

ارزیابی مدل‌های مختلف آماری در تهیه نقشه سیل‌گیری استان گیلان

- ❖ عیسی غلامی؛ دانشجوی کارشناسی ارشد آبخیزداری، دانشکده منابع طبیعی، دانشگاه تربیت مدرس، نور، ایران.
- ❖ مهدی وفاخواه*؛ استاد گروه آبخیزداری، دانشکده منابع طبیعی، دانشگاه تربیت مدرس، نور، ایران.
- ❖ سیدجلیل علوی؛ استادیار گروه جنگلداری، دانشکده منابع طبیعی، دانشگاه تربیت مدرس، نور، ایران.

چکیده

به دلیل کمبود اطلاعات در اکثر حوزه‌های آبخیز، بسیاری از محققین برای مطالعه‌های هیدرولوژیکی و سیل‌گیری به استفاده از تجزیه و تحلیل‌های مکانی در سیستم اطلاعات جغرافیایی روی آوردند. پژوهش حاضر به منظور مقایسه کارایی سه مدل ماشین بردار پشتیبان (SVM)، خطی تعمیم یافته (GLM) و جمعی تعمیم یافته (GAM) در تهیه نقشه سیل‌گیری استان گیلان برنامه‌ریزی شده است. بدین منظور لایه‌های اطلاعاتی درجه شیب، جهت شیب، شکل شیب، ارتفاع از سطح دریا، فاصله از رودخانه، تراکم زهکشی، زمین شناسی، کاربری اراضی، شاخص رطوبت توپوگرافی و شاخص توان آبراهه در محیط سامانه اطلاعات جغرافیایی (نرم‌افزارهای ArcGIS و SAGA-GIS) تهیه شدند. سپس بر اساس اطلاعات ۲۲۰ نقطه سیل‌گیر، از ۷۰ درصد تعداد کل نقاط به منظور واسنجی و ۳۰ درصد باقیمانده برای اعتبارسنجی و ارزیابی کارایی مدل‌ها مورد استفاده قرار گرفت. نتایج ارزیابی دقت مدل‌ها به ترتیب با استفاده از شاخص‌های سطح زیر منحنی (AUC) و کاپا (Kappa) نشان داد که از نظر شاخص سطح زیر منحنی (AUC)، مدل ماشین بردار پشتیبان (SVM) با ۰/۸۳۵ و مدل جمعی تعمیم یافته (GAM) با ۰/۸۲۷ دارای دقت خیلی خوب و مدل خطی تعمیم یافته (GLM) با ۰/۷۹۹ دارای دقت خوب می‌باشد. از نظر شاخص کاپا (Kappa) مدل ماشین بردار پشتیبان (SVM) با ۰/۵۸ داری دقت خوب، مدل جمعی تعمیم یافته (GAM) با ۰/۵۳ و مدل خطی تعمیم یافته (GLM) با ۰/۴۸ دارای دقت قابل قبول می‌باشند. بنابراین بر اساس شاخص‌های مذکور مدل ماشین بردار پشتیبان (SVM) نسبت به دو مدل دیگر در شناسایی مناطق سیل‌گیر کارایی بالاتری دارد. همچنین عوامل فاصله از رودخانه، ارتفاع از سطح دریا و شیب بیشترین تأثیر را بر سیل‌گیری منطقه مورد مطالعه دارند.

کلید واژگان: سیل‌گیری، داده‌کاوی، مدل‌های داده محور، مدل‌سازی، منحنی تشخیص عملکرد، استان گیلان.

۱. مقدمه

سیلاب به جریان شدید توأم با بالا آمدن نسبتاً زیاد آب در یک رودخانه اطلاق شده که معمولاً اراضی پست را فرا می‌گیرد [۹]. بین سال‌های ۲۰۰۰ تا ۲۰۰۸ تقریباً ۹۹ میلیون نفر از مردم در سال توسط سیلاب در جهان تحت تأثیر قرار گرفته‌اند. هر ساله در شهرهای شمالی ایران از قبیل استان‌های مازندران، گیلان و گلستان سیلاب‌های مهیبی اتفاق می‌افتند [۲۳]. انتظار می‌رود در شرایط فعلی، با توجه به توسعه شهرنشینی، افزایش جنگل‌زدایی و تداوم بارش‌های شدید به‌عنوان یک اثر ناشی از تغییرات آب و هوایی در مناطق مختلف و حساس، بیش از گذشته احساس شود. با در نظر گرفتن تأثیر منفی سیل، شناسایی مناطق سیل‌گیر به‌منظور کنترل و پیش‌گیری از آن ضروری می‌باشد. از طرفی با توجه به وقوع این سیل‌ها مکان‌های با بیشترین پتانسیل سیل‌گیری باید قبل از برنامه‌ریزی به‌وسیله تهیه نقشه پهنه‌بندی سیل مشخص شوند [۳]. در سال‌های اخیر مدل‌های آماری و برحسب احتمال زیادی برای تهیه نقشه سیل‌گیری به‌کار برده شده است [۱۷، ۱۸]. همچنین سیستم اطلاعات جغرافیایی (GIS) به‌عنوان ابزارهای آنالیز پایه به‌دلیل توانایی در تجزیه و تحلیل و مدیریت مقادیر زیادی از داده‌های مکانی به‌کار رفته است [۲۴]. از جمله این روش‌ها می‌توان به مدل خطی تعمیم یافته (GLM)، مدل جمعی تعمیم یافته (GAM) و مدل ماشین‌بردار پشتیبان (SVM) اشاره کرد. مدل‌های خطی تعمیم یافته و ماشین‌بردار پشتیبان نتایج خوبی را در مدل‌سازی فرآیندهای خطی و غیر خطی نشان داده‌اند [۳۱]. تهیه نقشه سیل‌گیری با استفاده از روش‌های تصمیم‌گیری چندمعیاره [۱۶]، رگرسیون لجستیک (LR) [۱۹، ۳۰]، شبکه عصبی مصنوعی (ANN) [۶؛ ۱۱؛ ۳۶]، نسبت فراوانی (FR) [۴]، شبکه عصبی-فازی تطبیقی [۳۴]، درخت رگرسیون تقویت شده [۲۶]، مدل خطی

تعمیم یافته (GLM) [۵]، ماشین‌بردار پشتیبان (SVM) [۳۱] و جنگل تصادفی (RF) [۵، ۲۶] انجام شده است. در پژوهشی نقشه سیل‌گیری شهر جهرم واقع در استان فارس را با بهینه‌سازی پارامترهای مدل سیستم استنتاج عصبی-فازی تطبیقی (ANFIS) توسط الگوریتم‌های مختلف فراشناختی کلونی مورچگان (ACO)، الگوریتم ژنتیک (GA) و بهینه‌سازی ذرات انبوه (PSO) تهیه کردند. نتایج این تحقیق بر اساس سطح زیر منحنی تشخیص عملکرد سیستم (ROC) برای مدل‌های ترکیبی ANFIS-ACO، ANFIS-GA و ANFIS-PSO به‌ترتیب برابر با ۰/۹۱۴، ۰/۹۱۸ و ۰/۹۴۵ گزارش شد [۳۲]. در مطالعه‌ای به مقایسه عملکرد سه مدل جنگل تصادفی (RF)، شبکه عصبی مصنوعی (ANN) و مدل ماشین‌بردار پشتیبان (SVM) در مناطق کوهستانی چین بر اساس تاریخچه سیلاب‌های گذشته از ۱۹۴۹ تا ۲۰۰۰ پرداختند. نتایج نشان داد که مدل جنگل تصادفی دارای دقت بالایی نسبت به مدل‌های شبکه عصبی مصنوعی و ماشین‌بردار پشتیبان می‌باشد [۳۶]. بنابراین بر اساس مطالب فوق هدف اصلی این تحقیق بررسی رفتار عوامل مختلف مؤثر بر سیل‌گیری و عملکرد مدل‌های خطی تعمیم یافته (GLM)، جمعی تعمیم یافته (GAM) و ماشین‌بردار پشتیبان (SVM) در شناسایی مناطق مستعد سیل در استان گیلان می‌باشد.

۲. روش‌شناسی

۲.۱. معرفی منطقه مورد مطالعه

استان گیلان با وسعت ۱۴۱۰۰ کیلومتر مربع در محدوده جغرافیایی " ۳۶° ۳۴' ۰۰" تا " ۳۸° ۲۷' ۰۰" عرض شمالی و " ۴۸° ۳۴' ۰۰" تا " ۴۸° ۳۶' ۰۰" طول شرقی در حاشیه جنوبی دریای خزر قرار دارد. ارتفاع از سطح آب‌های آزاد از ۲۸- متر در سواحل تا ۲۷۰۰ متر در

^۱Generalized Additive Model

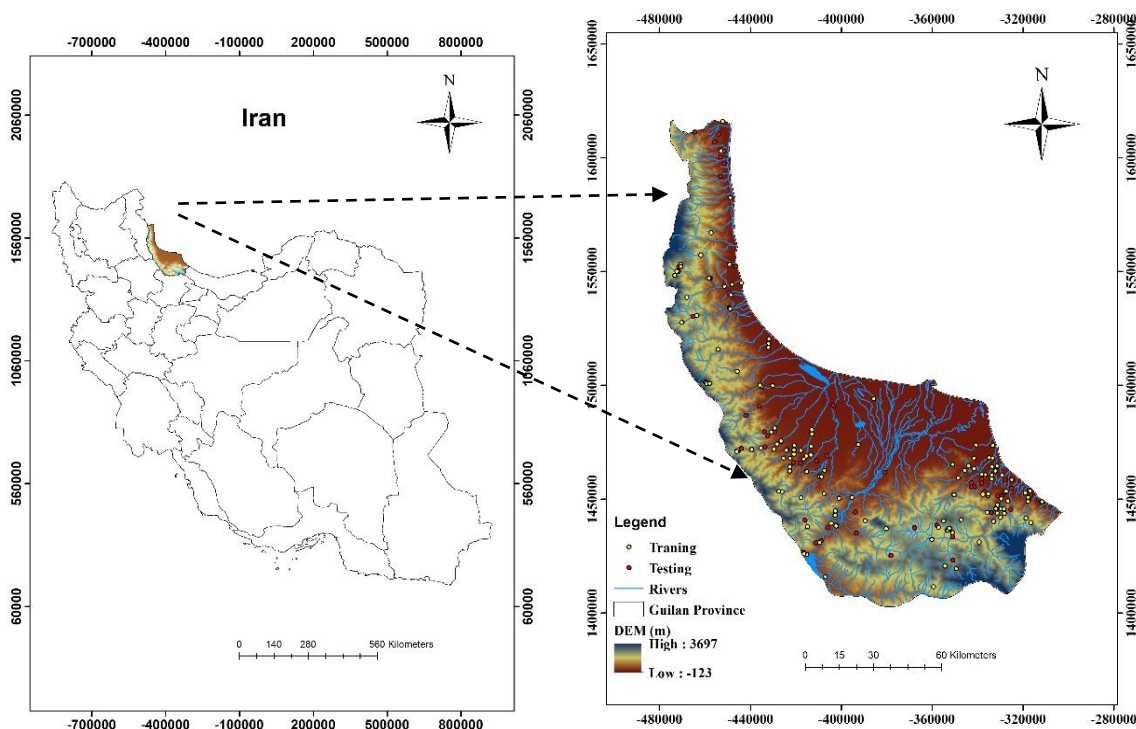
^۲Support Vector Machine

^۱Geographic Information System

^۲Generalized Linear Model

بارندگی در بین استان‌های شمالی کشور می‌باشد.

مناطق کوهستانی متغیر می‌باشد [۱]. میانگین بارندگی سالانه برابر با ۱۱۰۰ میلیمتر و دارای بیشترین میزان



شکل ۱. موقعیت کلی استان گیلان در ایران

کاربری اراضی، شاخص رطوبت توپوگرافی و شاخص توان آبراهه، در محیط سامانه اطلاعات جغرافیایی (نرم‌افزارهای ArcGIS و SAGA-GIS) تهیه شدند.

۳،۲. مدل‌های آماری

۱،۳،۲. مدل ماشین بردار پشتیبان (SVM)

یکی از روش‌های یادگیری با نظارت SVM بوده که از آن برای گروه‌بندی و برآورد تابع برازش داده‌ها استفاده می‌شود. به طوری که کمترین خطا در گروه‌بندی داده‌ها یا تابع برازش رخ دهد. این روش براساس تئوری یادگیری

۲،۲. ایجاد پایگاه داده

بر اساس پژوهش‌های [۵، ۲۶، ۳۶] و با توجه به ویژگی‌های محلی ۱۱ عاملی که در رخداد سیل تأثیرگذار بوده، شناسایی شدند. سپس نقشه پراکنش ۲۲۰ مکان سیل گیر شکل (۱)، از اداره کل منابع طبیعی استان گیلان تهیه شد. همچنین نقشه زمین‌شناسی منطقه در مقیاس ۱:۱۰۰۰۰۰ از سازمان زمین‌شناسی کشور و نقشه توپوگرافی در مقیاس ۱:۵۰۰۰۰ از سازمان نقشه برداری تهیه شدند. نقشه کاربری اراضی با استفاده از تصویر Landsat به دست آمد. سپس نقشه هر یک از عوامل مؤثر بر سیل‌گیری از قبیل درجه شیب، جهت شیب، شکل شیب (انحنای سطح و انحنای نیمرخ)، ارتفاع از سطح دریا، فاصله از رودخانه، تراکم زهکشی، زمین‌شناسی،

به وسیله فرمول ارائه می‌شود، می‌توان از طریق داده‌ها، منحنی تشخیص عملکرد سیستم (ROC) را تعیین کرد. تفاوت اساسی مدل‌های جمعی تعمیم یافته با مدل‌های پارامتری در این است که توابع خطی به وسیله توابع هموار نامعلوم جانشین می‌شوند که دارا بودن هموارسازها یکی از مزایای مهم مدل جمعی تعمیم یافته است که آن را از سایر مدل‌ها متمایز می‌کند [۱۴]. نکته مهم در مورد این مدل‌ها این است که به جای مدل محوری، داده محورند، یعنی نتایج پارامترهای به دست آمده از مدل پیشین استنباط نمی‌شود، بلکه ساختار داده‌ها آزمون می‌شود [۲۸]. لذا تمامی مدل‌ها با استفاده از بسته (sdm)^۴ نسخه ۳.۴.۱ در نرم افزار آماری R نسخه ۳.۵.۰ اجرا شدند.

۴.۲. اعتبارسنجی مدل‌ها

در این پژوهش ۷۰ درصد کل نقاط سیل به منظور واسنجی و ۳۰ درصد باقیمانده برای اعتبارسنجی و ارزیابی کارایی مدل‌ها مورد استفاده قرار گرفت [۳۱]. هر کدام از مدل‌ها ۱۰ بار اجراء و اعتبارسنجی شده و در نهایت میانگین اعتبارسنجی گزارش شد. سپس با استفاده از سطح زیر منحنی^۵ (AUC) تشخیص عملکرد سیستم (ROC) و شاخص کاپا^۶ صحت نقشه‌های تهیه شده، مشخص شد [۱۰]. با قرار دادن حساسیت مدل در محور xها و ویژگی در محور yها منحنی ROC ترسیم می‌شود. مقادیر سطح زیر منحنی ROC به پنج دسته عالی (۰/۹-۱)، خیلی خوب (۰/۸-۰/۹)، خوب (۰/۷-۰/۸)، متوسط (۰/۶-۰/۷) و ضعیف (۰/۵-۰/۶) طبقه‌بندی شده است [۲۵]. ضریب کاپا از توافق بین مقادیر پیش‌بینی شده و مشاهده شده به دست می‌آید. مقادیر آماره کاپا در محدوده ۱- به ۱+ قرار دارد که مقدار ۱+ نشان‌دهنده موافقت کامل و ارزش صفر یا کمتر از ۱ نشان‌دهنده عملکرد بهتر از تصادفی است [۲۲]. مقادیر کاپا پایین‌تر از

آماره^۱ است و از اصل کمینه سازی خطای ساختاری (SRM) بهره می‌گیرد تا یک جواب بهینه کلی حاصل شود [۳۵]. این مدل می‌تواند به صورت خطی به داده‌ها برازش شود و اگر چیدمان داده‌ها در فضای اصلی غیرخطی باشد آن‌ها را به فضای بالاتری می‌برد تا نحوه قرارگیری آن‌ها به صورت خطی شود. از طرفی به دلیل این که یافتن تابع ویژگی برای مسائل مختلف کار بسیار دشواری است بنابراین می‌توان از توابع کرنل^۱ استفاده نمود [۱۵].

۲.۳.۲. مدل خطی تعمیم یافته (GLM)

مدل GLM برای تجزیه و تحلیل داده‌های دارای توزیع غیر نرمال در چهارچوب مدل رگرسیون توسعه یافته است. هم‌چنین GLM طیف وسیعی از داده‌ها با توزیع نمایی را پوشش می‌دهد. در این مدل با استفاده از تابع پیوند (g)، متغیر پیشگو و متغیر اندازه‌گیری شده با هم ترکیب می‌شوند. این مدل ارتقاء یافته مدل خطی است که از حالت طبیعی خود داده‌ها استفاده می‌نماید و در نتیجه داده‌های استفاده شده می‌تواند غیرخطی بوده و یا واریانس آن‌ها ثابت نباشد. در این مدل امکان دارد که داده‌ها از توابع توزیع مختلفی از جمله توزیع نرمال، دو جمله‌ای، پواسون، دو جمله‌ای منفی و یا توزیع گاما پیروی کنند که بعضی از آن‌ها مناسب برای داده‌های هیدرولوژیکی می‌باشند. بنابراین مدل GLM انعطاف پذیرتر و برای تحلیل روابط هیدرولوژیکی که از توزیع نرمال پیروی نمی‌کنند مناسب است [۱۴].

۳.۳.۲. مدل جمعی تعمیم یافته (GAM)

مدل جمعی تعمیم یافته یک مدل پارامتری و از طرفی بسط یافته مدل‌های خطی تعمیم یافته می‌باشد. در مدل‌های جمعی تعمیم یافته برخلاف مدل رگرسیون خطی که در آن رابطه بین متغیرهای تبیینی و پاسخ

^۴Species Distribution Modelling

^۵Area Under Curve (AUC)

^۶Kappa Index

^۱Statistical Learning Theory

^۲Structural Risk Minimization

^۳Kernel Functions

مشاهده می‌شود.

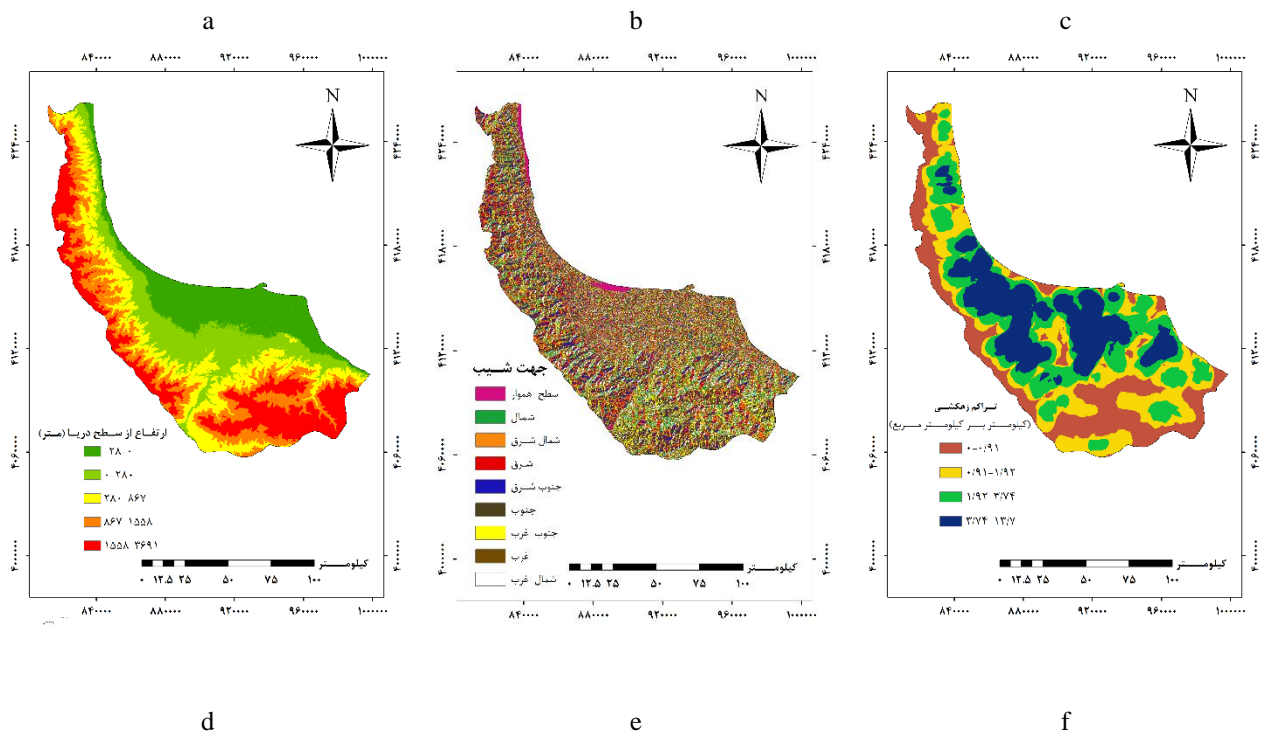
۰/۴ ضعیف، ۰/۴ تا ۰/۵۵ قابل قبول، ۰/۵۵ تا ۰/۷۰ تا ۰/۷۰ خوب، ۰/۷۰ تا ۰/۸۵ خیلی خوب و بالاتر از ۰/۸۵ عالی می‌باشد [۲۱].

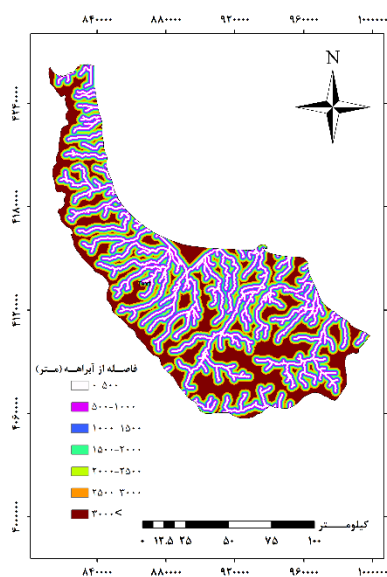
۱,۳. نتایج حاصل از اجرای مدل‌ها

نتایج ارزیابی دقت مدل‌ها به ترتیب با استفاده از شاخص‌های سطح زیر منحنی (AUC) و شاخص کاپا (Kappa) بر اساس جدول (۱) و شکل (۳) نشان داد که از نظر شاخص سطح زیر منحنی (AUC)، مدل ماشین بردار پشتیبان (SVM) با ۰/۸۳۳ و مدل جمعی تعمیم یافته (GAM) با ۰/۸۱۷ دارای دقت خیلی خوب می‌باشند.

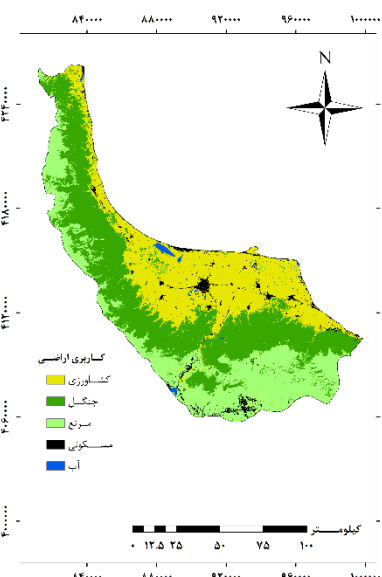
۳. نتایج

نقشه هر یک از عوامل مؤثر بر سیل‌گیری از قبیل درجه شیب، جهت شیب، شکل شیب (انحنای سطح و انحنای نیمرخ)، ارتفاع از سطح دریا، فاصله از رودخانه، تراکم زهکشی، زمین‌شناسی، کاربری اراضی، شاخص رطوبت توپوگرافی و شاخص توان آبراهه در شکل (۲)

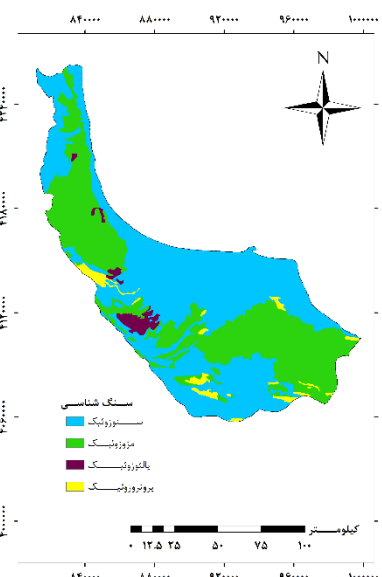




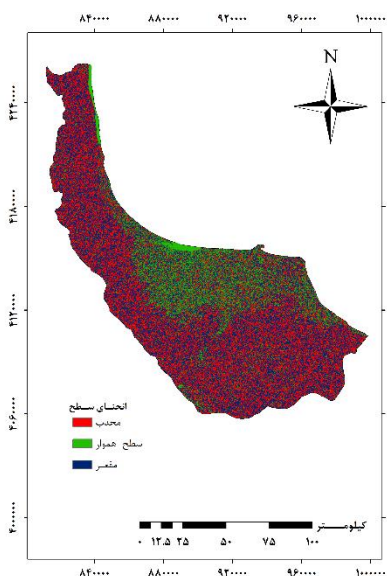
g



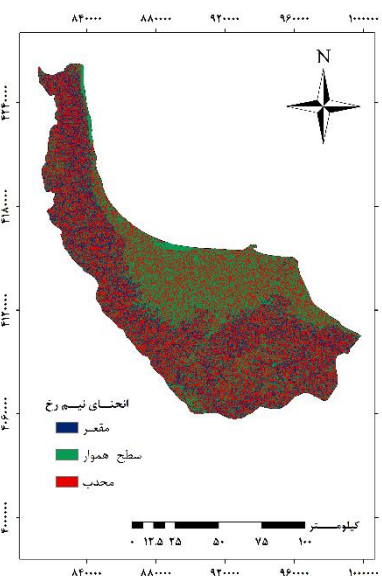
h



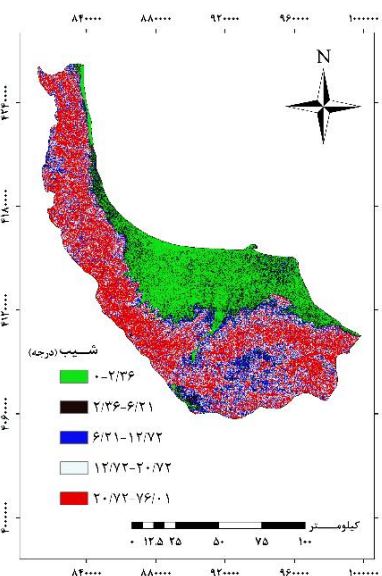
i

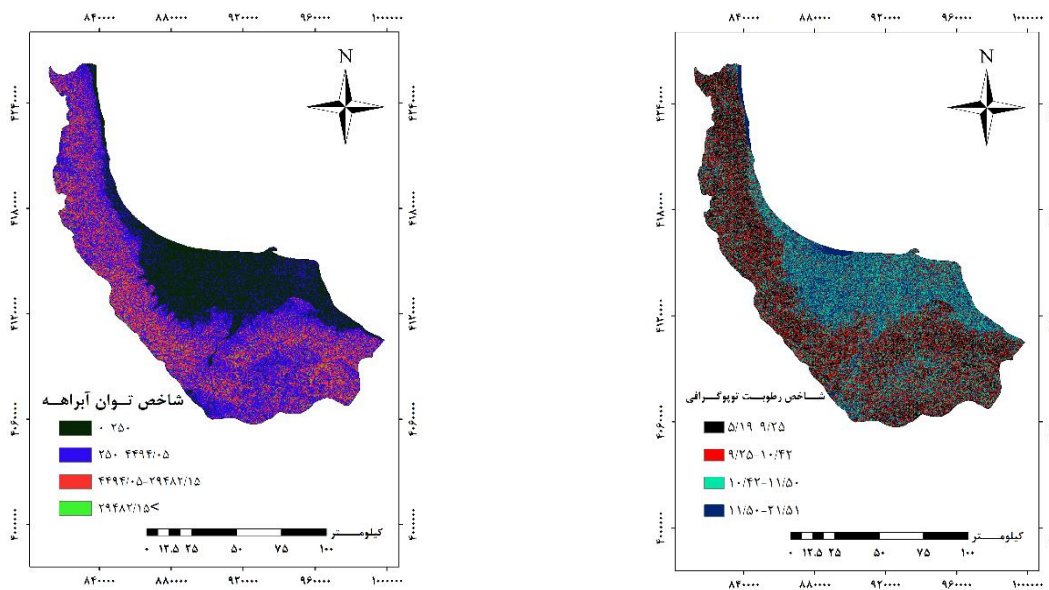


j



k





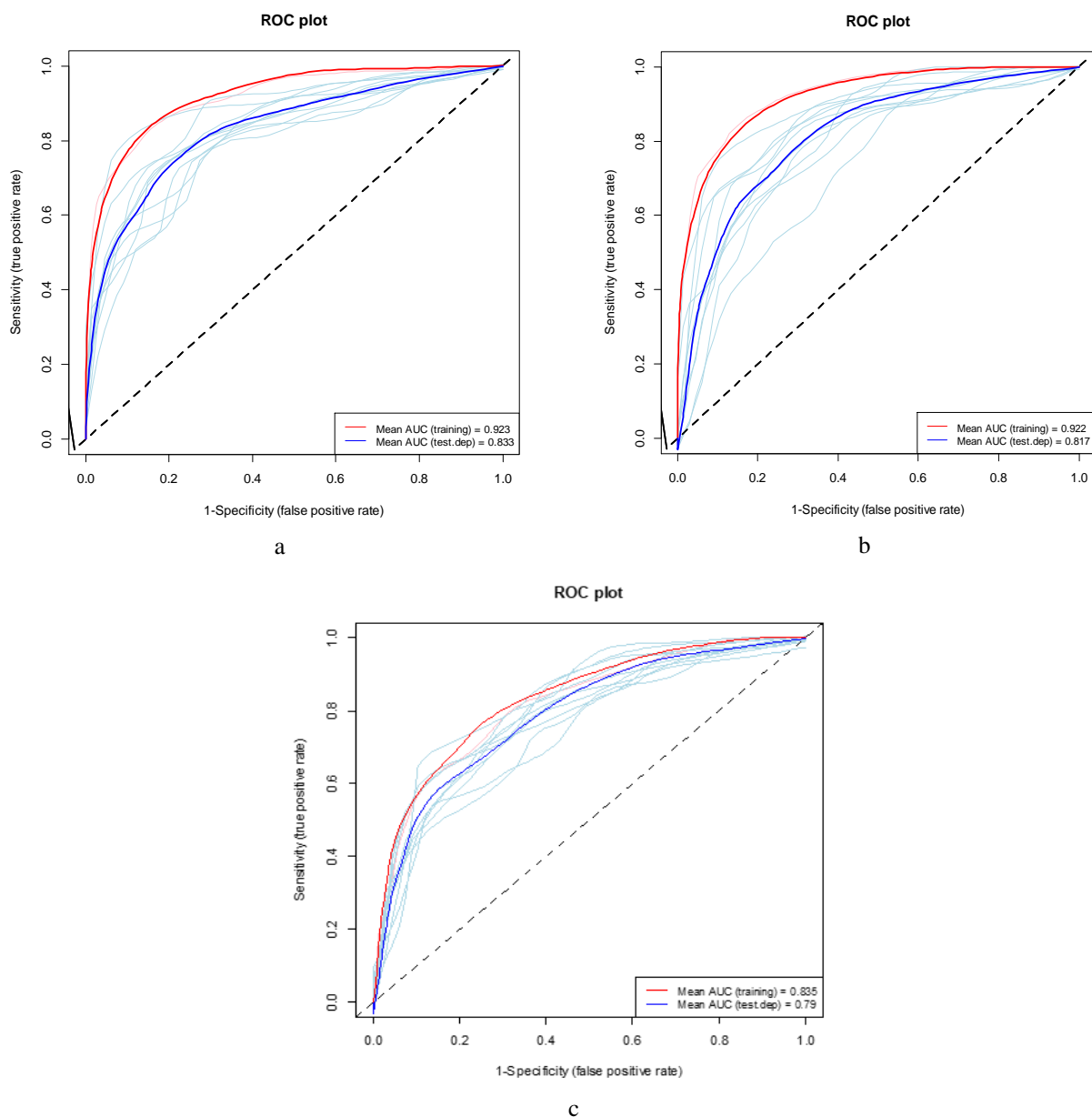
شکل ۲. نقشه‌های ورودی: (a) ارتفاع از سطح دریا، (b) جهت شیب، (c) تراکم زهکشی، (d) فاصله از آبراهه، (e) کاربری اراضی، (f) سنگ‌شناسی، (g) انحنا سطح، (h) انحنا نیم‌رخ، (i) شیب، (j) شاخص توان آبراهه، (k) شاخص رطوبت توپوگرافی

خوب، مدل جمعی تعمیم یافته (GAM) با 0.53 و مدل خطی تعمیم یافته (GLM) با 0.48 دارای دقت قابل قبول می‌باشند.

در حالی که مدل خطی تعمیم یافته (GLM) با 0.79 دارای دقت خوبی بوده و از نظر شاخص کاپا (Kappa) مدل ماشین بردار پشتیبان (SVM) با 0.58 دارای دقت

جدول ۱. نتایج حاصل از ارزیابی مدل‌ها بر اساس شاخص‌های سطح زیر منحنی (AUC) و شاخص کاپا (Kappa)

مدل	خطی تعمیم یافته (GLM)		جمعی تعمیم یافته (GAM)		ماشین بردار پشتیبان (SVM)			
	انحراف میانگین	انحراف معیار	انحراف میانگین	انحراف معیار	کرنل تابع شعاعی	کرنل چند جمله‌ای	کرنل خطی	انحراف میانگین معیار
سطح زیر منحنی (AUC)	0.79	0.03	0.82	0.04	0.02	0.76	0.77	0.02
شاخص کاپا (Kappa)	0.48	0.06	0.53	0.07	0.05	0.43	0.45	0.07



شکل ۳. منحنی ROC و سطح زیر منحنی (AUC) مربوط به مدل‌های (a) SVM(RBF)، (b) GAM، (c) GLM

۲،۳. بررسی اهمیت نسبی متغیرها

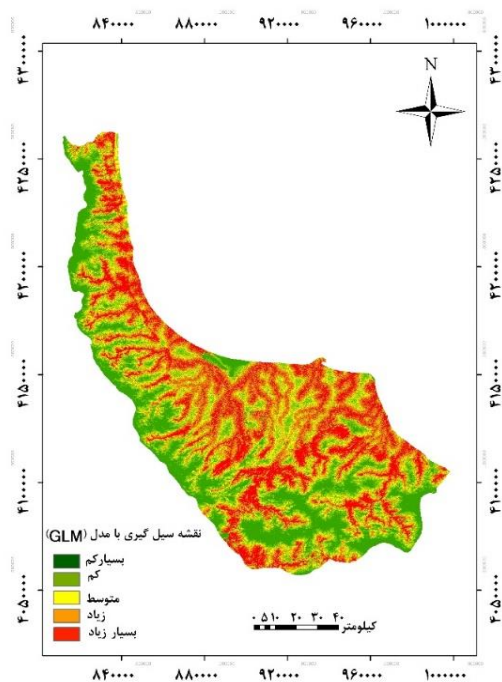
با توجه به عملکرد بالای مدل SVM در پیش‌بینی مناطق سیل‌گیر، با استفاده از این مدل اهمیت نسبی متغیرهای مورد استفاده در پیش‌بینی مناطق سیل‌گیر تهیه گردید که نتایج حاصل از آن در جدول (۲) گزارش

شده است. براساس نتایج، عوامل فاصله از آبراهه، ارتفاع از سطح دریا و شیب به ترتیب با وزن‌های ۰/۲۳، ۰/۱۸ و ۰/۰۵ بیش‌ترین تأثیر و عوامل شاخص توان آبراهه و انحنا ی پروفیل به ترتیب با وزن‌های ۰/۰۹ و ۰/۰۱ کمترین تأثیر را بر سیل‌گیری منطقه مورد مطالعه داشته‌اند.

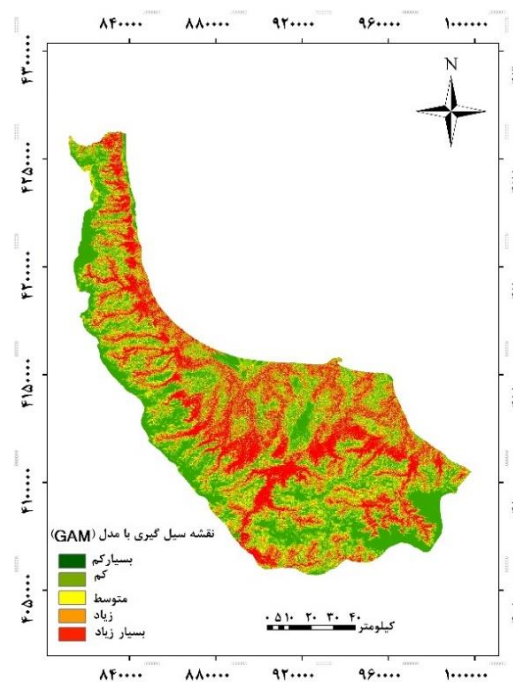
جدول ۲. نتایج حاصل از اهمیت نسبی عوامل در منطقه مورد مطالعه

مدل	ماشین بردار پشتیبان (SVM)، کرنل تابع شعاعی (RBF)
-----	--

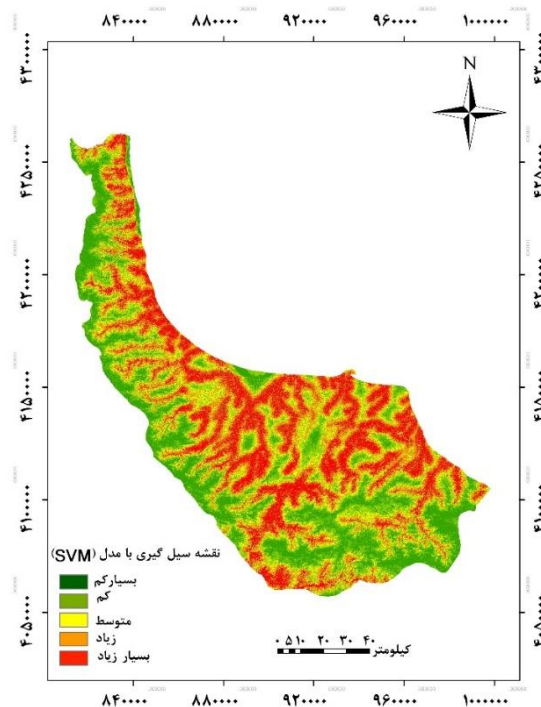
اهمیت نسبی	عامل
۰/۰۵	شیب
۰/۰۲	جهت شیب
۰/۱۸	ارتفاع از سطح دریا
۰/۰۴	کاربری اراضی
۰/۰۲	زمین شناسی
۰/۰۴	شاخص رطوبت توپوگرافی
۰/۰۰۹	شاخص توان آبراهه
۰/۰۱	انحنای پروفیل
۰/۰۳	انحنای مسطح
۰/۰۳	تراکم زهکشی
۰/۲۳	فاصله از آبراهه



شکل ۵. نقشه سیل گیری با مدل GLM منطقه مورد مطالعه



شکل ۴. نقشه سیل گیری با مدل GAM منطقه مورد مطالعه



شکل ۶. نقشه سیل گیری با مدل SVM منطقه مورد مطالعه

نشان دهنده آن است که مدل های مختلف می توانند نمره بالا را در یک جنبه خاص به دست آورند و در سایر جنبه ها نمره پایین تری به خود اختصاص دهند [۱۲]. نتایج مقایسه کرنل های در مدل SVM نشان داد که کرنل تابع شعاعی (RBF) عملکرد بهتری نسبت به کرنل های چندجمله ای و خطی دارد. نتایج مقایسه بین سه مدل نامبرده نشان داد که مدل SVM با کرنل تابع شعاعی (RBF) نسبت به سایر مدل ها عملکرد بالاتری داشته که نتایج تحقیق حاضر با نتایج تحقیقات [۲۹، ۳۱، ۲، ۲۰، ۳۳] مطابقت دارد. بهتر بودن عملکرد مدل SVM ناشی از ماهیت خود مدل بوده است، این مدل از یک ابزار به نام کرنل در تحلیل داده های ورودی استفاده می کند و هر کرنل از پارامتر و مؤلفه های مختلفی در فرایند مدل سازی استفاده می کنند که باعث افزایش توانایی مدل شده است. عملکرد متفاوت مدل ها نشان دهنده حساس بودن مدل ها به داده های ورودی می باشد. داده های هیدرولوژیکی ابعاد

۴. بحث و نتیجه گیری

در تحقیق حاضر به منظور تهیه نقشه سیل گیری از سه مدل GLM، GAM و SVM استفاده شد. عملکرد این مدل ها بر اساس داده های وقوع و عدم وقوع سیل بوده که این داده ها به راحتی قابل جمع آوری می باشند. در این تحقیق ۱۰ عامل درجه شیب، جهت شیب، شکل شیب، ارتفاع از سطح دریا، فاصله از رودخانه، تراکم زهکشی، زمین شناسی، کاربری اراضی، شاخص رطوبت توپوگرافی و شاخص توان آبراهه به عنوان عوامل مؤثر بر سیل گیری استفاده شدند. از میان عوامل مؤثر در دبی سیلاب، ویژگی های فیزیکی حوزه آبخیز به سبب پایداری و تغییرپذیری بسیار ناچیز به سایر عوامل از اهمیت خاصی برخوردارند. سپس هر یک از مدل های SVM، GLM و GAM بار اجرا و با استفاده از شاخص های سطح زیر منحنی (AUC) و شاخص کاپا (Kappa) ارزیابی شدند. بهره گیری از شاخص های مختلف برای ارزیابی مدل ها

متغیرها در نقشه سیل‌گیری تحت تأثیر روش مورد استفاده در تحقیق و ویژگی‌های منطقه مورد مطالعه می‌باشد. به عبارت دیگر شرایط مختلف زمین‌شناسی، توپوگرافی و آب و هوای یک منطقه، اولویت عوامل مؤثر در تهیه نقشه سیل‌گیری را تغییر می‌دهد. بر اساس نتایج حاصل، می‌توان اقدامات مدیریتی مناسب برای کاهش خسارات و تلفات ناشی از سیل را انجام داد. روند رو به افزایش سیل در سال‌های اخیر حاکی از آن است که اکثر مناطق کشور در معرض سیلاب‌های ادواری و مخرب قرار دارند، از این نظر بسیاری از شهرها، روستاها، تأسیسات صنعتی و کشاورزی و اماکن مسکونی در معرض خطر سیل‌گیری قرار گرفته‌اند. لذا شناسایی اصولی مناطق خطر ساز و دارای پتانسیل سیل‌گیری در منطقه از جمله اقدامات بسیار مهم در کنترل سیل و کاهش خسارات ناشی از آن محسوب می‌شود. همچنین برنامه‌ریزی برای مهار و مدیریت سیلاب‌ها، حفاظت از کیفیت و بهره‌برداری مناسب از آن‌ها مستلزم درک صحیح و دقیق از مدل‌سازی سیل‌گیری می‌باشد. مهم‌ترین عاملی که در این زمینه مدل‌سازها باید آن را در نظر بگیرند، انتخاب مدل مناسب است که بتواند گستره وسیعی از سیلاب‌ها را به طور صحیح شبیه‌سازی کند. بنابراین استفاده از مدل‌های فردی و تلفیقی برای دستیابی به نتایج بهتر برای مطالعات آینده پیشنهاد می‌شود تا نتایج حاصل از آن‌ها با اطمینان بیشتری در کارهای مدیریتی و اجرایی مورد استفاده قرار گیرد.

ز یادی را دارا بوده و روابط غیر خطی و پیچیده‌ای بین متغیرها حاکم می‌باشد و همچنین نبود داده‌های کافی برای پوشش کامل متغیرهای اندازه‌گیری شده باعث می‌شود که روش‌های آماری مانند GLM توانایی کافی برای کشف روابط و الگوهای مابین متغیرهای محیطی را نداشته باشند و همین امر باعث کاهش دقت پیش‌بینی برای مدل‌های خطی می‌شود [۷]. مدل GAM یک مدل ناپارامتری بوده و از فنون هموارساز برای توابع غیرخطی استفاده می‌نماید، به همین دلیل اجازه می‌دهد تا مدل‌سازی نزدیک‌تر به داده‌ها صورت بگیرد، در حالی که مدل GLM معمولاً ترکیبی از شرایط خطی، درجه دوم و یا درجه سوم استفاده می‌کند، اما به علت آن که مدل GAM انعطاف پذیری بیشتری دارد، نسبت به مدل GLM توانایی بالاتری در مدل‌سازی واکنش‌های پیچیده هیدرولوژیکی را دارا می‌باشد [۸]. همچنین اولویت‌بندی عوامل مؤثر نشان داد که عوامل فاصله از رودخانه، ارتفاع از سطح دریا و شیب به ترتیب بیشترین تأثیر و عوامل شاخص توان آبراهه و انحنا پروفیل کمترین تأثیر را بر سیل‌گیری منطقه مورد مطالعه داشته‌اند، که با نتایج [۱۳، ۳۱] مطابقت دارد. به دلیل این که در فاصله‌های نزدیک به آبراهه سیل‌گیری بیش‌تر است [۳۰]. سیل در مناطقی که دارای ارتفاعات کم بوده تشکیل شده و با افزایش ارتفاع تشکیل سیل کم‌تر می‌شود [۲۷]. شیب‌گیری در شیب‌های پایین تشکیل شده و در شیب‌های بالا سیل‌خیزی اتفاق می‌افتد. بنابراین اهمیت

References

- [1] Akbarzadeh, A., Jaimand, K., Hemmati, A. and Khanjani shiraz, B. (2009). Medicinal plants of Gilan province and their applications. Iranian Journal of Medicinal and Aromatic Plants, 26(3), 326-347. (In Persian)
- [2] Behzad, M., Asghari, K. and Coppola Jr, E. A. (2009). Comparative study of SVMs and ANNs in aquifer water level prediction. Journal of Computing in Civil Engineering, 24(5), 408-413.
- [3] Bubeck, P., Botzen, W. J. and Aerts, J. C. (2012). A review of risk perceptions and other factors that influence flood mitigation behavior. Risk Analysis, 32(9), 1481-1495.

- [4] Cao, C., Xu, P., Wang, Y., Chen, J., Zheng, L., and Niu, C. (2016). Flash flood hazard susceptibility mapping using frequency ratio and statistical index methods in coalmine subsidence areas. *Sustainability*, 8(9), 948.
- [5] Chapi, K., Singh, V. P., Shirzadi, A., Shahabi, H., Bui, D. T., Pham, B. T. and Khosravi, K. (2017). A novel hybrid artificial intelligence approach for flood susceptibility assessment. *Environmental Modelling Software*, 95, 229-245.
- [6] Costache, R. and Bui, D. T. (2019). Spatial prediction of flood potential using new ensembles of bivariate statistics and artificial intelligence: A case study at the Putna river catchment of Romania. *Science of the Total Environment*, 691, 1098-1118.
- [7] De'ath, G. and Fabricius, K. E. (2000). Classification and regression trees: a powerful yet simple technique for ecological data analysis. *Ecology*, 81(11), 3178-3192.
- [8] Elith, J., Graham, C. H., Anderson, R. P., Dudík, M., Ferrier, S., Guisan, A., Hijmans, R., Huettmann, F. R., Leathwick, J., Lehmann, A. and Li, J. (2006). Novel methods improve prediction of species' distributions from occurrence data. *Ecography*, 129-151.
- [9] Esmali, A. and Abdollahi, Kh., (2012). *Watershed Management and Soil Conservation*. First Edition, Mohaghegh Ardebili University press. (In Persian).
- [10] Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters*, 27(8), 861-874.
- [11] Falah, F., Rahmati, O., Rostami, M., Ahmadisharaf, E., Daliakopoulos, I. N. and Pourghasemi, H. R. (2019). Artificial Neural Networks for Flood Susceptibility Mapping in Data-Scarce Urban Areas. In *Spatial Modeling in GIS and R for Earth and Environmental Sciences* (pp. 323-336). Elsevier.
- [12] Fukuda, S., De Baets, B., Waegeman, W., Verwaeren, J., and Mouton, A. M. (2013). Habitat prediction and knowledge extraction for spawning European grayling (*Thymallus thymallus* L.) using a broad range of species distribution models. *Environmental Modelling and Software*, 47, 1-6.
- [13] Hasani Loor, S.M., (2016). Comparing Efficiency of Random Forest and Adaptive Neuro-fuzzy Inference System Models for Flood Susceptibility Mapping in the Guilan Province. Faculty of Natural Resources and Marine Sciences, Tarbiat Modares University, 92p. (In Persian).
- [14] Hastie, T. and Tibshirani, R., (1989). Non-parametric logistic and proportional odds regression. *Applied Statistics*, 260-276.
- [15] Hofmann, T., Schölkopf, B. and Smola, A. J. (2008). Kernel methods in machine learning. *The Annals of Statistics*, 1171-1220.
- [16] Khosravi, K., Shahabi, H., Pham, B. T., Adamowski, J., Shirzadi, A., Pradhan, B., Dou, J., Ly, H.B., Gróf, G., Ho, H.L. and Hong, H. (2019). A comparative assessment of flood susceptibility modeling using Multi-Criteria Decision-Making Analysis and Machine Learning Methods. *Journal of Hydrology*, 573, 311-323.
- [17] Lee, M. J., Kang, J. E. and Jeon, S. (2012). Application of frequency ratio model and validation for predictive flooded area susceptibility mapping using GIS. In *Geoscience and Remote Sensing Symposium, 2012 IEEE International*, 895-898.
- [18] Levy, J. K., Hartmann, J., Li, K. W., An, Y. and Asgary, A. (2007). Multi-Criteria decision Support Systems for Flood Hazard Mitigation and Emergency Response in Urban Watersheds *Journal of the American Water Resources Association*, 43(2), 346-358.
- [19] Mind'je, R., Li, L., Amanambu, A. C., Nahayo, L., Nsengiyumva, J. B., Gasirabo, A. and Mindje, M. (2019). Flood susceptibility modeling and hazard perception in Rwanda. *International Journal of Disaster Risk Reduction*, 38:101211.
- [20] Moharrampour, M., Sohrabi, S. and Vakili, J. (2013). Comparison of Support Vector Machines (SVM) and Autoregressive integrated moving average (ARIMA) in daily flow forecasting. *Journal of River Engineering*, 1.
- [21] Monserud, R. A., and Leemans, R. (1992). Comparing global vegetation maps with the Kappa statistic. *Ecological Modelling*, 62(4): 275-293.
- [22] Mouton, A.M., De Baets, B. and Goethals, P.L., (2010). Ecological relevance of performance criteria for species distribution models. *Ecological Modelling*, 221(16), pp.1995-2002.

- [23] Nohani, E., Darabi, F., Maroufinia, E. and Khosravi, Kh. (2016). Evaluation of Shannon Entropy in Order to Flood Probability and Susceptibility Mapping At Haraz Catchment. *Journal of Natural Environmental Hazards*, 5(10), 99-116.
- [24] Oh, H. J. and Pradhan, B., (2011). Application of a neuro-fuzzy model to landslide-susceptibility mapping for shallow landslides in a tropical hilly area. *Computers and Geosciences*, 37(9), 1264-1276.
- [25] Périé, C. and de Blois, S. (2016). Dominant forest tree species are potentially vulnerable to climate change over large portions of their range even at high latitudes. *PeerJ*, 4, e2218.
- [26] Rahmati, O. and Pourghasemi, H. R. (2017). Identification of Critical Flood Prone Areas in Data-Scarce and Ungauged Regions: A Comparison of Three Data Mining Models. *Water Resources Management*, 31(5), 1473-1487.
- [27] Rahmati, O., Pourghasemi, H. R. and Zeinivand, H. (2015). Flood susceptibility mapping using frequency ratio and weights-of-evidence models in the Golastan Province, Iran. *Geocarto International*, 31(1), 42-70. (In Persian)
- [28] Salehi, M., Vazirinasab, H., Khoshgam, M. and Rafati, N. (2012). Application of the generalized additive model in determination of the retinopathy risk factors relation types for Tehran diabetic patients. *Razi Journal of Medical Sciences*, 19(97), 1-9. (In Persian)
- [29] Suliman, A., Nazri, N., Othman, M., Abdul, M. and Ku-Mahamud, K. R. (2013). Artificial neural network and support vector machine in flood forecasting: A review. In *Proceedings of the 4th International Conference on Computing and Informatics, ICOCI* (pp. 28-30).
- [30] Tehrany, M. S., Pradhan, B. and Jebur, M. N. (2013). Spatial prediction of flood susceptible areas using rule based decision tree (DT) and a novel ensemble bivariate and multivariate statistical models in GIS. *Journal of Hydrology*, 504(2013), 69-79.
- [31] Tehrany, M. S., Pradhan, B. and Jebur, M. N. (2015). Flood susceptibility analysis and its verification using a novel ensemble support vector machine and frequency ratio method. *Stochastic Environmental Research and Risk Assessment*, 29(4), 1149-1165.
- [32] Termeh, S. V. R., Kornejady, A., Pourghasemi, H. R. and Keesstra, S. (2018). Flood susceptibility mapping using novel ensembles of adaptive neuro fuzzy inference system and metaheuristic algorithms. *Science of the Total Environment*, 615, 438-451.
- [33] Tien Bui, D., Pradhan, B., Lofman, O. and Revhaug, I. (2012). Landslide susceptibility assessment in vietnam using support vector machines, decision tree, and Naive Bayes models. *Mathematical Problems in Engineering*, 2012.
- [34] Tien Bui, D., Khosravi, K., Li, S., Shahabi, H., Panahi, M., Singh, V., Chapi, K., Shirzadi, A., Panahi, S., Chen, W. and Bin Ahmad, B. (2018). New hybrids of anfis with several optimization algorithms for flood susceptibility modeling. *Water*, 10(9), 1210.
- [35] Vapnik, V.N., (1999). *The Nature of Statistical Learning Theory*. 2nd Edition, Springer Verlag, New York, 314p.
- [36] Zhao, G., Pang, B., Xu, Z., Yue, J. and Tu, T. (2018). Mapping flood susceptibility in mountainous areas on a national scale in China. *Science of the Total Environment*, 615, 1133-1142.