

کاربرد الگوریتم‌های یادگیری ماشین در شبیه‌سازی خشکسالی مبتنی بر شاخص SPI در دشت مغان

پویا الله‌وردی پور^{۱*}، یعقوب دین پژوه^۲

۱- دانشجوی دکتری تخصصی مهندسی منابع آب، گروه مهندسی آب، دانشکده کشاورزی، دانشگاه تبریز، تبریز، ایران
۲- استاد، گروه مهندسی آب، دانشکده کشاورزی، دانشگاه تبریز، تبریز، ایران

*آدرس پست الکترونیک نویسنده مسئول: a.pouya1996@gmail.com

چکیده فارسی

خشکسالی، به‌عنوان پدیده‌ای پیچیده، تأثیرات عمیقی بر منابع آب و پایداری اکوسیستم‌ها دارد. این پژوهش با پیش‌بینی شاخص بارش استاندارد شده (SPI) با استفاده از مدل‌های یادگیری ماشین، به پیش‌بینی خشکسالی هواشناسی در دشت مغان طی یک دوره ۳۰ ساله (۲۰۲۲-۱۹۹۳) پرداخت. داده‌های بارش، دما، رطوبت نسبی و سرعت باد در مقیاس ماهانه از ایستگاه پارس‌آباد مغان استفاده شدند. مدل‌های هوشمند جنگل تصادفی (RF)، درختان اضافی (ET)، ماشین بردار پشتیبان (SVM) و شبکه عصبی پرسپترون چندلایه (MLP) با معیارهای ضریب همبستگی (r)، جذر میانگین مربعات خطاها (RMSE)، و ضریب نش-ساتکلیف (NSE) ارزیابی شدند. نتایج نشان داد که در مرحله آموزش، مدل SVM با $r = 0.95$ و $RMSE = 0.30$ و $NSE = 0.83$ بهترین عملکرد را داشت. در مرحله صحت‌سنجی نیز این مدل با $r = 0.88$ و $RMSE = 0.39$ و $NSE = 0.81$ دارای بهترین عملکرد بود. رطوبت نسبی برای مدل‌های RF، ET، SVM و ترکیب دما و رطوبت نسبی برای مدل MLP به‌عنوان بهترین سناریو ورودی شناسایی شدند. تحلیل پراکنش شاخص SPI نشان‌دهنده دقت بالای مدل SVM در ثبت شرایط شدید خشکسالی (مانند $SPI = -1.85$) و ترسالی (مانند $SPI = 1.85$) بود. این یافته‌ها بر توانایی مدل‌های یادگیری ماشین، به‌ویژه SVM، در پیش‌بینی خشکسالی و پشتیبانی از مدیریت منابع آب تأکید دارند.

کلید واژه‌ها: بردار پشتیبان، درختان اضافی، شبکه عصبی، منابع آب، هوش مصنوعی، یادگیری ماشین.

مقدمه

خشکسالی یکی از پیچیده‌ترین و مخرب‌ترین پدیده‌های طبیعی است که تأثیرات عمیقی بر منابع آب، تولیدات کشاورزی، امنیت غذایی و پایداری اکوسیستم‌ها در سراسر جهان دارد (رنجبر و همکاران، ۱۴۰۳). در مناطق کشاورزی محور مانند دشت مغان، که به‌عنوان یکی از مهم‌ترین مناطق تولید محصولات کشاورزی در ایران شناخته می‌شود، اثرات خشکسالی به دلیل وابستگی شدید به منابع آب سطحی و زیرزمینی برای آبیاری و تولید محصولات کشاورزی، به‌ویژه برجسته است. تغییرات اقلیمی، با افزایش دما، تغییر الگوهای بارش و تشدید رویدادهای شدید هواشناسی، چالش‌های جدیدی را در پیش‌بینی و مدیریت خشکسالی ایجاد کرده است (IPCC, 2021). این عوامل، همراه با افزایش تقاضای آب در بخش‌های کشاورزی، صنعتی و شهری، ضرورت توسعه ابزارهای دقیق و قابل اعتماد برای ارزیابی و پیش‌بینی خشکسالی را بیش از پیش آشکار می‌سازد (راهواره و همکاران، ۱۴۰۱). شاخص استاندارد بارش (SPI)، که توسط مک‌کی و همکاران (۱۹۹۳) معرفی شد، به دلیل انعطاف‌پذیری در تحلیل خشکسالی در مقیاس‌های زمانی مختلف (از ماهانه تا سالانه) و توانایی در شناسایی شدت، مدت و گستره خشکسالی، به یکی از پرکاربردترین شاخص‌ها در مطالعات هیدرولوژیکی تبدیل شده است (خلیلی و همکاران، ۱۴۰۲). با این حال، مدل‌سازی دقیق SPI در مناطق پیچیده‌ای مانند دشت مغان، که تحت تأثیر عوامل متعددی از جمله تغییرات فصلی بارش، توپوگرافی متنوع و فعالیت‌های انسانی قرار دارد، چالش‌برانگیز است. روش‌های سنتی مدل‌سازی، مانند مدل‌های آماری خطی، اغلب در ثبت روابط غیرخطی و دینامیک‌های پیچیده بارش ناکام می‌مانند، که این امر منجر به کاهش دقت پیش‌بینی‌ها می‌شود (Sheffield et al., 2012).

در دهه‌های اخیر، پیشرفت‌های چشمگیر در حوزه یادگیری ماشین، امکان مدل‌سازی روابط غیرخطی و پیچیده در داده‌های هواشناسی و هیدرولوژیکی را فراهم کرده است. الگوریتم‌های یادگیری ماشین، از جمله شبکه‌های عصبی مصنوعی (ANN)، ماشین‌های بردار پشتیبان (SVM) و جنگل‌های تصادفی (RF)، به دلیل توانایی در یادگیری الگوهای پنهان و پیش‌بینی دقیق‌تر در مقایسه با روش‌های سنتی، به‌طور فزاینده‌ای در مطالعات خشکسالی مورد استفاده قرار گرفته‌اند (Hao et al., 2018). این روش‌ها به‌ویژه در مناطقی مانند دشت مغان، که دارای تنوع اقلیمی و داده‌های بلندمدت هواشناسی است، می‌توانند به بهبود پیش‌بینی‌ها و ارائه بینش‌های عمیق‌تر در مورد الگوهای خشکسالی کمک کنند.

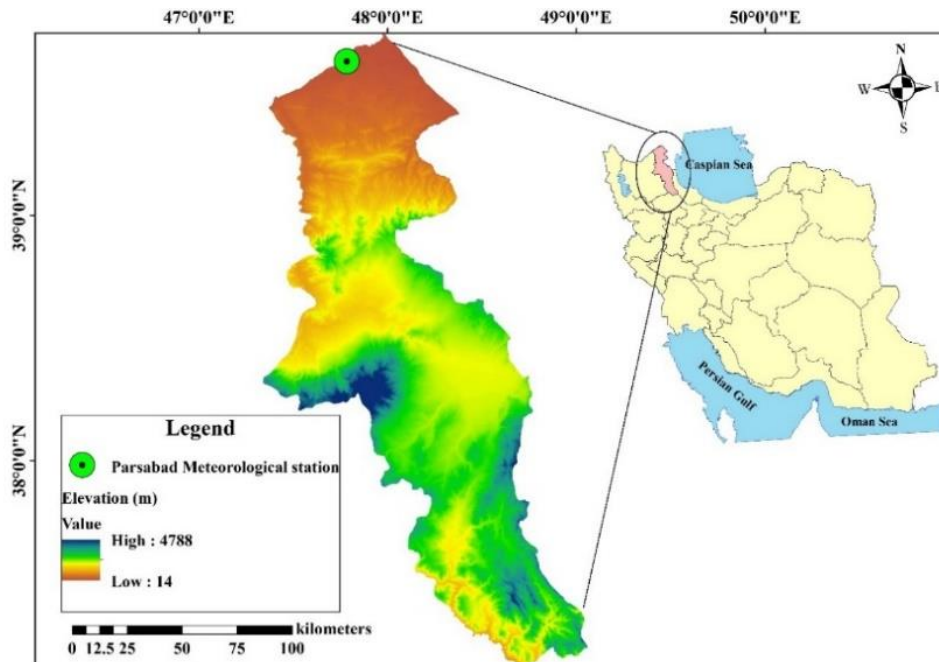
این مطالعه با هدف مدل‌سازی شاخص استاندارد بارش (SPI) در دشت مغان با استفاده از الگوریتم‌های پیشرفته یادگیری ماشین شامل: RF, SVM, MLP و RT انجام شده است. با بهره‌گیری از داده‌های هواشناسی بلندمدت و مقایسه عملکرد مدل‌های مختلف یادگیری ماشین، این پژوهش به دنبال شناسایی الگوهای خشکسالی، و ارائه راهکارهای علمی برای مدیریت پایدار منابع آب در این حوضه حیاتی است.

مواد و روش‌ها

منطقه مورد مطالعه و داده‌های مورد استفاده

دشت مغان در شمال استان اردبیل و در همسایگی مرز جمهوری آذربایجان قرار دارد و به‌عنوان یکی از قطب‌های اصلی کشاورزی کشور شناخته می‌شود. این ناحیه که بیش از ۲۰۰ هزار هکتار وسعت دارد، سهم قابل‌توجهی در تولیدات زراعی و تأمین غذای کشور بر عهده دارد. وجود خاک‌های پرقدردت، دسترسی به آب رودخانه ارس و شرایط آب‌وهوایی مناسب، زمینه را برای توسعه فعالیت‌های کشاورزی در شمال‌غرب ایران فراهم کرده است. با این حال، تغییر در الگوهای بارش و دگرگونی‌های اقلیمی سال‌های اخیر، پایداری تولیدات کشاورزی منطقه را با مسائل تازه‌ای روبه‌رو کرده است.

در این تحقیق، داده‌های آب‌وهوایی ایستگاه همدید پارس‌آباد مغان—که بازتاب‌دهنده شرایط اقلیمی بخش وسیعی از دشت است—مورد بررسی قرار گرفته است. این ایستگاه در حدود عرض ۳۹ درجه و ۳۶ دقیقه شمالی و طول ۴۷ درجه و ۴۷ دقیقه شرقی و در ارتفاع حدود ۷۲ متر از سطح دریا واقع شده است (شکل ۱). طبق آمار بلندمدت، میانگین بارش سالانه آن نزدیک به ۲۶۱ میلی‌متر و میانگین دمای سالانه حدود ۵/۱۵ درجه سانتی‌گراد گزارش شده است. ترکیب این ویژگی‌های اقلیمی و تغییرات فصلی بارندگی و دما، دشت مغان را به فضایی مناسب برای بررسی روندهای اقلیمی تبدیل کرده است. گستردگی دوره ثبت داده‌ها و دقت اندازه‌گیری‌های ایستگاه پارس‌آباد نیز زمینه قابل‌اعتمادی برای تحلیل‌های هیدرولوژیکی و ارزیابی وضعیت خشکسالی فراهم می‌آورد.



شکل ۱- موقعیت جغرافیایی منطقه مورد مطالعه، دشت مغان، استان اردبیل، ایران (دین پژوه و همکاران، ۱۴۰۴).

در این پژوهش، به منظور ارزیابی خشکسالی هواشناسی در دشت مغان، از شاخص استاندارد بارش (SPI) در مقیاس زمانی یک‌ماهه استفاده شده است. داده‌های بارش مورد استفاده در این مطالعه، از ایستگاه همدید پارس‌آباد مغان در یک بازه زمانی ۳۰ ساله (۱۹۹۳-۲۰۲۲) گردآوری شده‌اند. این دوره طولانی‌مدت، امکان تحلیل دقیق الگوهای بارش و شناسایی دوره‌های خشکسالی و ترسالی را فراهم می‌کند. برای مدل‌سازی شاخص SPI، از الگوریتم‌های پیشرفته یادگیری ماشین شامل شبکه عصبی مصنوعی پرسپترون چندلایه (MLP)، ماشین بردار پشتیبان (SVM)، جنگل تصادفی (RF)، و درختان اضافی (ET) استفاده شده است. متغیرهای اقلیمی دما، رطوبت نسبی، و سرعت باد به‌عنوان متغیرهای ورودی این مدل‌ها انتخاب شده‌اند، زیرا این متغیرها به‌طور مستقیم بر الگوهای بارش و شدت خشکسالی تأثیر می‌گذارند (Hao et al., 2018). مشخصات آماری متغیرهای مورد استفاده شامل بارش (میلی‌متر)، دما (درجه سانتی‌گراد)، رطوبت نسبی (درصد) و سرعت باد (متر بر ثانیه) در جدول ۱ ارائه شده است، که مبنای تحلیل‌های بعدی در این مطالعه قرار گرفته است.

جدول ۱- مشخصات آماری متغیرهای اقلیمی دشت مغان در مقیاس ماهانه در دوره ۱۹۹۳-۲۰۲۲

متغیر	واحد	کمینه	بیشینه	میانگین	انحراف معیار
بارش	میلی‌متر	0	65	20.14	16.74
دما	درجه سانتی‌گراد	-1.40	29.14	14.99	8.63
رطوبت نسبی	درصد	51.43	86.42	71.46	7.78
سرعت باد	متر بر ثانیه	0.45	4.08	2.03	0.86

این مشخصات آماری نشان‌دهنده تنوع قابل‌توجه در متغیرهای اقلیمی دشت مغان است که پیچیدگی‌های مدل‌سازی خشکسالی را برجسته می‌کند. انحراف معیار بالای بارش و دما نشان‌دهنده تغییرات اقلیمی قابل‌توجه در منطقه است که نیاز به مدل‌های یادگیری ماشین پیشرفته مانند MLP، SVM، RF و ET را برای ثبت روابط غیرخطی بین متغیرهای ورودی و SPI تأیید می‌کند. این تحلیل اولیه، مبنای مناسبی برای ارزیابی عملکرد مدل‌های یادگیری ماشین در پیش‌بینی خشکسالی فراهم می‌کند و بر اهمیت انتخاب متغیرهای ورودی مرتبط با الگوهای اقلیمی تأکید دارد.

شاخص خشکسالی SPI

شاخص استاندارد بارش (Standardized Precipitation Index - SPI)، که توسط مک‌کی و همکاران (۱۹۹۳) معرفی شد، یکی از پرکاربردترین ابزارهای هواشناسی برای پایش و ارزیابی خشکسالی است. این شاخص به دلیل سادگی، انعطاف‌پذیری و توانایی تحلیل خشکسالی در مقیاس‌های زمانی مختلف (از یک ماه تا چندین سال)، در مطالعات اقلیمی به‌طور گسترده استفاده می‌شود. SPI بر اساس استانداردسازی داده‌های بارش نسبت به توزیع آماری آن‌ها در یک دوره زمانی بلندمدت (معمولاً حداقل ۳۰ سال) عمل می‌کند و شدت خشکسالی یا ترسالی را با تبدیل مقادیر بارش به توزیع نرمال استاندارد (با میانگین صفر و انحراف معیار یک) اندازه‌گیری می‌کند. این استانداردسازی امکان مقایسه خشکسالی در مناطق مختلف با اقلیم‌های متنوع را فراهم می‌سازد. محاسبه SPI شامل جمع‌آوری داده‌های بارش، فیت کردن آن‌ها به یک توزیع آماری (معمولاً توزیع گاما)، تبدیل به مقادیر احتمال تجمعی، و نگاهش به مقادیر استاندارد نرمال است. مزایای SPI شامل سادگی، نیاز به تنها داده‌های بارش، قابلیت مقایسه بین مناطق مختلف، و حساسیت به تغییرات کوتاه‌مدت و بلندمدت بارش است (Shi et al., 2022).

ترکیب SPI با مدل‌های یادگیری ماشین، می‌تواند دقت پیش‌بینی خشکسالی را بهبود بخشد. این رویکرد در دشت مغان امکان تحلیل روابط غیرخطی بین متغیرهای اقلیمی و SPI را فراهم کرده و به شناسایی الگوهای پیچیده خشکسالی کمک می‌کند. برای مثال، در مناطق مشابه مانند دشت مغان، داده‌های بارش با میانگین سالانه ۲۶۱ میلی‌متر و انحراف معیار بالا، نشان‌دهنده تغییرات قابل‌توجه اقلیمی است که SPI به‌خوبی آن‌ها را ثبت می‌کند. این مدل‌سازی به سیاست‌گذاران در توسعه استراتژی‌های مدیریت منابع آب و کاهش اثرات خشکسالی کمک می‌کند. برای محاسبه این شاخص از رابطه (۱) استفاده می‌شود.

$$SPI = \frac{P_i - \bar{P}}{\sigma} \quad \text{رابطه (۱)}$$

که در آن، SPI شاخص خشکسالی، P_i میانگین بارندگی ماه مورد نظر در هر سال، \bar{P} میانگین بارندگی ماه مورد نظر در کل دوره و σ انحراف معیار برای کل دوره است. مقادیر شاخص SPI نیز به صورت جدول ۲ تفسیر می‌شوند.

جدول ۲- طبقه‌بندی خشکسالی بر اساس شاخص SPI (مک کی و همکاران، ۱۹۹۳)

SPI	Category
$\geq(2)$	Extreme wet
$(1.5)-(1.99)$	Severe wet
$(1)-(1.49)$	Middle wet
$(0)-(0.99)$	Little wet
$(-0.99)-(0)$	Little drought
$(-1.49)-(-1)$	Middle drought
$(-1.99)-(-1.5)$	Severe drought
$\leq(-2)$	Extreme drought

مدل‌های یادگیری ماشین

ماشین بردار پشتیبان (SVM)

الگوریتم ماشین بردار پشتیبان (SVM) یکی از روش‌های مطرح در یادگیری نظارتی است که در کاربردهای طبقه‌بندی و همچنین رگرسیون، از جمله تحلیل‌های هیدرولوژیکی و پیش‌بینی شاخص‌هایی مانند SPI برای بررسی خشکسالی، جایگاه ویژه‌ای دارد (Vapnik, 1995). اساس کار این روش بر یافتن مرزی استوار است که بتواند داده‌ها را با بیشترین فاصله از نقاط حساس یا همان بردارهای پشتیبان جدا کند؛ ویژگی‌ای که موجب پایداری مدل و کاهش تمایل آن به بیش‌برازش می‌شود.

در حالت رگرسیونی این روش (SVR)، هدف ساخت تابعی است که بتواند مقادیر خروجی را با حداقل خطا و در قالب یک نوار تحمل مشخص برآورد کند. در فضای ویژگی‌ها، SVM تلاش می‌کند سطحی را تعیین کند که گروه‌های مختلف داده را از هم تفکیک کند. هنگامی که داده‌ها رفتاری غیرخطی دارند، از ترنند کرنل استفاده می‌شود تا داده‌ها به فضایی با ابعاد بالاتر منتقل شوند؛ فضایی که در آن جداسازی خطی امکان‌پذیر می‌شود. کرنل‌های متنوعی برای این کار به کار می‌روند، از جمله کرنل‌های خطی، چندجمله‌ای و تابع پایه شعاع (RBF)، که کرنل RBF به دلیل انعطاف‌پذیری مناسب در نمایش روابط پیچیده میان متغیرها به طور گسترده مورد استفاده قرار می‌گیرد (Pai and Hong, 2007).

پس از اتمام مرحله آموزش، مدل ایجادشده می‌تواند داده‌های جدید را پیش‌بینی کند و برای برآورد مقادیری مانند SPI در مجموعه آزمایشی به کار گرفته می‌شود.

جنگل تصادفی (RF)

روش جنگل تصادفی (Random Forest) از جمله الگوریتم‌های پرکاربرد در یادگیری ماشین است که بر پایه مجموعه‌ای از درخت‌های تصمیم عمل می‌کند و در هر دو حوزه رگرسیون و طبقه‌بندی عملکرد بسیار مناسبی دارد. در مطالعات هیدرولوژیکی، به‌ویژه در پیش‌بینی شاخص‌هایی مانند SPI برای پایش خشکسالی، این مدل یکی از گزینه‌های قابل‌اعتماد محسوب می‌شود. قدرت اصلی جنگل تصادفی در آن است که با اتکا به پیش‌بینی‌های تعداد زیادی درخت تصمیم‌گیری، خطای مدل کاهش یافته و پایداری آن افزایش می‌یابد؛ موضوعی که موجب مقاومت بیشتر آن در برابر بیش‌برازش می‌شود (Breiman, 2001). این رویکرد بر اساس چارچوب یادگیری جمعی از نوع بگینگ توسعه یافته است. در این فرآیند، برای ساخت هر درخت، داده‌ها به صورت بوت‌استرپ نمونه‌گیری می‌شوند و هر درخت تنها با بخشی از داده‌ها و مجموعه‌ای تصادفی از ویژگی‌ها آموزش می‌بیند. انتخاب تصادفی ویژگی‌ها در هر گره، موجب تنوع بالای درخت‌ها و در نتیجه کاهش هم‌بستگی بین آن‌ها می‌شود. درخت‌ها معمولاً بدون هرس تا عمق کامل رشد می‌کنند و در پایان، خروجی تک‌تک آن‌ها تجمیع شده و نتیجه نهایی مدل حاصل می‌شود (Cutler et al., 2012; Yin, 2023).

درختان اضافی (ET)

مدل Extra Trees یا Extremely Randomized Trees یک الگوریتم یادگیری جمعی مبتنی بر درخت‌های تصمیم است که برای افزایش سرعت و بهبود کارایی نسبت به جنگل تصادفی معرفی شده است (Geurts et al., 2006). این مدل با استفاده از تصادفی‌سازی بیشتر در انتخاب ویژگی‌ها و نقاط تقسیم، درخت‌ها را سریع‌تر و با تنوع بالاتر ایجاد می‌کند و به همین دلیل در برابر بیش‌برازش مقاوم‌تر است. Extra Trees در مسائل رگرسیون و طبقه‌بندی، به‌ویژه در داده‌های پیچیده یا نویزی مانند داده‌های هواشناسی، عملکرد بسیار خوبی دارد و در مدل‌سازی‌هایی نظیر پیش‌بینی شاخص SPI برای ارزیابی خشکسالی به‌طور گسترده استفاده می‌شود.

شبکه‌های عصبی مصنوعی از نوع پرسپترون چندلایه (MLP)

شبکه عصبی MLP یا پرسپترون چندلایه یکی از رایج‌ترین مدل‌های یادگیری نظارت‌شده است که در مسائل رگرسیون و طبقه‌بندی، از جمله پیش‌بینی شاخص SPI برای تحلیل خشکسالی، کاربرد فراوان دارد. این نوع شبکه با داشتن لایه‌های پنهان قادر است روابط غیرخطی و پیچیده میان متغیرهای اقلیمی و هیدرولوژیکی را به‌خوبی مدل‌سازی کند و برای داده‌های چندبعدی عملکرد بسیار مناسبی دارد. پژوهش‌ها نیز نشان داده‌اند که MLP در بسیاری از موارد دقت بالاتری نسبت به روش‌های آماری سنتی مانند رگرسیون خطی ارائه می‌دهد (Jain et al., 1999).

طراحی سناریوهای ورودی مدل‌های یادگیری ماشین

در این پژوهش، به‌منظور بررسی تأثیر ترکیب‌های مختلف متغیرهای اقلیمی بر دقت پیش‌بینی شاخص SPI در ارزیابی خشکسالی هواشناسی، هفت سناریوی ورودی با استفاده از سه متغیر اقلیمی کلیدی شامل دما (T)، رطوبت نسبی (RH)، و سرعت باد (U) طراحی و مورد ارزیابی قرار گرفتند. این متغیرها به دلیل ارتباط مستقیم و غیرمستقیم با الگوهای بارش و تأثیرات آن‌ها بر شدت و مدت خشکسالی انتخاب شدند. دما به‌طور مستقیم بر نرخ تبخیر و تعرق تأثیر می‌گذارد، که می‌تواند شدت خشکسالی را تشدید کند (Vicente-Serrano et al., 2010). رطوبت نسبی، به‌عنوان شاخصی از میزان رطوبت موجود در جو، بر الگوهای بارش و تبخیر اثر می‌گذارد و نقش مهمی در پویایی‌های خشکسالی دارد. سرعت باد نیز با تأثیر بر نرخ تبخیر و انتقال رطوبت جوی، به‌ویژه در مناطق نیمه‌خشک مانند دشت مغان، می‌تواند بر شدت خشکسالی تأثیر بگذارد. انتخاب این متغیرها با توجه به مشخصات آماری آن‌ها (جدول ۱) و تنوع بالای اقلیمی در منطقه مورد مطالعه، منطقی و توجیه‌پذیر است.

انتخاب ترکیب‌های مختلف ورودی امکان تحلیل نقش هر متغیر و تعاملات آن‌ها در بهبود عملکرد مدل‌های یادگیری ماشین، از جمله پرسپترون چندلایه (MLP)، ماشین بردار پشتیبان (SVM)، جنگل تصادفی (RF) و درختان اضافی (ET)، را فراهم می‌کند. جدول ۳ سناریوهای مختلف ورودی به مدل‌ها را در مطالعه حاضر نشان می‌دهد و به‌عنوان مبنای تحلیل‌های بعدی برای ارزیابی عملکرد مدل‌ها مورد استفاده قرار گرفته است. هفت سناریوی ورودی طراحی شده به‌گونه‌ای هستند که امکان ارزیابی تأثیر هر متغیر به‌تنهایی (سناریوهای ۱ تا ۳)، ترکیب‌های دوتایی (سناریوهای ۴ تا ۶)، و ترکیب کامل متغیرها (سناریو ۷) را فراهم می‌کنند. این رویکرد چندگانه به ارزیابی تعاملات بین متغیرها (مانند اثر هم‌افزایی دما و رطوبت نسبی) و تحلیل حساسیت مدل‌های یادگیری ماشین نسبت به متغیرهای ورودی کمک می‌کند و امکان شناسایی ترکیب بهینه برای پیش‌بینی SPI را فراهم می‌سازد.

جدول ۳- سناریوهای مختلف برای ترکیب‌های ورودی در مطالعه حاضر

Scenario	Inputs combination
1	T
2	RH
3	U
4	T, RH
5	T, U
6	RH, U
7	T, RH, U

معیارهای ارزیابی

برای ارزیابی عملکرد مدل‌های یادگیری ماشین مورد استفاده در این پژوهش، از سه معیار آماری پرکاربرد شامل ضریب همبستگی پیرسون (r)، جذر میانگین مربعات خطاها (Root Mean Square Error - RMSE)، و ضریب نش-ساتکلِف (Nash-Sutcliffe Efficiency - NSE) استفاده شده است. این معیارها به دلیل توانایی در سنجش دقت پیش‌بینی‌ها و تطابق بین مقادیر مشاهداتی و شبیه‌سازی‌شده، به‌طور گسترده در مطالعات هیدرولوژیکی و اقلیمی، از جمله مدل‌سازی شاخص SPI، مورد استفاده قرار می‌گیرند (Moriasi et al., 2007). این معیارها به دلیل تکمیل یکدیگر در ارزیابی جنبه‌های مختلف عملکرد مدل، انتخاب شده‌اند. ضریب r میزان قدرت و جهت رابطه خطی بین دو متغیر X (مانند مقادیر پیش‌بینی‌شده) و Y (مانند مقادیر مشاهداتی) را اندازه‌گیری می‌کند، RMSE خطای مطلق را اندازه‌گیری می‌کند و NSE کارایی مدل را نسبت به یک معیار پایه ارزیابی می‌کند. در مدل‌سازی شاخص SPI، این معیارها امکان مقایسه دقیق بین مدل‌های یادگیری ماشین را فراهم می‌کنند. این معیارها در روابط ۶ تا ۸ ارائه شده‌اند.

$$r = \frac{\sum_{i=1}^N (O_i - \bar{O}_i)(P_i - \bar{P}_i)}{\sqrt{\sum_{i=1}^N (O_i - \bar{O}_i)^2 \sum_{i=1}^N (P_i - \bar{P}_i)^2}} \quad (6)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (P_i - O_i)^2}{N}} \quad (7)$$

$$NSE = 1 - \frac{\sum_{i=1}^N (O_i - P_i)^2}{(O_i - \bar{O}_i)^2} \quad (8)$$

در این روابط، پارامتر $O(i)$ مقادیر مشاهداتی (واقعی)، $P(i)$ مقادیر شبیه‌سازی شده، \bar{O} میانگین مقادیر مشاهداتی، \bar{P} میانگین مقادیر شبیه‌سازی شده و N تعداد مشاهدات است. مقدار r بین -۱ و +۱ قرار می‌گیرد، که مقدار ۱ نشان‌دهنده همبستگی خطی کامل و مثبت مقادیر پیش‌بینی‌شده با مشاهدات و عملکرد قوی مدل، مقدار -۱ همبستگی خطی کامل و منفی و $r = 0$ نشان‌دهنده عدم وجود همبستگی بین مقادیر پیش‌بینی‌شده با مشاهدات است. RMSE مقیاسی از خطای پیش‌بینی است و مقادیر کمتر آن نشان‌دهنده دقت بالاتر مدل است. این معیار به واحد داده‌های مشاهداتی وابسته است و امکان مقایسه مستقیم بین مدل‌ها را در یک مجموعه داده فراهم می‌کند. NSE مقیاسی از کارایی مدل نسبت به یک مدل پایه (میانگین مشاهدات) است. مقادیر NSE نزدیک به ۱ نشان‌دهنده عملکرد عالی مدل، مقادیر بین ۰/۵ تا ۱ نشان‌دهنده عملکرد خوب، و مقادیر منفی نشان‌دهنده عملکرد ضعیف‌تر از میانگین مشاهدات است.

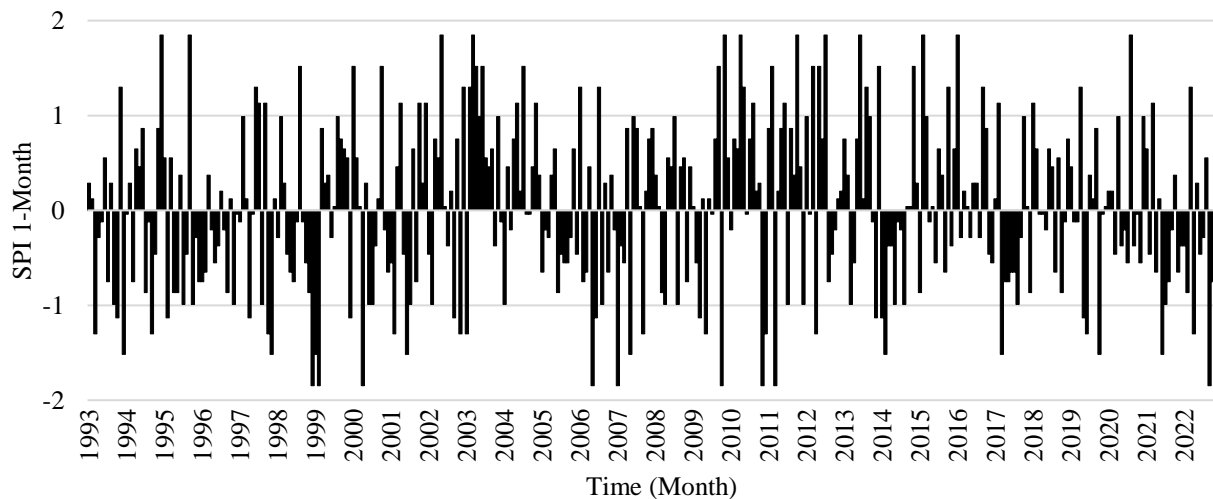
نتایج

شکل ۲ نمودار شاخص SPI در مقیاس زمانی یک‌ماهه را ارائه می‌دهد که به‌عنوان ابزاری کلیدی برای پیش‌بینی خشکسالی هواشناسی در دشت مغان مورد استفاده قرار گرفته است. این شاخص، با استانداردسازی داده‌های بارش ماهانه نسبت به توزیع بلندمدت آن‌ها، شدت و فراوانی دوره‌های خشکسالی و ترسالی را در منطقه مشخص می‌کند. داده‌های مورد استفاده در این مطالعه از ایستگاه همدید پارس‌آباد مغان در دوره زمانی ۳۰ ساله (۱۹۹۳-۲۰۲۲) استخراج شده‌اند، که شامل ۳۶۰ ماه است. این بازه زمانی طولانی‌مدت امکان تحلیل دقیق الگوهای خشکسالی و شناسایی تغییرات اقلیمی در این منطقه کشاورزی‌محور را فراهم می‌کند.

دشت مغان در طول این دوره مورد مطالعه، دوره‌های متعدد خشکسالی و ترسالی را تجربه کرده است. شاخص SPI یک‌ماهه در ۵۰٪ از ماه‌های این دوره (۱۸۰ ماه از ۳۶۰ ماه) مقادیر منفی داشته است، که نشان‌دهنده وقوع خشکسالی هواشناسی در نیمی از دوره مورد بررسی است. این نسبت بالا بیانگر آسیب‌پذیری اقلیمی منطقه و تأثیرات بالقوه آن بر کشاورزی و منابع آب است. مقادیر منفی SPI نشان‌دهنده کاهش بارش نسبت به میانگین بلندمدت بوده و می‌تواند به کاهش رطوبت خاک و محدودیت منابع آب سطحی و زیرزمینی منجر شود (Hayes et al., 2011). شدیدترین دوره‌های خشکسالی، با مقدار SPI برابر با -۱.۸۵، در طبقه‌بندی خشکسالی بسیار شدید (Severe Drought - SD) قرار می‌گیرند. وقوع این دوره‌های بسیار شدید در ماه‌های مختلف، به‌ویژه در فصول گرم و سرد، نشان‌دهنده تنوع زمانی خشکسالی در دشت مغان است. این تنوع می‌تواند به تغییرات فصلی در الگوهای بارش، تحت تأثیر عواملی مانند جبهه‌های جوی، مرتبط باشد.

توزیع زمانی دوره‌های خشکسالی (۱۸۰ ماه با SPI منفی) نشان‌دهنده پراکندگی یکنواخت خشکسالی در طول دوره ۳۰ ساله است، که بیانگر فقدان یک الگوی فصلی مشخص است. این امر می‌تواند به پیچیدگی اقلیمی دشت مغان، که تحت تأثیر موقعیت جغرافیایی آن در مجاورت دریای خزر و منطقه قفقاز در مرز ایران و جمهوری آذربایجان و همچنین نفوذ جبهه‌های سرد از سیبری قرار دارد، نسبت داده شود. وقوع ۹ ماه با خشکسالی بسیار شدید ($SPI = -1.85$) در بازه‌های زمانی پراکنده (۱۹۹۸ تا ۲۰۲۲) نشان‌دهنده تأثیرات متناوب تغییرات اقلیمی، مانند کاهش بارش‌های فصلی یا افزایش دما، است که نرخ تبخیر را افزایش داده و شدت خشکسالی را تشدید می‌کند (IPCC, 2021).

دشت مغان به‌عنوان یکی از مراکز اصلی تولید کشاورزی در ایران، به شدت به بارش و منابع آب وابسته است. وقوع خشکسالی در ۵۰٪ از ماه‌های دوره مورد مطالعه، به‌ویژه دوره‌های بسیار شدید، می‌تواند پیامدهای قابل‌توجهی برای تولید محصولات کشاورزی، مانند گندم و جو، و پایداری منابع آب داشته باشد. مقادیر SPI منفی، به‌ویژه در محدوده خشکسالی شدید و بسیار شدید، با کاهش دسترسی به آب برای آبیاری و افت سطح آب زیرزمینی مرتبط است، که این امر نیاز به استراتژی‌های مدیریت منابع آب، مانند آبیاری کارآمد و ذخیره‌سازی آب، را برجسته می‌کند.



شکل ۲- نمودار شاخص خشکسالی SPI یک‌ماهه در دشت مغان در دوره ۲۰۲۲-۱۹۹۳.

ارزیابی عملکرد مدل‌های یادگیری ماشین برای پیش‌بینی شاخص SPI

نتایج ارزیابی عملکرد مدل‌ها در مراحل آموزش و صحت‌سنجی، با در نظر گرفتن سناریوهای مختلف ورودی معرفی شده در جدول ۲ (شامل ترکیب‌های تک‌متغیره، دومتغیره، و سه‌متغیره از دما (T)، رطوبت نسبی (RH)، و سرعت باد (U))، در جدول ۳ ارائه شده است. این نتایج بهترین ساختار ورودی برای هر مدل را مشخص کرده و امکان مقایسه عملکرد مدل‌ها را فراهم می‌کند.

نتایج جدول ۳ نشان‌دهنده عملکرد قوی مدل‌های RF، ET، SVM و SVM در مقایسه با MLP است. مدل SVM با $NSE = 0.806$ ، $RMSE = 0.326$ ، و $r = 0.91$ در مرحله صحت‌سنجی، بهترین عملکرد را در میان مدل‌ها ارائه می‌دهد. این امر می‌تواند به توانایی SVM در مدل‌سازی روابط غیرخطی با استفاده از کرنل مناسب RBF و بهینه‌سازی حاشیه حداکثری نسبت داده شود (Vapnik, 1995). مدل‌های RF و ET نیز با مقادیر NSE به ترتیب ۰/۷۸۴ و ۰/۷۹۱ و به ترتیب ۰/۸۵ و ۰/۸۷ در مرحله صحت‌سنجی، عملکرد قابل‌قبولی دارند، که نشان‌دهنده مقاومت این مدل‌های مبتنی بر درخت در برابر بیش‌برازش و توانایی آن‌ها در مدیریت داده‌های اقلیمی نویزی است (Breiman, 2001). در مقابل، مدل MLP با $NSE = 0.773$ ، $RMSE = 0.461$ ، و $r = 0.82$ عملکرد ضعیف‌تری نسبت به سایر مدل‌ها دارد، که ممکن است به حساسیت آن به تنظیم‌های اولیه (مانند تعداد لایه‌ها و نوروها) یا نیاز به داده‌های آموزشی بیشتر مربوط باشد (Haykin, 1999).

در مورد تأثیر سناریوهای ورودی، نکته قابل‌توجه در جدول ۳، انتخاب سناریو ۲ (رطوبت نسبی - RH) به‌عنوان بهترین ساختار ورودی برای مدل‌های RF، ET، و SVM است. این امر نشان‌دهنده نقش کلیدی رطوبت نسبی در پیش‌بینی شاخص SPI در دشت مغان است. رطوبت نسبی، به‌عنوان شاخصی از

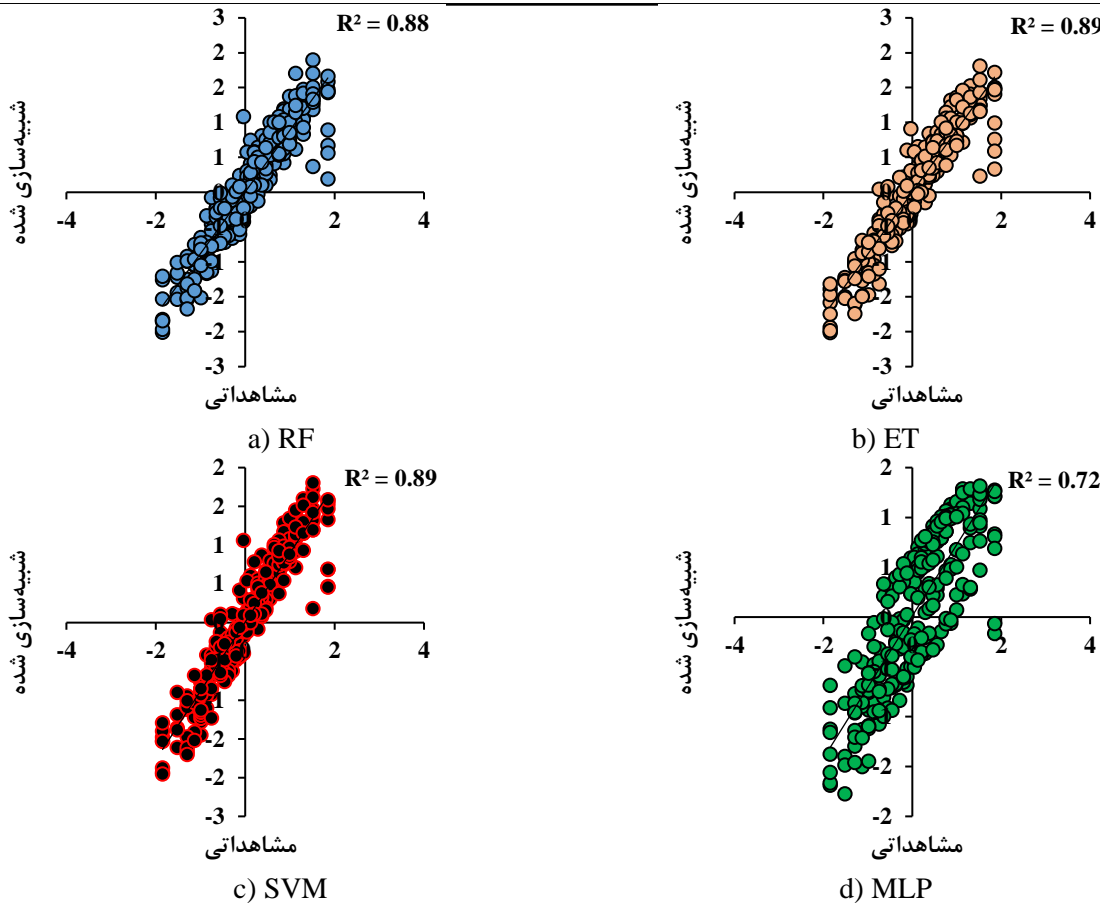
میزان رطوبت موجود در جو، با الگوهای بارش و نرخ تبخیر ارتباط نزدیکی دارد و می‌تواند تأثیر قابل‌توجهی بر شدت خشکسالی داشته باشد (Vicente-Serrano et al., 2010). در مقابل، مدل MLP بهترین عملکرد را با سناریو ۴ (ترکیب دما و رطوبت نسبی - T, RH) نشان می‌دهد، که بیانگر توانایی این مدل در استفاده از تعاملات بین متغیرها برای ثبت روابط غیرخطی پیچیده‌تر است. عدم انتخاب سناریو ۷ (ترکیب کامل T, RH, U) به‌عنوان بهترین سناریو برای هیچ‌کدام از مدل‌ها، ممکن است به افزایش پیچیدگی مدل و خطر بیش‌برازش یا حضور نویز در متغیر سرعت باد (U) نسبت داده شود، به‌ویژه با توجه به انحراف معیار پایین این متغیر (۰/۸۶ متر بر ثانیه در جدول ۱).

تمامی مدل‌ها در مرحله آموزش عملکرد بهتری (NSE و r بالاتر، RMSE پایین‌تر) نسبت به مرحله صحت‌سنجی نشان می‌دهند، که امری متداول در مدل‌های یادگیری ماشین است و به تفاوت بین داده‌های آموزشی و آزمایشی مربوط می‌شود. با این حال، کاهش اندک در معیارهای عملکرد (مانند NSE و r) در مرحله صحت‌سنجی برای مدل SVM نشان‌دهنده تعمیم‌پذیری بالای این مدل است. در مقابل، کاهش قابل‌توجه در r برای مدل RF ممکن است به حساسیت بیشتر این مدل به تغییرات در داده‌های آزمایشی اشاره داشته باشد. نتایج جدول ۳ نشان‌دهنده برتری نسبی مدل SVM با ورودی رطوبت نسبی در پیش‌بینی SPI در دشت مغان است، که می‌تواند به سیاست‌گذاران و مدیران منابع آب در این منطقه کشاورزی‌محور کمک کند تا استراتژی‌های مؤثری برای مدیریت خشکسالی توسعه دهند. انتخاب رطوبت نسبی به‌عنوان متغیر کلیدی در اکثر مدل‌ها، با توجه به میانگین ۷۱/۴۶ درصد و دامنه تغییرات ۵۱/۴۳ تا ۸۶/۴۲ درصد (جدول ۱)، نقش این متغیر در دینامیک‌های اقلیمی منطقه را تأیید می‌کند. این یافته‌ها می‌توانند در برنامه‌ریزی آبیاری، مدیریت منابع آب، و کاهش اثرات خشکسالی در دشت مغان مورد استفاده قرار گیرند.

جدول ۴- نتایج ارزیابی مدل‌های مورد استفاده در مطالعه حاضر در مراحل آموزش و صحت‌سنجی

Model	Best scenario	Stage	NSE	RMSE	r
RF	2	Train	0.83	0.314	0.94
		Test	0.784	0.464	0.85
ET	2	Train	0.82	0.31	0.94
		Test	0.791	0.397	0.87
SVM	2	Train	0.83	0.307	0.94
		Test	0.806	0.326	0.91
MLP	4	Train	0.69	0.491	0.85
		Test	0.67	0.461	0.82

شکل ۳ نمودار پراکنش متقابل مقادیر شاخص SPI مشاهداتی و پیش‌بینی‌شده توسط مدل‌های یادگیری ماشین را در مرحله آموزش نشان می‌دهد. این نمودارها امکان ارزیابی بصری و کمی تطابق بین مقادیر مشاهداتی و پیش‌بینی‌شده SPI را فراهم می‌کنند و مبنایی برای سنجش دقت مدل‌ها در پیش‌بینی خشکسالی هواشناسی در دشت مغان ارائه می‌دهند. مشاهده می‌شود که مدل‌های RF، ET، SVM و SVM در مرحله آموزش تطابق بسیار خوبی با مقادیر مشاهداتی SPI دارند. این تطابق خوب و همبستگی قوی، بیانگر توانایی این مدل‌ها در ثبت الگوهای خطی و غیرخطی موجود در داده‌های SPI است. در مقادیر حدی مانند $SPI = -1.85$ ، مدل‌های RF، ET، SVM پیش‌بینی‌های نزدیک به مقدار واقعی ارائه می‌دهند (مانند $-1/82$ ، $-1/75$ و $-1/53$ برای یکی از مشاهدات)، در حالی که MLP با مقادیری مانند $-0/91$ یا $-1/15$ انحراف بیشتری دارد. این نشان می‌دهد که مدل‌های مبتنی بر درخت و SVM ثبت شرایط شدید خشکسالی دقیق‌تر هستند. در مورد ترسالی‌های شدید ($SPI \geq 1.5$)، در مواردی مانند $SPI = 1.85$ ، تمامی مدل‌ها پیش‌بینی‌های نسبتاً دقیقی ارائه می‌دهند (مانند $1/66$ ، $1/47$ ، $1/47$ و $1/26$ برای RF، ET، SVM و MLP)، اما MLP در برخی موارد (مانند $0/69$ یا $-0/06$) انحراف بیشتری نشان می‌دهد. در محدوده مقادیر نزدیک به نرمال ($SPI \approx 0$) مدل‌ها عموماً عملکرد مشابهی دارند، اما MLP گاهی پیش‌بینی‌های نادرستی ارائه می‌دهد (مانند $-0/43$ برای $SPI = 0.12$).



شکل ۳- نمودار پراکنش متقابل مقادیر شاخص SPI مشاهداتی و شبیه‌سازی شده با استفاده از مدل‌های هوشمند در مرحله آموزش.

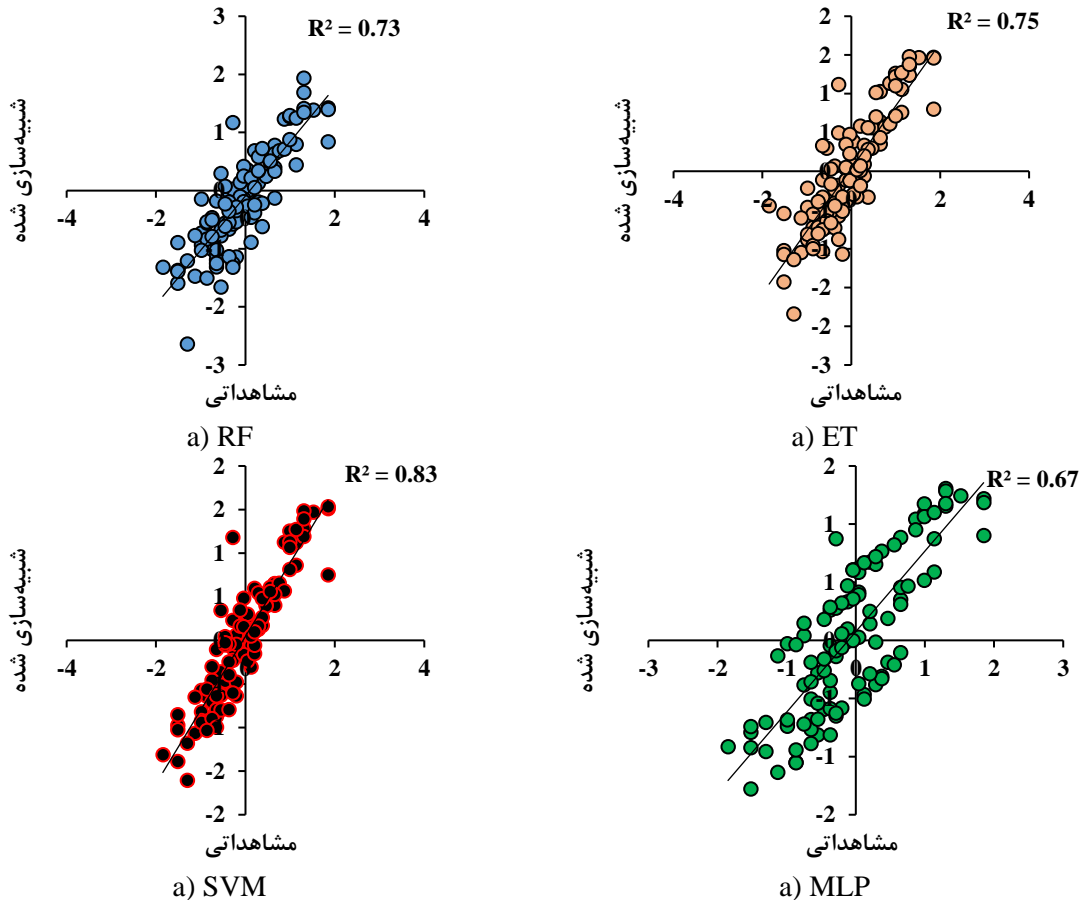
شکل ۴ نمودارهای پراکنش مقادیر شاخص SPI مشاهداتی و پیش‌بینی شده توسط مدل‌های یادگیری ماشین را در مرحله صحت‌سنجی برای دوره زمانی ژانویه ۲۰۱۴ تا دسامبر ۲۰۲۲ (۱۰۸ ماه) در دشت مغان نشان می‌دهد که نقاط نزدیک به خط $(y=x)$ بیانگر پیش‌بینی‌های دقیق هستند. مشاهدات نشان‌دهنده تنوع قابل توجه در SPI است، با مقادیر شدید خشکسالی (مانند $-1/85$ در سپتامبر ۲۰۲۲) و ترسالی (مانند $1/85$ در فوریه ۲۰۱۵ و ژانویه ۲۰۱۶). مدل‌های RF، ET و SVM به‌طور کلی روندهای زمانی را به‌خوبی دنبال می‌کنند، به‌ویژه در دوره‌های ترسالی (مانند اکتبر ۲۰۱۵ با $SPI = 1.30$ و پیش‌بینی‌های $1/93$ ، $1/47$ و $1/48$ برای RF، ET و SVM). با این حال، MLP در برخی موارد، مانند ژوئیه ۲۰۱۵ ($SPI = 0.65$ ، پیش‌بینی $MLP = -0.11$) و ژوئیه ۲۰۱۶ ($SPI = 0.29$ ، پیش‌بینی $MLP = -0.38$)، انحرافات قابل توجهی نشان می‌دهد. این انحرافات ممکن است به حساسیت MLP به تنظیم‌های اولیه یا ناکافی بودن داده‌های آموزشی برای ثبت روابط غیرخطی پیچیده مرتبط باشد. مدل SVM با $R^2 = 0.83$ بهترین تطابق را نشان می‌دهد، به‌ویژه در مقادیر نزدیک به نرمال ($SPI \approx 0$) و دوره‌های ترسالی شدید (مانند $SPI = 1.85$ با پیش‌بینی $1/51$ در فوریه ۲۰۱۵). مدل‌های RF و ET نیز با $R^2 = 0.73$ و 0.75 عملکرد قابل قبولی دارند، اما در برخی موارد، مانند آوریل ۲۰۲۲ ($SPI = -1.30$ ، پیش‌بینی $RF = -2.64$)، انحرافات بیشتری نشان می‌دهند. با $R^2 = 0.67$ کمترین تطابق را دارد، به‌ویژه در مقادیر مثبت SPI (مانند ژوئیه ۲۰۱۵ با پیش‌بینی $-0/11$ برای $SPI = 0.65$).

در پیش‌بینی خشکسالی‌های شدید ($SPI \leq -1.5$)، در مواردی مانند سپتامبر ۲۰۲۲ ($SPI = -1.85$)، مدل‌های SVM، ET و MLP پیش‌بینی‌های متفاوتی ارائه می‌دهند، که SVM نزدیک‌ترین پیش‌بینی به مقدار واقعی را دارد. در این مورد ($-1/32$) نیز عملکرد خوبی دارد، اما در مارس ۲۰۱۷ ($SPI = -1.52$)، پیش‌بینی ($RF = -1.38$) دقت بیشتری نشان می‌دهد. در پیش‌بینی رویدادهای ترسالی شدید ($SPI \geq 1.5$) مثل فوریه ۲۰۱۵ و ژانویه ۲۰۱۶ ($SPI = 1.85$)، تمامی مدل‌ها پیش‌بینی‌های نزدیک به واقعی ارائه می‌دهند، اما MLP در برخی موارد (مانند نوامبر ۲۰۱۴ با $SPI = 1.52$) انحراف کمتری دارد. در محدوده مقادیر نزدیک به نرمال ($SPI \approx 0$)، مدل‌های SVM و ET پیش‌بینی‌های دقیق‌تری ارائه می‌دهند (مانند اکتبر ۲۰۱۴ با $SPI = 0.04$)، در حالی

انحرافات بیشتری

که MLP گاهی

نشان می‌دهد (مانند پیش‌بینی ۰/۶۱ برای $SPI = -0.04$ در مارس ۲۰۱۸). این نتایج توانایی مدل‌های یادگیری ماشین، به‌ویژه SVM، را در ثبت الگوهای خشکسالی و ترسالی در دشت مغان تأیید می‌کنند و می‌توانند به بهبود پیش‌بینی‌های هیدرولوژیکی و مدیریت پایدار منابع آب کمک کنند.



شکل ۵- نمودار پراکنش متقابل مقادیر شاخص SPI مشاهداتی و شبیه‌سازی شده با استفاده از مدل‌های هوشمند در مرحله آزمون.

به‌طور کلی، اگرچه SVM بهترین عملکرد را در بین مدل‌های مورد بررسی نشان می‌دهد، اما انحرافات در برخی موارد (مانند سپتامبر ۲۰۲۲) نشان‌دهنده محدودیت‌های مدل‌ها در ثبت شرایط شدید است. MLP به دلیل انحرافات بیشتر، نیاز به تنظیم دقیق‌تر معماری شبکه یا داده‌های آموزشی بیشتری دارد. پیشنهاد می‌شود در مطالعات آینده، متغیرهای ورودی اضافی (مانند فشار جو یا تابش خورشیدی) بررسی شوند. همچنین، استفاده از روش‌های تحلیل حساسیت می‌تواند نقش متغیرهای ورودی (مانند رطوبت نسبی، که در جدول ۳ به‌عنوان بهترین ورودی برای RF، ET، SVM و شناسایی شد) را روشن‌تر کند.

بحث و نتیجه‌گیری

این پژوهش، با تمرکز بر پیش‌بینی شاخص بارش استاندارد شده (SPI) در دشت مغان، نشان‌دهنده توانایی بالای مدل‌های یادگیری ماشین، به‌ویژه ماشین بردار پشتیبان (SVM)، جنگل تصادفی (RF)، و درختان اضافی (ET)، در ثبت الگوهای خشکسالی و ترسالی هواشناسی است. این مدل‌ها با استفاده از ترکیب‌های مختلف متغیرهای اقلیمی (دما، رطوبت نسبی، و سرعت باد)، به‌ویژه در شرایط اقلیمی متنوع منطقه، توانستند تطابق قابل‌قبولی بین مقادیر مشاهداتی و پیش‌بینی‌شده ایجاد کنند. نقش کلیدی رطوبت نسبی به‌عنوان یک متغیر ورودی برجسته، بر اهمیت تحلیل دقیق متغیرهای اقلیمی در پیش‌بینی خشکسالی تأکید دارد. این یافته‌ها نه تنها درک عمیق‌تری از دینامیک‌های اقلیمی دشت مغان ارائه می‌دهند، بلکه مبنایی برای توسعه استراتژی‌های مدیریت منابع آب و کشاورزی در مناطق نیمه‌خشک فراهم می‌کنند.



پنجمین کنفرانس ملی هیدرولوژی ایران
The 5th National Conference on Hydrology of Iran
ارومیه ۱۳-۱۲ آذر ۱۴۰۴



اهمیت استفاده از

نتایج این پژوهش بر

رویکردهای چندمدلی و تحلیل حساسیت متغیرهای ورودی برای تقویت پیش‌بینی‌های هیدرولوژیکی و حمایت از تصمیم‌گیری‌های پایدار در برابر تغییرات اقلیمی تأکید دارد. با وجود عملکرد قوی مدل‌های یادگیری ماشین، برخی محدودیت‌ها، مانند انحرافات در پیش‌بینی مقادیر شدید SPI توسط مدل پرسپترون چندلایه (MLP) و وابستگی به داده‌های ورودی خاص، نشان‌دهنده نیاز به تحقیقات بیشتر است. پیشنهاد می‌شود در مطالعات آینده، متغیرهای اقلیمی اضافی، مانند فشار جو یا تابش خورشیدی، و تکنیک‌های پیشرفته‌تر سازمان‌بندی مدل‌ها بررسی شوند تا دقت پیش‌بینی‌ها بهبود یابد.

منابع

- ۱- خلیلی، س.، مهتابی، ق.، نیکبخت، ج.، و محمدی، م. (۱۴۰۲). مطالعه ارتباط زمانی خشکسالی هیدرولوژیکی و هواشناسی با استفاده از شاخص‌های RDI، SPI و SDI با تأکید بر مقایسه دو روش مختلف محاسبه ET_0 (مطالعه موردی: حوضه آبریز صوفی چای مراغه). مهندسی آبیاری و آب ایران، ۱۴(۲)، ۲۸۰-۲۵۷.
- ۲- دین پژوه، ی.، و الله‌ویردی پور، پ. (۱۴۰۴). پایش و پیش‌آگاهی از تغییرات تبخیر-تعرق مرجع دشت مغان مطابق گزارش ششم IPCC. محیط زیست و مهندسی آب، ۱۱(۱)، ۴۷-۵۶.
- ۳- راهواره، م.، معتمدوزیری، ب.، مقدم نیا، ع.، و مریدی، ع. (۱۴۰۱). بررسی خشکسالی هواشناسی و هیدرولوژیکی حوضه زرنه‌رود با استفاده از شاخص‌های SPI و SRI تحت سناریوهای تغییر اقلیم. مهندسی اکوسیستم بیابان، ۱۱(۳۷)، ۲۶-۱۵.
- ۴- رنجبر، آ.، ابراهیمی پاک، ن.، و تافته، آ. (۱۴۰۳). بررسی تغییرات فراوانی بارش‌های با شدت کم و رابطه آن با روند بارش و شاخص بارش استاندارد (SPI) در استان کردستان. هواشناسی کشاورزی، ۱۲(۱)، ۸۶-۷۵.

- 1- Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5–32.
- 2- Dikici, M., & Aksel, M. (2021). Comparison of drought indices in the case of the Ceyhan Basin. *Environment and Geoinformatics*, 8(2), 113-125.
- 3- Dinpashoh, Y., & Allahverdi-pour, P. (2025). Monitoring and Predicting Changes in Reference Evapotranspiration in the Moghan Plain According to CMIP6 of IPCC. *Environment and Water Engineering*, 11(1), 47-56. doi: 10.22034/ewe.2024.466037.1947
- 4- Hao, Z., Singh, V. P., & Xia, Y. (2018). Seasonal drought prediction: Advances, challenges, and future prospects. *Reviews of Geophysics*, 56(1), 108-141.
- 5- Hayes, M., Svoboda, M., Wall, N., & Widhalm, M. (2011). The Lincoln declaration on drought indices: universal meteorological drought index recommended. *Bulletin of the American Meteorological Society*, 92(4), 485-488.
- 6- Haykin, S. (1994). *Neural networks: a comprehensive foundation*. Prentice hall PTR.
- 7- IPCC. (2021). *Climate Change 2021: The Physical Science Basis*. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change.
- 8- McKee, T. B., Doesken, N. J., & Kleist, J. (1993). The relationship of drought frequency and duration to time scales. *Proceedings of the 8th Conference on Applied Climatology*, 179–184.
- 9- Moriasi, D. N., Arnold, J. G., Van Liew, M. W., Bingner, R. L., Harmel, R. D., & Veith, T. L. (2007). Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Transactions of the ASABE*, 50(3), 885-900.
- 10- Shi, X., Ding, H., Wu, M., Shi, M., Chen, F., Li, Y., Yang, Y. (2022). A comprehensive drought monitoring method integrating multi-source data. *PeerJ*, 10, e13560.
- 11- Spinoni, J., Naumann, G., Vogt, J. V., & Barbosa, P. (2016). *Meteorological droughts in Europe. Events and impacts, past trends and future projections*. Publications Office of the European Union, Luxembourg, EUR 27748 EN.
- 12- Vapnik, V. (2013). *The nature of statistical learning theory*. Springer science & business media.
- 13- Vicente-Serrano, S. M., Beguería, S., & López-Moreno, J. I. (2010). A multiscalar drought index sensitive to global warming: the standardized precipitation evapotranspiration index. *Journal of climate*, 23(7), 1696-1718.
- 14- Vogt, J. V., Naumann, G., Masante, D., Spinoni, J., Cammalleri, C., Erian, W., & Barbosa, P. (2018). *Drought risk assessment and management: A conceptual framework*. Publications Office of the European Union.



Application of Machine Learning Algorithms in Simulating Drought Based on the SPI Index in the Moghan Plain

Pouya Allahverdipour^{*1}, Yagob Dinpashoh²

- 1- Ph.D. Candidate, Department of Water Engineering, Faculty of Agriculture, Tabriz University, Tabriz, Iran.
- 2- Professor, Department of Water Engineering, Faculty of Agriculture, Tabriz University, Tabriz, Iran

* Corresponding author: a.pouya1996@gmail.com

ABSTRACT

Drought, a complex natural phenomenon, profoundly impacts water resources and ecosystem sustainability worldwide. In agriculturally dependent regions like the Moghan Plain, reliance on water resources exacerbates drought effects, posing significant challenges to sustainable resource management. This study evaluates the performance of machine learning models in predicting the Standardized Precipitation Index (SPI) for monitoring meteorological drought in the Moghan Plain over a 30-year period (1993–2022). Monthly precipitation, temperature, relative humidity, and wind speed data from the Parsabad Moghan station were used. Random Forest (RF), Extra Trees (ET), Support Vector Machine (SVM), and Multi-Layer Perceptron (MLP) models were assessed using correlation coefficient (r), Root Mean Square Error (RMSE), and Nash-Sutcliffe Efficiency (NSE) metrics. Results indicate that in the training phase, SVM exhibited superior performance with $r = 0.95$, $RMSE = 0.30$, and $NSE = 0.83$. In the validation phase, SVM maintained its lead with $r = 0.88$, $RMSE = 0.39$, and $NSE = 0.81$. Relative humidity was identified as the optimal input for RF, ET, and SVM, while the combination of temperature and relative humidity was best for MLP. These findings underscore the efficacy of machine learning models, in drought prediction and support sustainable water resource management.

Keywords: Support Vector, Extra Trees, Neural Network, Water Resources, Artificial Intelligence, Machine Learning.