data", *Remote Sensing*, Vol. 13 (9), 1575, 2021.
- [15] K. Soleimani, S. Darvishi, F. Shokrian, "Analysis of Agricultural Drought using Remote Sensing Indexes (case study: Marivan city)", *The Journal of RS and GIS for Natural Resources*, Vol. 10(2), pp. 15-33, 2019 (Persian)
- [16] D. Zuo, S. Cai, Z. Xu, D. Peng, G. Kan, and W. Sun, "Assessment of meteorological and agricultural droughts using in-situ observations and remote sensing data", *Agricultural Water Management*, Vol. 222, pp. 125–38, 2019.
- [17] A. Poortinga, N. Clinton, D. Saah, P. Cutter, F. Chishtie, and K.N. Markert, "An operational before-after-control-impact (BACI) designed platform for vegetation monitoring at planetary scale". *Remote Sensing*, Vol. 10(5), 760, 2018.
- [18] A. Huete, C. Justice, and H. Liu, "Development of Vegetation and Soil

- Indices for MODIS-EOS*”, *Remote Sensing of Environment*, Vol. 234, pp. 224–34, 1994.
- [19] A. Huete, K. Didan, T. Miura, E.P. Rodriguez, X. Gao, and L.G. Ferreira, “Overview of the radiometric and biophysical performance of the MODIS vegetation indices”, *Remote Sensing of Environment*, Vol. 83, pp. 195–213, 2002.
- [20] M. Belgiu, and L. Drăgu, “Random forest in remote sensing: A review of applications and future directions”, *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 114, pp. 24–31, 2016.
- [21] F.N. Kogan, “Global Drought Watch from Space”, *Bulletin of the American Meteorological Society*, Vol. 78, pp. 621–36, 1994.
- [22] R. Shen, A. Huang, B. Li, and J. Guo, “Construction of a drought monitoring model using deep learning based on multi-source remote sensing data”, *International Journal of Applied Earth Observation and Geoinformation*, Vol. 79(219), pp. 48–57, 2019.
- [23] S. Aksoy, O. Gorucu, and E. Sertel, “Drought monitoring using MODIS derived indices and google earth engine platform”, 2019 8th International Conference on Agro-Geoinformatics (Agro-Geoinformatics 2019), IEEE, 2019.
- [24] N. Oceanic, “Application of Vegetation Index and Brightness Temperature for Drought Detection”, *Advances in Space Research*, Vol. 15(11), pp. 91-100, 1995.
- [25] R.P. Singh, S. Roy, and F. Kogan, “Vegetation and temperature condition indices from NOAA AVHRR data for drought monitoring over India”, *International journal of remote sensing*, Vol. 24(22), pp. 4393–402, 2003.
- [26] M.C. Anderson, J.M. Norman, J.R. Mecikalski, J.A. Otkin, and W.P. Kustas, “A climatological study of evapotranspiration and moisture stress across the continental United States based on thermal remote sensing: 2. Surface moisture climatology”, *Journal of Geophysical Research Atmospheres*, Vol. 112(11), pp. 1–13, 2007.
- [27] R.G. Allen, L.S. Pereira, D. Raes, and M. Smith, “Crop evapotranspiration guidelines for computing crop water requirements”, *FAO Irrigation & drainage Paper 56*, Roma: FAO, Food and Agriculture Organization of the United Nations, 1998.
- [28] T. Jiang, J.L. Gradus, and A.J. Rosellini, “upervised machine learning: a brief primer”, *Behavior Therapy*, Vol. 51, pp. 675-687, 2020.
- [29] W. Yu, M. Shao, M. Ren, H. Zhou, Z. Jiang, and D. Li, “Analysis on spatial and temporal characteristics drought of Yunnan Province”, *Acta Ecologica Sinica*, Vol. 33(6), 3pp. 17–24, 2013.
- [30] A. Elnashar, H. Zeng, B. Wu, N. Zhang, F. Tian, and M. Zhang, “Downscaling TRMM monthly precipitation using google earth engine and google cloud computing”, *Remote Sensing*, Vol. 12(23), pp. 1–22, 2020.
- [31] H. Han, J. Bai, J. Yan, H. Yang, and G. Ma, “A combined drought monitoring index based on multi-sensor remote sensing data and machine learning”, *Geocarto International*, Vol. 36 (10), pp. 1161–1177, 2019.
- [32] E. Boser, N. Vapnik, I.M. Guyon, and T.B. Laboratories, “A Training Algorithm Margin for Optimal Classifiers”, *Proceedings of the fifth annual workshop on Computational learning theory*, Vol. 32, pp. 144–52, 2002.
- [33] C. Cortes, and V. Vapnik, “Support-vector networks”. *Machine learning*, Vol. 20, pp. 273–97, 1995.
- [34] Z. Nikraftar, M. Hasanlou, and M. Esmaeilzadeh, “Novel snow depth retrieval method using time series SSMI passive microwave imagery”. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* -

- ISPRS Archives, Vol. 41, pp. 525–30. 2016.*
- [35] A.J. Smola, and B.S.H. Olkopf, “A tutorial on support vector regression”, *Statistics and Computing, Vol. 14 (3), pp. 199–222, 2004.*
- [36] X. Zhou, X. Zhu, Z. Dong, W. Guo, “Science Direct Estimation of biomass in wheat using random forest regression algorithm and remote sensing data”, *The Crop Journal, Vol. 4(3), pp. 1–8, 2016.*
- [37] L. Breiman, “Random Forests”, *Machine learning, Vol. 45, pp. 5–32, 2001.*
- [38] E. Izquierdo-verdiguier, and R. Zurita-milla, “An evaluation of Guided Regularized Random Forest for classification and regression tasks in remote sensing”, *International Journal of Applied Earth Observation and Geoinformation, Vol. 88, pp. 102051, 2020*
- [39] G. Tsakiris, and H. Vangelis, “Towards a Drought Watch System based on spatial SPI”, *Water Resources Management, Vol. 18(1), pp. 1–12, 2004.*
- [40] P.F. Smith, S. Ganesh, and P. Liu, “A comparison of random forest regression and multiple linear regression for prediction in neuroscience”, *Journal of Neuroscience Methods, Vol. 220(1), pp. 85–91, 2013.*
- [41] Y. Yang, C. Cao, X. Pan, X. Li, and X. Zhu, “Downscaling land surface temperature in an arid area by using multiple remote sensing indices with random forest regression”, *Remote Sensing, Vol. 9(8), pp. 789, 2017.*
- [42] H. Tamiminia, B. Salehi, M. Mahdianpari, L. Quackenbush, S. and Adeli, B. Brisco, “Google Earth Engine for geo-big data applications: A meta-analysis and systematic review”, *ISPRS Journal of Photogrammetry and Remote Sensing, Vol. 164, pp. 152–70, 2020.*



## ***Evaluation of remote sensing-based drought monitoring indexes using support vector regression and random forest models (Case study: Marivan city)***

Jamal Seyedi Ghaldareh <sup>1</sup>, Salman Ahmadi <sup>2\*</sup>, Mehdi Gholamnia <sup>3</sup>

1- Master student of Remote Sensing, Department of Civil Engineering, Faculty of Engineering, University of Kurdistan

2- Assistant Professor, Department of Civil Engineering, Faculty of Engineering, University of Kurdistan

3 - Assistant Professor, Department of Civil Engineering, Faculty of Engineering, Azad University, Sanandaj Branch

### **Abstract**

Drought is a natural and climatic phenomenon that occurs in large areas around the world every year, and its occurrence is caused by the shortage of rainfall and increased evaporation and transpiration at high temperatures. The purpose of this research is evaluating the remote sensing data in drought monitoring for Marivan city and analyzing the spatial-temporal distribution of the drought conditions and identifying its severity. In this study, we used different drought indicators produced from MADIS and TRMM satellite data, which were extracted from Google Earth Engine platform to analyze the drought conditions in Marivan city from February to November for the years 2001 to 2017. In this research, remote sensing indices such as normalized difference index of vegetation, index of vegetation conditions, index of temperature conditions, index of improved vegetation, index of evaporation and transpiration and index of rainfall status were selected as independent variables. Furthermore, the standard rainfall index obtained from meteorological data has been calculated as a dependent variable to evaluate drought conditions. Random forest methods and support vector regression were used to compare the remote sensing data and the ground data and to check the correlation between them and the importance of the remote sensing indicators for drought monitoring. The result of the modeling was obtained using the support vector regression algorithm with the values of the explanatory coefficient of 0.88 and the mean square error of 0.313. The results of the random forest model with the values of the coefficient of explanation of 0.909 and the mean square error of 0.259 indicated the high efficiency of this model. Then, the correlation between the remote sensing indices and the meteorological index was investigated. And PCI, ET, EVI, NDVI indices had the most correlation among the other variables. Therefore, the remote sensing indicators can be used to predict the drought situation in the research area.

**Key words :** Drought, Remote sensing images, Random Forest, Support vector regression.

**Correspondence Address:** Department of Civil Engineering, Faculty of Engineering, University of Kurdistan, Sanandaj, Iran.

**Tel. :** +989126721769

**Email :** s.ahmadi@uok.ac.ir