

## تجزیه و تحلیل خطر سیل گیری با استفاده از روش یادگیری ماشین جنگل تصادفی (مطالعه موردی: شهر مشهد)

نرگس عرب<sup>۱</sup>، عبدالرسول سلمان ماهینی<sup>۲\*</sup>، علیرضا میکائیلی تبریزی<sup>۳</sup>، توما ویتته<sup>۴</sup>

۱. دکتری آمایش محیط زیست، دانشکده شیلات و محیط زیست، دانشگاه گرگان

۲. استاد، دانشکده شیلات و محیط زیست، دانشگاه گرگان

۳. دانشیار، دانشکده شیلات و محیط زیست، دانشگاه گرگان

۴. دانشیار، دانشکده جغرافیا، دانشگاه رن ۲

(تاریخ دریافت ۱۴۰۱/۰۲/۲۸؛ تاریخ بازنگری ۱۴۰۱/۰۵/۲۸؛ تاریخ تصویب ۱۴۰۱/۰۹/۲۸)

### چکیده

سیل از رایج ترین بلایای طبیعی است و خسارت های مالی و جانی فراوانی به جای می گذارد. اگرچه میزان بارندگی در بسیاری از مناطق ایران کم است، در بسیاری از مناطق، بیشترین میزان بارندگی سالانه تنها در یک روز یا مدت کوتاهی رخ می دهد که منجر به سیل می شود. آب روان در جریان سیل به دلیل ساختار زمین شناسی و همچنین، تخریب اکوسیستم می تواند بسیار آلوده باشد و اغلب گل ولای زیادی به همراه دارد که بر خسارت های سیل می افزاید. برای کاهش خسارت های احتمالی سیل، برنامه ریزان و تصمیم گیرندگان باید از زمان و مکان وقوع سیل آگاه باشند. این امر مستلزم استفاده از روش های جدید پیش بینی سیل و جلوگیری از خسارت های آن است. در این مطالعه، از روش یادگیری ماشین درخت تصادفی یا Random Forest (RF) برای پیش بینی مکان وقوع سیل در شهر مشهد استفاده شد و عملکرد آن مورد بررسی قرار گرفت. همچنین تأثیر هر یک از عوامل ارتفاع و شیب متوسط حوضه، جهت شیب، شاخص رطوبت توپوگرافی، شاخص خشکسالی، فاصله از آبراهه ها، زمین شناسی، کاربری اراضی، تراکم آبراهه ها، آبراهه ها و میزان بارش حداکثر متوسط سالانه در این پیش بینی مورد بررسی قرار گرفت. نتایج ارزیابی خروجی مدل RF نشان داد مقدار  $AUC=0.95$  درصد است. به طور کلی، نتایج نشان داد مدل RF دارای دقت زیادی در تعیین مناطق حساس به وقوع سیل در حوضه شهر مشهد است.

**کلیدواژگان:** ارزیابی خطر سیل گیری، جنگل تصادفی، یادگیری ماشین، شهر مشهد، پهنه بندی سیل.

## ۱. مقدمه

بررسی جنبه‌های مختلف زندگی انسان روی زمین نشان می‌دهد به نوعی با پدیده‌های اقلیمی مرتبط است و هنگام برنامه‌ریزی برای زندگی انسان، شناخت آن‌ها یک ضرورت اساسی است. سیل از رایج‌ترین بلایای طبیعی است که بر اساس گزارش پایگاه بین‌المللی خطرات بلایای طبیعی، بیشترین میزان زلزله و خشکسالی از نظر خسارت‌های مالی و جانی را در خود جای داده است. سیل یک خطر طبیعی است که همواره جوامع بشری را تهدید می‌کند. سیلاب‌های طبیعی معمولاً پس از بارندگی شدید اتفاق می‌افتد که بیشتر از شکستن سدها که یکی از عوامل سیل مصنوعی است، رخ می‌دهد. با گسترش سریع شهرها و رشد

جمعیت، بسیاری از مناطق با خطر سیلاب روبه‌رو هستند. سیلاب‌های شهری از شدیدترین بلایای مربوط به تغییرات اقلیمی هستند که به دلیل رواناب‌های سطحی زیاد در مناطق ساخته‌شده، تهدیدی برای شرایط محیطی و حتی زندگی انسان‌ها هستند. سیلاب‌های شدید که با افزایش سریع سطح آب رواناب شهری مشخص می‌شوند به دلیل ویژگی‌های ژئومورفولوژی و مشکلات شهری ایجاد شده و همواره در حال افزایش هستند. رودخانه‌های شهری بیشتر با کاهش ظرفیت هیدرولیکی کانال‌ها روبه‌رو می‌شوند و مجبور به جاری شدن در زمین می‌شوند [۱]. سیلاب‌ها می‌توانند به خسارت اقتصادی و اجتماعی زیادی منجر شوند، زیرا به طور قابل توجهی به زیرساخت‌های شهری از جمله حمل‌ونقل، املاک و مستغلات و منابع برق و آب آسیب می‌رسانند [۲]. در سال‌های اخیر (۱۹۹۵-۲۰۱۵) سیل مسئول ۴۷ درصد از بلایای مربوط به اقلیم است و ۲/۳ میلیارد نفر را در سطح جهان تحت تأثیر قرار داده است [۳]. سیل هر دو کشورهای در حال توسعه و توسعه‌یافته را تحت تأثیر قرار داده است. به عنوان مثال، در سال ۱۹۹۸، حدود ۳۰ میلیون نفر توسط سیل در کشورهای در حال توسعه مانند بنگلادش [۴] تحت تأثیر قرار گرفته‌اند. به عنوان یک کشور توسعه‌یافته، از سال ۱۹۰۰ در ایالات متحده بیش از ۱۰ هزار نفر بر اثر سیل جان خود را از دست داده‌اند [۵]. طبق گزارش‌های آژانس محیط زیست اروپا (EEA) از سال ۱۹۸۰ تا کنون تقریباً ۱۵۰۰ سیل در اروپا رخ داده که بیش از نیمی از آن‌ها پس از سال ۲۰۰۰ بوده است. در اروپا خسارت سالانه پیش‌بینی‌شده (EAD) از طغیان رودخانه

حدود ۶/۴ میلیارد یورو است، در حالی که جمعیت سالانه پیش‌بینی‌شده (EAP) در معرض طغیان حدود ۱۹۵ هزار نفر است. طی دهه‌های گذشته، صدها نفر در نتیجه سیلاب‌ها جان خود را از دست داده‌اند [۵]. تمام این آمارها و مطالعات نشان می‌دهد سیل از برجسته‌ترین بلایای طبیعی است که هم مردم و هم اموال آن‌ها را تحت تأثیر قرار می‌دهد. به‌خوبی مشخص شده است که طغیان رودخانه‌ها توسط هر دو عوامل طبیعی و انسانی ایجاد می‌شوند. بر گویجس و همکاران (۲۰۱۷) تغییرات اخیر در طغیان شدید سیلاب‌ها در چندین قاره را بررسی کردند و دریافتند که تغییرات طولانی‌مدت در اقلیم می‌تواند منجر به تغییر در میزان وقوع سیل شود [۶]. در حالی که مدیریت و تغییر شرایط طبیعی بسیار دشوار است، تعداد زیادی از مقالات اشاره کرده‌اند که سیلاب‌های شهری نیز با عوامل مختلف انسانی مانند سرمایه‌گذاری در سیستم زهکشی با ساخت سطوح نفوذناپذیر و برنامه‌ریزی پیکربندی پوشش زمین، بسیار مرتبط است. با این حال، تعدادی از مطالعات استدلال می‌کنند که سیستم‌های زهکشی در بسیاری از کشورهای در حال توسعه کمک‌چندانی به مدیریت خطر سیل نمی‌کنند. آن‌ها دریافتند که سیستم‌های تخلیه آب باران به دلیل طراحی استاندارد پایین و نگهداری ناکافی به‌سختی می‌توانند با حوادث شدید سیل کنار بیایند و از همه مهم‌تر، به‌روزرسانی شبکه‌های خط لوله شهری وقت‌گیر و پرمحتوا است [۷]. علاوه بر تجدید سیستم‌های زهکشی، بهبود ترکیب و پیکربندی پوشش زمین نیز برای کاهش خطرات سیل مهم است. به دلیل شهرنشینی سریع و رشد جمعیت، فضاهای سبز (به عنوان مثال، جنگل، مرتع و تالاب) که می‌توانند به طور مؤثر روان‌آب‌ها را کاهش دهند، همواره توسط سطوح نفوذناپذیر مصنوعی (به عنوان مثال سیمان، آسفالت) جایگزین می‌شوند [۸]. چنین پدیده‌ای پوشش اصلی زمین و شرایط هیدرولوژیک را تغییر می‌دهد و بنابراین، خطر طغیان رودخانه را می‌افزاید [۹]. به طور کلی، سیستم‌های زهکشی ضعیف، ظرفیت کم برای ذخیره‌سازی و نفوذ باران و رشد ناهمگن شهری، از دلایل اصلی جاری شدن سیل در مناطق شهری است. علاوه بر این، تغییرات اقلیمی و کاربری اراضی خطرات ناشی از سیلاب شهری را افزایش می‌دهد که باید در مدیریت سیلاب در نظر گرفته شود [۱۰]. با این حال، غالباً

سنجش از دور (RS) و سیستم اطلاعات جغرافیایی (GIS) علم هیدرولوژی، به‌ویژه مدیریت سیل را متحول کرده است. در بیشتر مطالعات سیل از GIS برای ارزیابی خطر سیل استفاده شده است. یوسف و دیگران (۲۰۱۱) از تکنیک‌های سنجش از دور در یک محیط GIS استفاده کردند تا خطر سیلاب را تخمین بزنند [۱۴]. به‌تازگی دارابی و همکاران (۲۰۱۹) برای ارزیابی خطر سیل در شهر ساری دو مدل یادگیری ماشین، روش الگوریتم ژنتیک و تولید مجموعه تنظیم (GARP) و درخت سریع و بی‌طرفانه آماری (QUEST) را معرفی کرد. آن‌ها در روش‌های خود چندین فاکتور را در نظر گرفتند و دریافتند که فاصله کانال و تراکم جمعیت در بیشترین خطر سیل تأثیر دارد [۱۵]. تحلیل مبتنی بر تصمیم چندمعیاره مانند فرایند تحلیلی سلسله‌مراتبی [۱۶] در GIS نیز برای ارزیابی خطر سیل رایج است. از میان مدل‌های مختلف سیل مبتنی GIS، شبکه‌های عصبی مصنوعی، نسبت فرکانس (FR) [۱۷]، رگرسیون لجستیک [۱۸] سیستم استنتاج فازی مبتنی بر شبکه تطبیقی [۱۹]، درخت تصمیم‌گیری و ماشین‌های بردار پشتیبان (SVMs) [۲۰] گسترده‌ترین تکنیک‌هایی هستند که از RS و GIS استفاده می‌کنند. با توسعه هوش مصنوعی، مدل‌های یادگیری ماشین (ML) شامل بردار پشتیبانی (SVM)، جنگل تصادفی (RF) و درخت تصمیم (DT)، پیشنهاد شده‌اند و در ارزیابی خطر سیل استفاده می‌شوند. مدل‌های یادگیری ماشینی از تعیین ذهنی وزن، مانند روش‌های AHP، با یادگیری رابطه بین خطر سیل و عوامل توضیحی اجتناب می‌کنند. اگرچه مدل‌های پیش‌بینی سیل زیادی وجود دارند، اما دقت نقشه‌های پیش‌بینی سیل همچنان مسئله مهمی است. در مدل‌سازی سیل، دقت زیادی برای نقشه‌های پیش‌بینی سیل باید وجود داشته باشد و بنابراین، مدل‌های جدید و کارآمدی باید بررسی شوند تا دقت را افزایش دهند. مدل‌های هیدرولیکی ممکن است منجر به عدم اطمینان شوند، زیرا به داده‌های هیدرولوژیک کامل و کافی نیاز دارند [۱۳]. بنابراین، استفاده از داده‌های RS و مدل‌های مبتنی بر GIS می‌تواند یک رویکرد مکمل برای مدل‌سازی سیل در نظر گرفته شود [۲۲]. در این مطالعه، از مدل‌های یادگیری ماشین جنگل تصادفی (RF) برای ارزیابی ریسک استفاده شد. هدف از تحقیق حاضر، تعیین میزان خطر سیل در شهر

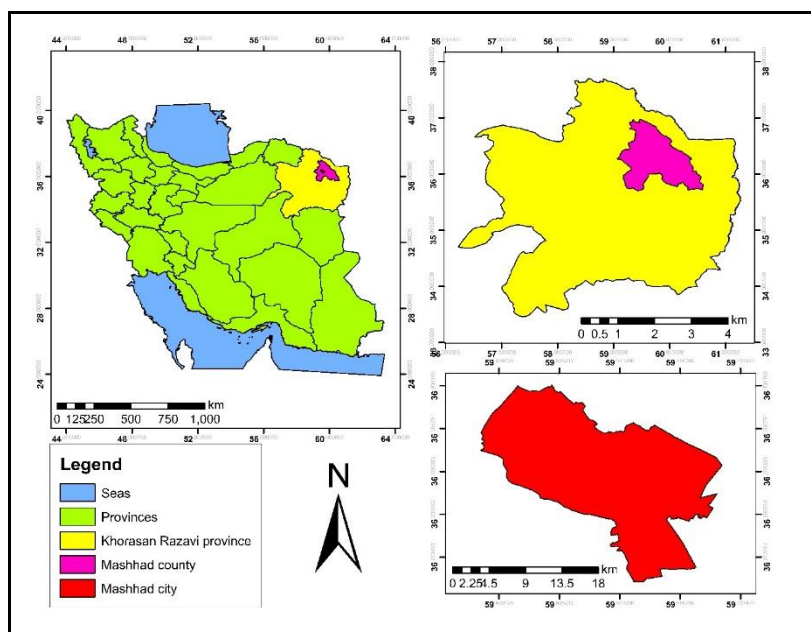
سیلاب‌های شهری در برنامه‌ریزی‌ها نادیده گرفته می‌شوند، اگرچه تأثیر بسزایی در جوامع شهری دارد. عدم برنامه‌ریزی شهری در فرایند شهرنشینی دلیل دیگری است که باعث آسیب‌پذیری مناطق شهری در برابر سیل می‌شود [۱۱]. پیش‌بینی موقعیت مناطق مستعد سیل با استفاده از نقشه‌های خطر سیل می‌تواند اثرات سیل را در مناطق شهری کاهش دهد و می‌تواند برای برنامه‌ریزی شهری مورد استفاده قرار گیرد. تأثیرات سیلاب‌های شهری به آسیب‌پذیری افراد آسیب‌دیده و زیرساخت‌ها بستگی دارد. آسیب‌پذیری تابعی از حساسیت یک سیستم شهری در برابر خطر سیل و کمبود توانایی واکنش و سازگاری است. با وجود بهبود روش‌های استفاده‌شده در نقشه‌برداری خطر سیل، دانش آسیب‌پذیری به عنوان از مهم‌ترین چالش‌ها در ارزیابی خطر سیل باقی مانده است [۱۰]. تجزیه و تحلیل ریسک سیل و کاهش خطرات، دو مؤلفه مدیریت خطر سیل است. تجزیه و تحلیل خطر سیل برای بررسی مناطقی که خطر وقوع سیل به طور غیر قابل قبولی زیاد است، مورد نیاز است. بنابراین، تجزیه و تحلیل جامع خطر سیل با شناسایی مناطق خطرپذیر و دارای ریسک، بخشی اساسی در مدیریت ریسک برای تخمین میزان خسارت‌هایی است که می‌تواند به دلیل جاری شدن سیل رخ دهد. مطالعات زیادی در مورد سیل و زمینه‌های مختلف مرتبط با آن انجام شده است. چن گوانگلی و همکاران (۲۰۱۶) از الگوریتم کلونی مورچه‌ها برای استخراج منطقه خطر سیل در حوضه رودخانه دونگ جینگ در جنوب چین بر اساس استخراج قانون (مورچه کیش) استفاده کرد. بر اساس نتایج مطالعه آن‌ها، مناطق پرخطر با نقشه‌های خطر و سیل‌هایی که طی تاریخ رخ داده، مطابقت داده شده است [۱۲]. خطر فاجعه سیل اساساً مفهومی سه‌بعدی از عدم قطعیت و پیچیدگی است. این تعدد، پیچیدگی، عدم قطعیت، عدم دقت و تنوع روش‌های ارزیابی است که ارزیابی خطر سیل را ارائه می‌کند، که یک مشکل جهانی در علوم طبیعی و فناوری است. تجزیه و تحلیل سیل برای بررسی مکان‌هایی انجام می‌شود که در آن خطر سیل قبول نیست و اقدامات کاهش خطر مورد نیاز است. بنابراین، تجزیه و تحلیل جامع خطر سیل با شناسایی مناطق خطرناک، بخش اساسی مدیریت ریسک برای برآورد میزان خسارتی است که ممکن است بر اثر سیل رخ دهد [۱۳]. پیشرفت‌های اخیر در کارایی فناوری‌های

مشهد است. نتیجه این تحقیق تصویری کلی و دقیق از مناطق آسیب‌پذیر سیل با استفاده از داده‌های سنجش از دوری با رویکرد مکانی مبتنی بر یادگیری ماشین ارائه می‌دهد. این مقاله روش پیشرفته یادگیری ماشین جنگل تصادفی را برای تهیه نقشه خطر سیلاب شهر مشهد ارزیابی می‌کند. این ارزیابی خطر سیل برای مناطق شهری ضروری است، زیرا به سیاستگذاران کمک می‌کند تا سطح ریسک سیل و تأثیرات آن را بر ساختمان‌ها و جمعیت‌های مسکونی و تجاری را بشناسند و سپس اقدامات حفاظتی را طراحی کنند که به کاهش تلفات و خسارت‌ها کمک می‌کند. ارزیابی ریسک سیلاب‌های شهری و انتشار این اطلاعات به همه ذی‌نفعان در مدیریت سیلاب شهری اهمیت زیادی دارد.

## ۲. مواد و روش

### ۲.۱. منطقه مورد مطالعه

مطالعه حاضر در شهر مشهد، یکی از کلان‌شهرهای مهم ایران با آب‌وهوای سرد و خشک با میانگین دمای سالانه



شکل ۱. موقعیت منطقه مورد مطالعه

استفاده از یک پایگاه داده دقیق GIS می‌تواند بر دقت و کیفیت روش‌ها تأثیر بگذارد. انتخاب عوامل مختلف به شرایط و ویژگی‌های هر منطقه بستگی دارد. با این حال، دستورالعمل خاصی برای انتخاب عوامل متأثر از سیل در مناطق شهری وجود ندارد. داده‌های مورد استفاده در این

## روش‌ها

در سال‌های اخیر روش‌های مکانی زیادی توسط محققان برای تهیه نقشه حساسیت سیل و مناطق مستعد سیل‌گیری ارائه شده است. برای ساخت مدل ارزیابی سیل نیاز به پارامترهای مؤثر بر سیل وجود دارد [۲۱]. نحوه

شیب کمتر به این معناست که زمین نسبتاً مسطح است و خطر سیلاب بیشتری در آن وجود دارد.

#### جهت

نقشه جهت نیز در ارزیابی سیل‌گیری منطقه نقش دارد (شکل ۲. C). جهت بر جاری شدن سیلاب تأثیر می‌گذارد، چرا که خود بر رطوبت خاک اثر می‌گذارد [۲۹] و از این جنبه به طور غیرمستقیم طغیان را تحت تأثیر قرار می‌دهد. به عنوان مثال، مناطق شیب‌دار در سایه، که در آن رطوبت خاک نسبتاً زیاد است، رواناب زیادی تولید می‌کند.

#### شاخص رطوبت توپوگرافی (TWI)

شاخص رطوبت توپوگرافی (TWI)، تأثیر توپوگرافی را بر تولید رواناب و میزان تجمع جریان در هر مکان در حوضه رودخانه نشان می‌دهد. این فرایند در (شکل ۲. D) نشان داده شده است. این شاخص ابزار مفیدی برای تعیین وضعیت رطوبت خاک و یکی از شاخص‌های مؤثر بر وقوع سیل احتمالی در حوضه آبریز است [۳۰]. شاخص رطوبت طبق رابطه زیر تعریف می‌شود. در این شاخص، TWI شاخص رطوبت توپوگرافی،  $As$  سطح ویژه حوضه و  $\beta$  شیب زمین (بر حسب درجه) است.

#### کاربری اراضی

یکی دیگر از عوامل اصلی مرتبط که به شدت در جاری شدن سیل نقش دارد، کاربری اراضی و پوشش منطقه (Land Use Land Cover) است که به اختصار LULC نامیده می‌شود. درک دقیق LULC برای خطرات محیط زیستی و طبیعی بسیار ضروری است [۳۱]. انواع کاربری‌های زمین بر سیل مناطق شهری تأثیر می‌گذارد. به عنوان مثال، ساخت‌وسازها و ساختمان‌ها در مناطق شهری با از بین بردن پوشش گیاهی، رواناب سطحی حاصل از بارندگی را افزایش می‌دهد. بنابراین میزان نفوذ آب به الگوهای استفاده از زمین بستگی دارد. مناطق پوشیده از گیاه به دلیل همبستگی منفی بین وقوع سیل و تراکم پوشش گیاهی، کمتر در معرض سیل قرار دارند. با این حال، مناطق شهری معمولاً از سطوح نفوذناپذیر و زمین‌های لخت تشکیل شده است که باعث افزایش رواناب‌های سیل آسا می‌شود. بنابراین نقشه کاربری اراضی نقش مهمی در مدل‌سازی خطر سیل ایفا می‌کند و به

تحقیق شامل ۹ معیار است. معیارهای مؤثر در این تحقیق با نظر کارشناسان و مرور مطالعات پیشین تعیین شدند. معیارهای مؤثر عبارت‌اند از: ارتفاع، شیب، جهت شیب، شاخص رطوبت توپوگرافی، شاخص خشکسالی، فاصله از آبراهه‌ها، زمین‌شناسی، کاربری اراضی، تراکم آبراهه‌ها، آبراهه‌ها و میزان بارش. در زیر به بررسی هر یک از این عوامل پرداخته می‌شود.

#### ارتفاع

یکی از پارامترهای مهم تأثیرگذار در مطالعات سیل، ارتفاع از سطح زمین است (شکل ۲. A). وقوع سیل در مناطق بسیار مرتفع تقریباً غیرممکن است. آب از زمین‌های بسیار مرتفع به سمت مناطق پایین‌تر جریان می‌یابد و بنابراین، احتمال وقوع سیل به طور طبیعی در مناطق مسطح بیشتر است. علاوه بر این، پارامترهای توپوگرافی که مستقیم تحت تأثیر میزان جریان و سرعت رواناب قرار می‌گیرند، نقش مهمی در وقوع سیل دارند [۴]. همه پارامترهای توپوگرافی مربوط به وقوع سیل در هر منطقه مستقیم از نقشه مدل رقومی ارتفاع استخراج می‌شوند. بنابراین در این مطالعات نقشه مدل رقومی ارتفاع باید دقیق باشد. در این تحقیق، از نقشه مدل رقومی ارتفاع (DEM) با وضوح ۱۰ متری استفاده شد [۲۴]. ارتفاع در پنج کلاس طبقه‌بندی شد و با توجه به سطح خطرات مقادیر مناسبی به آن‌ها اختصاص داده شد. نواحی بسیار مرتفع همان‌طور که دارای مقدار کمتری هستند، ریسک کمتر و مناطق مسطح مقادیر بیشتری را دریافت می‌کنند، زیرا دارای خطر بیشتری هستند [۱۱].

#### شیب

شیب یکی دیگر از پارامترهای توپوگرافی است که به دلیل تأثیر آن در تولید و سرعت رواناب در یک منطقه به عنوان پارامتری مهم در هیدرولوژی شناخته می‌شود [۲۵]. خطر آبگرفتگی نیز به شیب بستگی دارد [۲۶ و ۲۷]. شیب‌های تند سرعت رواناب سطحی را افزایش می‌دهند و در آن‌ها میزان نفوذ آب کاهش می‌یابد. از این رو، مناطق پایین‌تر با شیب‌های کمتر و پهنه‌های مسطح رواناب بیشتری را تولید می‌کنند و خطر سیل را افزایش می‌دهند [۲۸]. در این تحقیق، درصد شیب در پنج دسته طبقه‌بندی شد (شکل ۲. B). در این طبقه‌بندی درصد شیب بیشتر به معنای خطر کمتری از نظر سیل‌گیری است. به عکس، درصد

### تراکم آبراهه‌ها

در مناطقی با تراکم زهکشی بیشتر، به دلیل فراوانی زیاد رواناب‌ها که از مناطق مختلف به یک نقطه وارد می‌شوند، امکان سرریز شدن جوی‌ها و کانال‌ها هنگام وقوع سیل بیشتر است و ممکن است بر اثر یک بارش شدید این مناطق دچار آب‌گرفتگی بیشتری نسبت به مناطقی که تراکم زهکشی کمتری دارند، شوند [۳۲]. علاوه بر این، در مناطقی که تراکم زهکشی بیشتر است، پتانسیل تولید رواناب زیاد است. هر چه تراکم زهکشی بیشتر باشد، رواناب سریع‌تر خارج می‌شود، ولی بر اثر تمرکز رواناب در یک زمان از چند منطقه و کارآمد نبودن کانال‌ها در منطقه مورد مطالعه امکان آب‌گرفتگی بیشتر می‌شود [۳۲]. شکل ۲.۲ H تراکم آبراهه‌ها و شکل ۲.۱ آبراهه‌ها در منطقه را نشان می‌دهد.

### فاصله از رودخانه و آبراهه‌ها

فاصله از رودخانه یکی از عوامل مهم در مطالعات سیل‌گیری در نظر گرفته می‌شود. مناطقی که در نزدیکی رودخانه قرار دارند، اغلب احتمال بیشتری در وقوع سیل دارند و می‌توانند سیل‌های بیشتری را تجربه کنند. از این‌رو، فاصله تا رودخانه عامل مهمی در تعیین خطرات ناشی از سیل است [۱۶]. فاصله تا رودخانه با استفاده از GIS محاسبه شد. مناطقی که در نزدیکی رودخانه واقع شده‌اند، مناطق پرخطر برای طغیان محسوب می‌شوند و مقادیر بیشتری دریافت می‌کنند. به‌عکس، مناطقی که از رودخانه دورتر هستند، خطر کمتری دارند. بنابراین، مقادیر کمتری به این مناطق اختصاص داده می‌شود. علاوه بر رودخانه‌ها، کانال‌ها می‌توانند خطر طغیان را به دلیل احتمال طغیان داشته باشند. فاصله کانال‌ها نیز با استفاده از ابزار فاصله اقلیدسی در نرم‌افزار ArcGIS 10.3 محاسبه شد [۱۰] (شکل ۲.۲). J.

### شاخص استاندارد بارش (SPI)

SPI شاخص اندازه‌گیری قدرت فرسایش و شدت جریان سطحی است (شکل ۲.۲ K). SPI نسبت بین سطح رو به بالای یک نقطه و شیب بالادست را به گونه‌ای نشان می‌دهد که نمایه عرضی و نیمرخ طولی آبراهه را منعکس می‌کند [۳۳]. این شاخص از طریق توابع موجود در سیستم اطلاعات جغرافیایی (GIS) و با خطای انباشت

عنوان یکی از پارامترهای مهم در ارزیابی آسیب‌پذیری در نظر گرفته می‌شود. در پژوهش حاضر از تصاویر ماهواره‌ لندست برای استخراج نقشه کاربری و پوشش زمین استفاده شد. پس از انجام پیش‌پردازش‌های لازم روی تصاویر، از الگوریتم SVM برای طبقه‌بندی تصاویر در نرم‌افزار ENVI استفاده شد. نقشه نهایی کاربری اراضی در شکل ۲.۲ E مشاهده می‌شود.

### بارندگی

همان‌طور که در شکل ۲.۲ F نشان داده شده، بارش یک عامل مهم در تجزیه و تحلیل حساسیت سیل است. در این تحقیق، نقشه میانگین بارندگی سالانه با استفاده از داده‌های بارش بلندمدت (۱۳۶۷-۱۳۹۸) از ۳۲ ایستگاه باران‌سنجی با استفاده از روش کریجینگ محاسبه شد. روش کریجینگ با داشتن میزان خطای RMSE و MAE کم برای تهیه نقشه بارندگی مطلوب تشخیص داده شد. داده‌های بارندگی از سازمان هواشناسی ایران (IRIMO) و شرکت مدیریت منابع آب ایران (IDWRMC) به دست آمد. نقشه بارندگی براساس داده‌های بارندگی روزانه استخراج شده از داده‌های سازمان ملی هواشناسی مرتبط با ایستگاه‌های سطح مشهد تهیه شد.

### خاک‌شناسی، زمین‌شناسی و سنگ‌شناسی

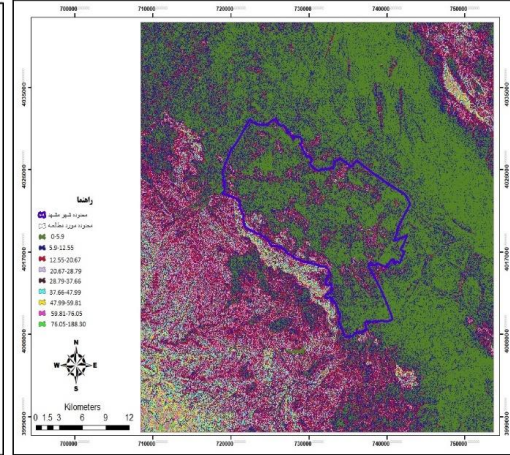
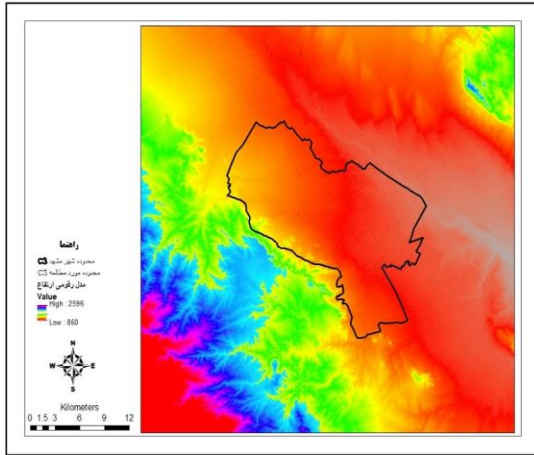
انواع خاک یکی دیگر از عوامل مؤثر بر وقوع سیل است که غیر مستقیم روی آن تأثیر می‌گذارد. انواع مختلف خاک می‌تواند سطح خطر سیل‌خیزی را افزایش یا کاهش دهد. به عنوان مثال، سنگ‌های نفوذناپذیر به رواناب سطحی کمک می‌کنند، در حالی که سنگ‌های نفوذپذیر نفوذ آب را افزایش می‌دهند [۱۱]. نقشه‌های سنگ‌شناسی و خاک پارامترهای بسیار مهمی در یافتن مناطق حساس مستعد به سیلاب هستند. نوع خاک به دلیل ویژگی‌های خاک، مانند بافت، درجه نفوذپذیری و ساختار، مستقیم بر روند زهکشی تأثیر می‌گذارد. در ارزیابی خطر سیل به اطلاعات سنگ‌شناسی در مورد نفوذپذیری سنگ‌ها نیز نیاز است. بنابراین، داشتن اطلاع از انواع خاک و سنگ‌شناسی برای انجام تجزیه و تحلیل در این تحقیق حیاتی است. شکل ۲.۲ G نقشه زمین‌شناسی را نشان می‌دهد.

شاخص SPI به شش گروه طبقه بندی می شود [۲۸]:

- $0.01 \geq SPI \geq -$
- $99.65 \geq SPI \geq 0.01$
- $199.30 \geq SPI \geq 99.66$
- $498.25 \geq SPI \geq 199.31$
- $25410.81 \geq SPI \geq 498.26$

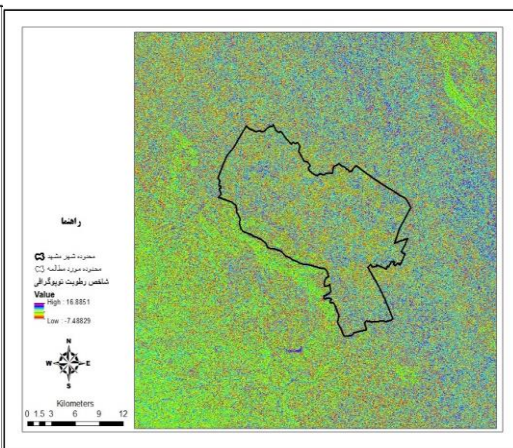
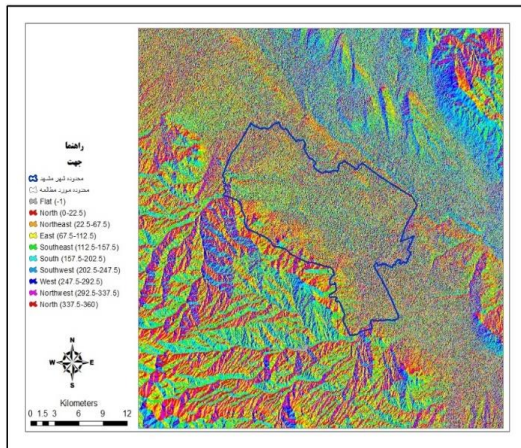
مطابق معادله زیر محاسبه می شود: در این شاخص As نشان دهنده سطح ویژه حوضه (m<sup>2</sup>/m) و نشان دهنده زاویه شیب (بر حسب درجه) است.

$$SPI = A \tan \beta$$



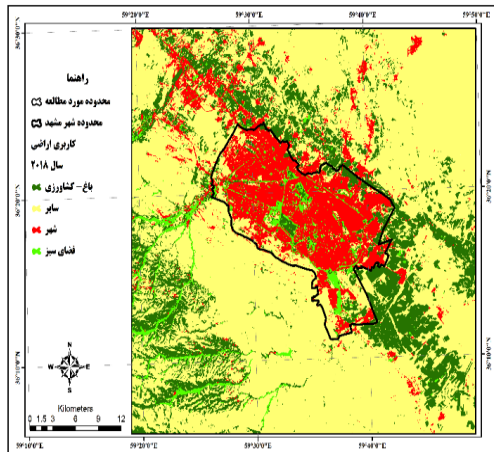
B: نقشه شیب

A: نقشه ارتفاع

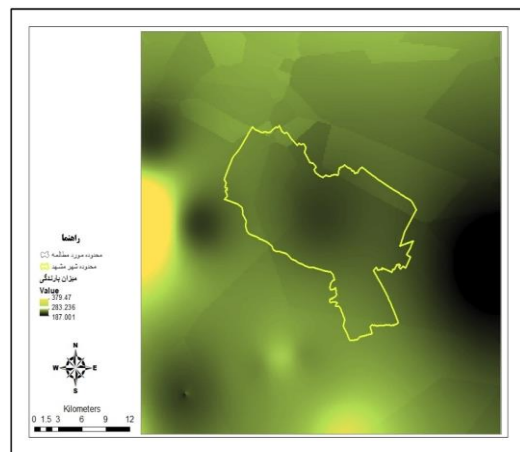


D: نقشه شاخص رطوبت توپوگرافی

C: نقشه جهت

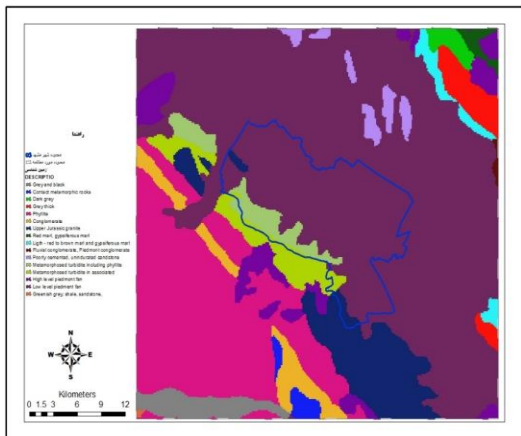


E: نقشه کاربری اراضی

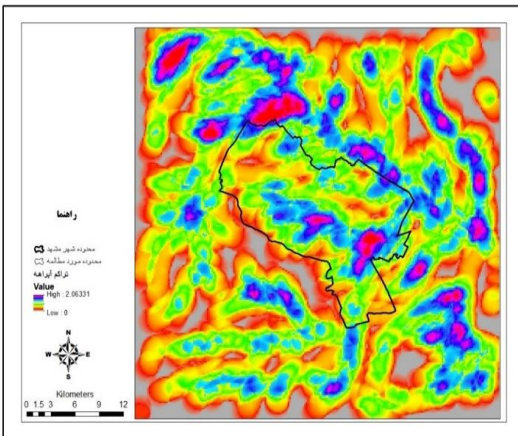


F: نقشه میزان بارندگی

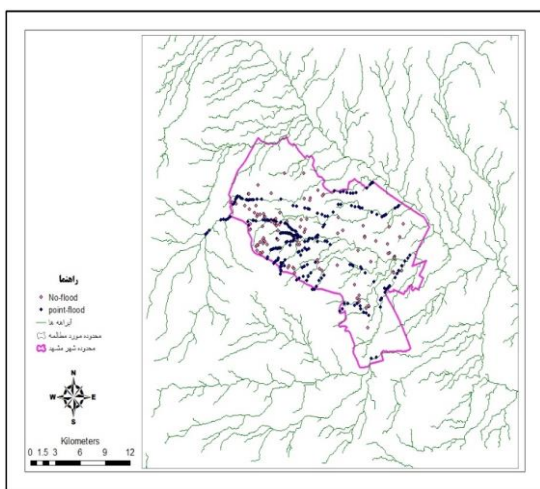




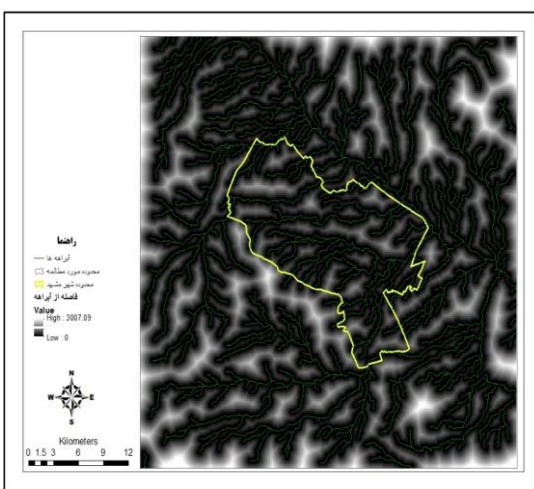
H: نقشه تراکم آبراهها



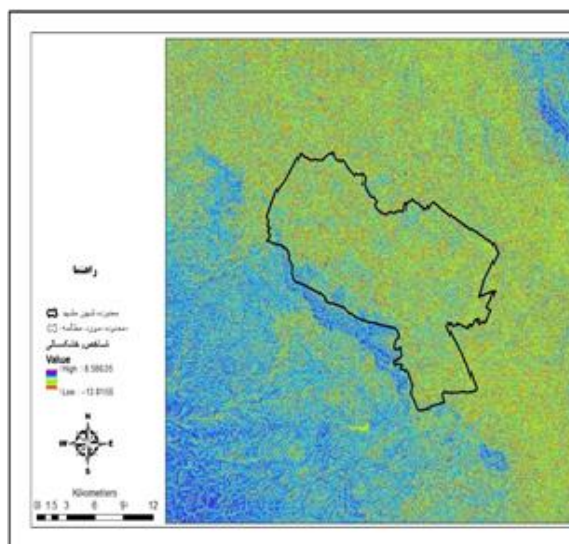
G: نقشه زمین شناسی



I: نقشه فاصله از آبراهها



J: نقشه آبراهها



K: نقشه شاخص خشکسالی

شکل ۲. نقشه‌های پایه



## ۲.۲. روش‌های یادگیری ماشین

ریشه‌های یادگیری ماشین در سنجش از دور به دهه ۱۹۹۰ برمی‌گردد. این روش‌ها ابتدا به عنوان راهی برای خودکارسازی دانش‌بنیان برای سنجش از دور معرفی شد. هوانگ و جنسن اولین بار در سال ۱۹۹۷ در مورد چگونگی ایجاد یک پایگاه دانش با استفاده از حداقل ورودی از متخصصان انسانی صحبت کردند و سپس، درختان تصمیم‌گیری برای استنباط قوانین از ورودی‌های انسانی برای سیستم خبره ایجاد شدند. قوانین تولیدشده در یک سایت مطالعاتی در رودخانه ساوانا استفاده شد. نتیجه‌گیری این مطالعات نشان می‌دهد چگونه روش یادگیری ماشین پیشنهادی به کمک سیستم خبره بیشترین دقت را نسبت به روش‌های متداول در آن زمان به دست آورده است. پس از تحولات مشابه، یادگیری ماشین به‌زودی به عنوان ابزاری مهم توسط جامعه سنجش از دور پذیرفته شد. در این پژوهش از روش جنگل تصادفی برای بررسی و تهیه نقشه خطر سیلاب در شهر مشهد استفاده می‌شود که در ذیل به تعریف این روش پرداخته می‌شود.

## ۲.۲.۱. مدل جنگل تصادفی

الگوریتم‌های جنگل تصادفی (Random Forest) یا به اختصار RF به دلیل صحت طبقه‌بندی‌هایشان، در جامعه سنجش از دور کاربرد زیادی پیدا کرده‌اند. در این الگوریتم‌ها، از چندین درخت تصمیم‌گیری استفاده می‌شود. الگوریتم‌های جنگل تصادفی (Random Forest) توسط بریمن (۲۰۰۱) پیشنهاد شد و امروزه محبوبیت زیادی پیدا کرده است [۳۴]. مدل جنگل تصادفی یک تکنیک یادگیری گروهی است که برای آن تعداد زیادی درخت تصمیم برای توضیح روابط فضایی بین وقوع سیل و عوامل مرتبط برای طبقه‌بندی و رگرسیون ساخته می‌شود. با ساخت تعداد زیادی درخت تصمیم در زمان آموزش و خروجی کلاس که حالت کلاس‌ها (طبقه‌بندی) یا میانگین پیش‌بینی (رگرسیون) درخت‌های منفرد است، عمل می‌کند. جنگل‌های تصمیم‌گیری تصادفی، درخت‌های تصمیم‌گیری زیادی برای مجموعه آموزشی درست می‌کنند. الگوریتم جنگل تصادفی یک روش یادگیری گروهی است که مجموعه‌ای از پیش‌بینی‌کننده‌های درختی ساده را برای طبقه‌بندی و رگرسیون تشکیل

می‌دهد [۳۴]. اولین الگوریتم برای جنگل‌های تصمیم‌گیری تصادفی با استفاده از روش زیرفضای تصادفی ایجاد شد [۳۵]. از فرمول هو، پیاده‌سازی بر اساس رویکرد تصادفی-تبعیض طبقه‌بندی به دست می‌آید [۳۶]. توسعه الگوریتم با یک ایده بسته‌بندی و انتخاب تصادفی ویژگی‌ها برای ساخت مجموعه‌ای از سیستم‌های درختی تصمیم‌گیری با واریانس کنترل‌شده توسعه داده شد [۳۴] و این الگوریتم برای این مطالعه استفاده می‌شود. نتیجه‌ای با گروهی از مقادیر پیش‌بینی‌کننده برای هر دو الگوریتم طبقه‌بندی و رگرسیون تولید می‌شود. در این مطالعه، آسیب‌پذیری سیل بر اساس مقادیر فاکتورهای داده‌شده طبقه‌بندی می‌شود. به طور کلی، مسئله طبقه‌بندی با هدف به دست آوردن مرز تصمیم است. به عکس، الگوریتم رگرسیون یک مقدار پیوسته برای مقدار خروجی یا هدف دارد. بنابراین، این مطالعه میزان آسیب‌پذیری را به صورت یک مقدار پیوسته با الگوریتم رگرسیون نشان می‌دهد [۳۷]. همان‌طور که گفته شد، ابتدا برای الگوریتم طبقه‌بندی، گروهی از مقادیر پیش‌بینی مستقل با یکی از گروه‌هایی که در متغیر وابسته وجود دارد، طبقه‌بندی می‌شود. به عنوان یک جایگزین، برای الگوریتم رگرسیون، درخت با تخمینی از متغیرهای وابسته با خروجی کلاس‌بندی‌شده که حالت پیش‌بینی میانگین است، به پیش‌بینی‌کننده‌ها پاسخ می‌دهد. تعداد درختان تصادفی ساده، پیش‌بینی‌کننده‌های نهایی یک الگوریتم جنگل تصادفی را تعیین می‌کنند. الگوریتم طبقه‌بندی، درخت‌های تصمیم را می‌سازد و کلاس‌بندی را انجام می‌دهد. پاسخ، محبوب‌ترین طبقه، توسط گروه درختان ساده تعیین می‌شود. برای الگوریتم رگرسیون، از میانگین نتایج برای به دست آوردن تخمینی از متغیر وابسته استفاده می‌شود. بهبود قابل توجهی در دقت پیش‌بینی می‌تواند از استفاده از مجموعه‌های درختی حاصل شود. یک جنگل تصادفی به هیچ فرضی در مورد روابط بین متغیرهای توضیحی و پاسخ نیاز ندارد و روش مناسبی برای تجزیه و تحلیل تعاملات سلسله‌مراتبی و غیرخطی در مجموعه داده‌های بزرگ است.

مجموعه‌ای از مقادیر پیش‌بینی‌کننده به طور مستقل انتخاب می‌شود و بر پاسخ هر درخت تأثیر می‌گذارد. همچنین، توزیع یکسانی را برای همه درختان جنگل نشان

داده برای برخی پیش‌بینی‌کننده‌ها از فرایند جنگل تصادفی ضروری نیست [۳۸].

### ۲.۲.۲. ارزیابی مدل

به منظور ارزیابی مدل تهیه‌شده از منحنی تشخیص عملکرد نسبی (ROC) استفاده شد. منحنی ROC یک نمایش گرافیکی از موازنه بین نرخ خطای منفی و مثبت برای هر مقدار احتمالی از برش‌ها است. سطح زیرمنحنی ROC (AUC) بیانگر مقدار پیش‌بینی سیستم از طریق توصیف توانایی آن در تخمین درست وقایع رخ داده (وقوع سیل) و عدم وقوع رخداد (عدم وقوع سیل) آن است. ایده‌آل‌ترین مدل بیشترین سطح زیرمنحنی را دارد و مقادیر سطح زیرمنحنی بین ۰/۵ تا ۱ متغیر است. هر چه سطح زیرمنحنی به یک نزدیک‌تر باشد، بیانگر دقت بیشتر نقشه تهیه‌شده است [۳۹]. گسترش سریع مناطق شهری مشهود خطرات سیل را افزایش می‌دهد. به این منظور ۳۱۷ نقطه آب‌گرفتنی از اداره آب منطقه‌ای شهر مشهد گرفته شده است که در شکل (I.۲) نشان داده شده است.

### ۳. یافته‌ها

سیلاب‌ها پدیده‌های طبیعی هستند که از جنبه‌های آب‌شناختی، زمین‌ریخت‌شناختی و بوم‌شناختی اهمیت دارند. سیلاب‌ها با ویژگی‌هایی همچون سرعت جریان بالا، مسیرهای جریانی مختلف و فرایندهای بسیار فعال فرسایشی، حمل و رسوب‌گذاری همراه هستند. به طور کلی، عوامل متعددی در یک حوضه آبخیز بر جریان آب در سطح حوضه اثرگذار هستند که هر یک از این عوامل توان متفاوتی در تغییر جریان و رخداد سیلاب دارند. تعیین مهم‌ترین مؤثرترین عوامل تأثیرگذار بر سیلاب برای انجام مدل‌سازی بسیار ضروری است. با توجه به نتایج به‌دست‌آمده در مطالعات پیشین، مقدار ارتفاع، شیب و فاصله از رودخانه با وقوع سیل رابطه عکس دارند، به طوری که هر چه ارتفاع، شیب و فاصله از رودخانه و مسیل‌ها بیشتر باشد، احتمال خطر و به طبع آن خسارت کاهش پیدا می‌کند. در این پژوهش از مدل معتبر RF، برای تخمین میزان خطرپذیری سیلاب استفاده شد (شکل ۳: A). نتایج به شکل نقشه حساسیت به سیل، برای مدل یادگیری ماشین تهیه شد. مقادیر حساسیت به سیل (شکل ۳: B) در این مدل طبق مطالعات گذشته به ۵ گروه حساسیت طبقه‌بندی شدند که

می‌دهد و زیرمجموعه‌ای از مقادیر پیش‌بینی‌کننده مجموعه داده اولیه است. شرایط بهینه زیرمجموعه متغیرهای پیش‌بینی‌کننده با  $\log_2(M + 1)$  به دست می‌آید، که در آن  $M$  تعداد ورودی‌های الگوریتم است و میانگین مربعات خطای یک جنگل تصادفی با معادله زیر به دست می‌آید:

$$\varepsilon = (V_{Observed} - V_{response})^2 \quad (1)$$

که در آن  $\varepsilon$  میانگین مربعات خطای الگوریتم است،  $v$  مشاهده شده متغیرهای داده‌های مشاهده شده است و  $v$  پاسخ متغیرهای حاصل است.

میانگین پیش‌بینی‌ها از درختان به صورت زیر محاسبه می‌شود:

$$S = \frac{1}{K} \sum K^{th} V_{response} \quad (2)$$

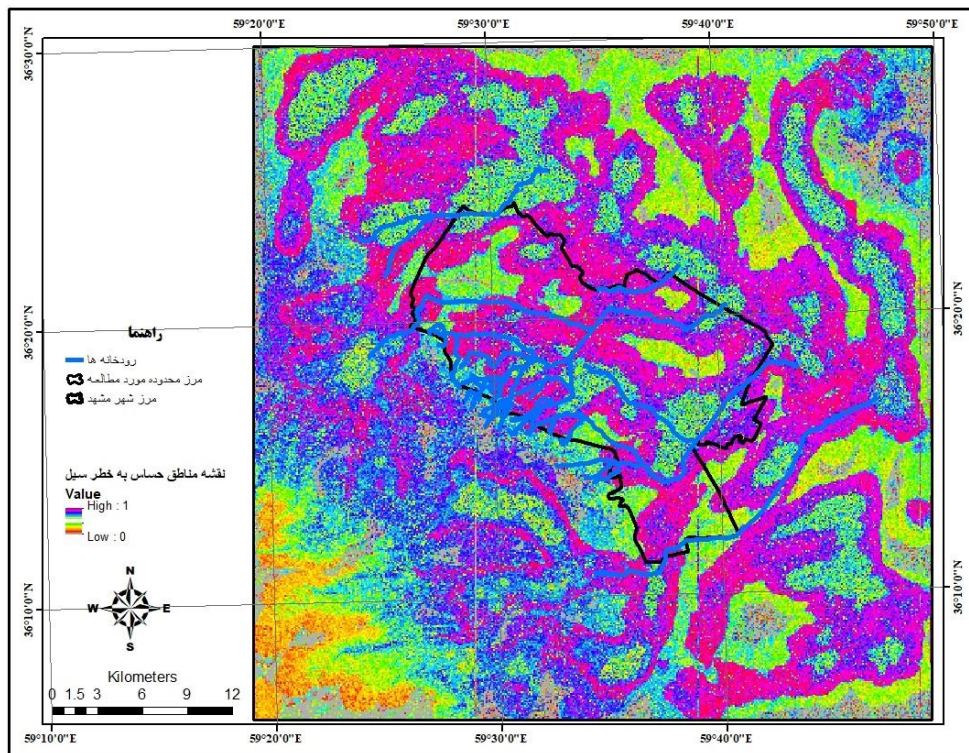
که در آن  $S$  پیش‌بینی جنگل تصادفی است و  $K$  برای تک تک درختان جنگل اعمال می‌شود.

برای الگوریتم طبقه‌بندی، تعریف تابع حاشیه‌ای که میانگین تعداد آرای اضافی را برای کلاس صحیح در مقایسه با میانگین رأی کلاس‌های دیگر در متغیر وابسته محاسبه می‌کند، پس از تعیین مجموعه‌ای از درختان ساده و متغیرهای پیش‌بینی‌کننده تصادفی انجام می‌شود. الگوریتم طبقه‌بندی یک روش عملی برای پیش‌بینی‌ها ارائه می‌کند و همچنین، می‌تواند روشی برای ارتباط یک معیار اطمینان با پیش‌بینی‌ها باشد.

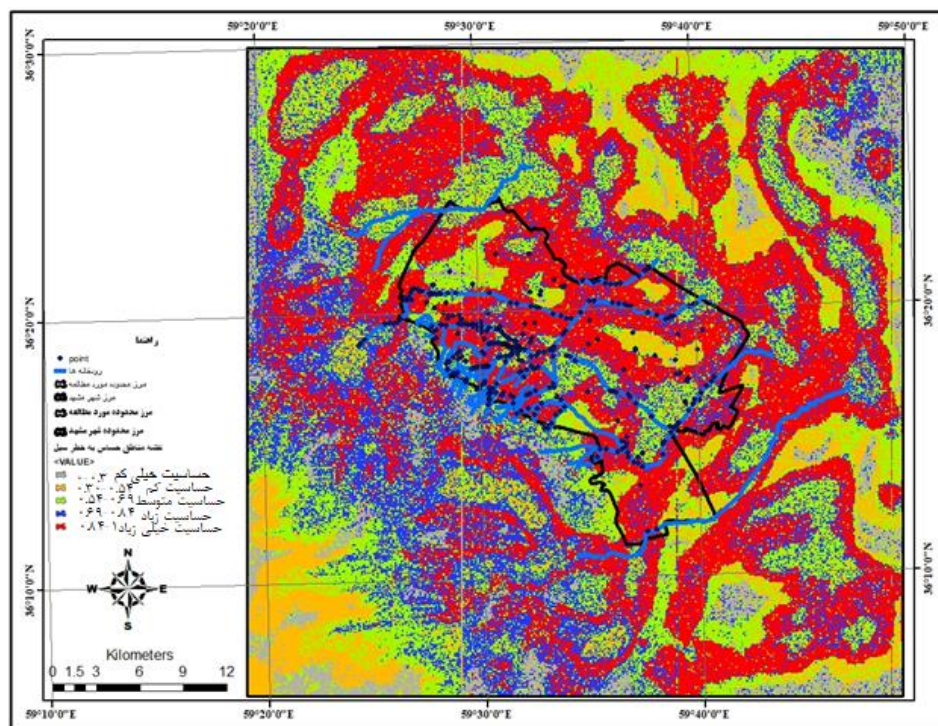
برای الگوریتم رگرسیون، از درختان ساده‌ای که قادر به تولید یک مقدار پاسخ عددی هستند برای ساخت جنگل‌های تصادفی استفاده می‌شود. مانند الگوریتم طبقه‌بندی، یک مجموعه پیش‌بینی به طور تصادفی برای همه درختان از توزیع برابر انتخاب می‌شود. معمولاً، داده‌های گمشده از متغیرهای پیش‌بینی‌کننده در جنگل تصادفی می‌توانند به طور انعطاف‌پذیری یکپارچه شوند. هنگامی که الگوریتم یک مدل می‌سازد، پیش‌بینی یک مورد خاص بر اساس آخرین گره قبلی (غیر پایانی) در درخت مربوطه زمانی که داده‌های از دست‌رفته گنجانده می‌شود، ایجاد می‌شود. بنابراین، میانگین کلی در گره ریشه برای استخراج متغیرهای پیش‌بینی زمانی استفاده می‌شود که هیچ داده معتبری در یک نقطه خاص در دنباله درختان وجود نخواهد داشت. بنابراین، حذف موارد بدون

هشداردهنده و برنامه‌های نظارتی باید در وهله اول در این قسمت‌های بسیار حساس به سیل در منطقه متمرکز شوند.

شامل حساسیت خیلی کم، کم، متوسط، زیاد و خیلی زیاد است. بر این اساس پیشنهاد می‌شود که سیستم‌های



شکل ۳: A: نقشه حساسیت به سیل شهر مشهد با مدل RF

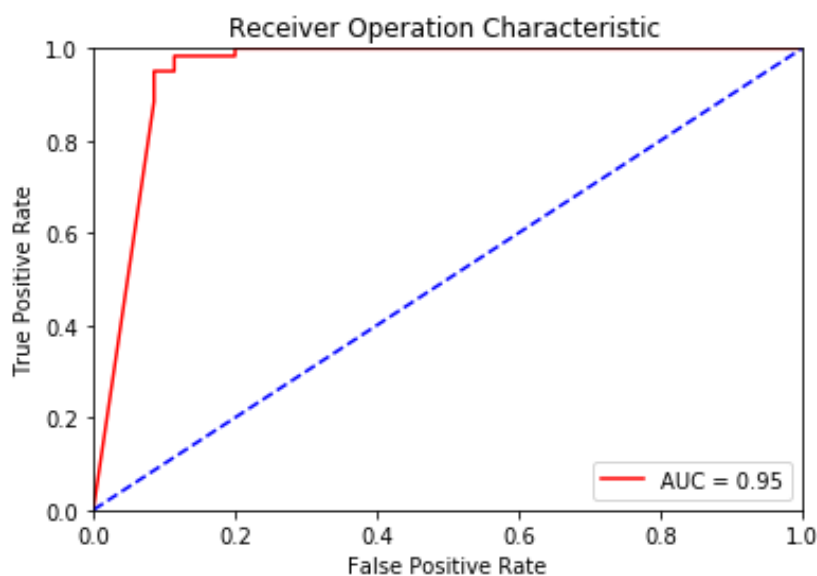


شکل ۳: B: نقشه طبقه‌بندی حساسیت به سیل شهر مشهد با مدل RF

## ۱.۳. اعتبارسنجی

اعتبارسنجی نقشه تهیه شده گامی اساسی در توسعه و تعیین مناطق حساس است. به منظور اعتبارسنجی مدل‌های مورد استفاده از منحنی ROC (شکل ۴) و مساحت زیرمنحنی (AUC) برای ارزیابی استفاده شده است. نتایج اعتبارسنجی مدل استفاده شده با استفاده از داده‌های آزمون نشان داد مدل RF، دارای سطح زیرمنحنی (۹۵ درصد) است. همچنین، نتایج اعتبارسنجی با استفاده از داده‌های آموزشی

برای مدل RF، نشان دهنده سطح زیرمنحنی (۹۵ درصد) بود. نتایج نشان داد مدل RF دارای دقت زیادی در تعیین مناطق حساس به وقوع سیل در شهر مشهد است. از نتایج این پژوهش می‌توان به منظور شناسایی مناطق مناسب و امن برای توسعه عمرانی استفاده کرد. با توجه به نتایج قابل قبول مدل جنگل تصادفی (RF) پیشنهاد می‌شود از این روش به منظور تهیه نقشه حساسیت و خطر سیل در مناطق مشابه دیگر نیز استفاده شود.



شکل ۴. منحنی ROC

## ۴. بحث و نتیجه‌گیری

شناسایی عوامل مؤثر بر وقوع سیل موجود در یک منطقه و پهنه‌بندی خطر آن یکی از عوامل اساسی برای دستیابی به راهکارهای کنترل این پدیده طبیعی است. ارزیابی خطر سیل یک فرایند پیچیده و چندمرحله‌ای است که توسط بسیاری از پژوهشگران مورد بررسی قرار گرفته و تا به امروز روش‌های مختلفی برای برآورد خطر پدیده سیل و عملکرد آن‌ها مطرح شده است. با وجود کاربرد جهانی فنون یادگیری ماشین برای پیش‌بینی سیل، تولید نقشه حساسیت به سیل قابل اعتماد هنوز یک امر چالش‌برانگیز است. مطالعه موردی این پژوهش در شهر مشهد انجام شد و ۱۱ عامل برای مدل سازی انتخاب شدند. داده‌های توپوگرافی یا مدل رقومی ارتفاع از تصاویر ماهواره‌ای با اندازه پیکسل ۱۰\*۱۰ انتخاب شدند. علاوه بر این، نقشه پوشش و کاربری اراضی که یکی از پارامترهای اساسی و

مؤثر بر سیلاب بود از تصاویر ماهواره Landsat 8 در آگوست ۲۰۱۸ تهیه شدند. در این مطالعه روش یادگیری ماشین RF اعمال شد و عملکرد آن در شهر مشهد مورد بررسی قرار گرفت. همچنین از عوامل ارتفاع، جهت، شیب، کاربری اراضی، زمین‌شناسی، تراکم آبراهه‌ها و فاصله از رودخانه و آبراهه‌ها، میزان بارندگی، شاخص خشکسالی (SPI) شاخص رطوبت توپوگرافی (TWI)، در این فرایند استفاده شد. نتایج نشان می‌دهد مدل RF برای منطقه مطالعاتی و براساس داده‌های به‌کاررفته دقت مطلوبی دارد و عملکرد و اعتبارسنجی این یافته با حساسیت سیل گذشته مطابقت داشته است. الگوریتم جنگل تصادفی یک الگوریتم بسیار قوی و ترکیبی است که از سیستم درختی تصمیم‌گیری برای الگوریتم‌های خود استفاده می‌کند. در این الگوریتم مجموعه‌ای از درخت‌های تصمیم با هم یک جنگل را تولید می‌کنند و این جنگل می‌تواند تصمیم‌های

- [4]. Khan A, Xun W, Ahsan H, Vineis P. Climate change, sea-level rise, & health impacts in Bangladesh. *Environ. Sci. Policy Sustain. Dev.* 2011; 53 (5): 18–33. <https://doi.org/10.1080/00139157.2011.604008>. Retrieved from.
- [5]. Adeoye N O, Ayanlade A, Babatimehin O. Climate change and menace of floods in Nigerian cities: socio-economic implications. *Adv. Nat. Appl. Sci.* 2009; 3 (3): 369–378.
- [6]. Berghuijs W R, et al. Recent changes in extreme floods across multiple continents. *Environmental Research Letters.* 2017; 12: 114035.
- [7]. Zhang Q. et al. Identifying dominant factors of waterlogging events in metropolitan coastal cities: The case study of Guangzhou, China. *Journal of Environmental Management.* 2020; 271: 110951.
- [8]. Lin J. Wu W. Investigating the land use characteristics of urban integration based on remote sensing data: experience from Guangzhou and Foshan. *Geocarto International.* 2019; 34: 1608-1620.
- [9]. Brody S D, et al. Evaluating the effects of open space configurations in reducing flood damage along the Gulf of Mexico coast. *Landscape and Urban Planning.* 2017; 167: 225-231.
- [10]. Mohammad Eini, Hesam Seyed Kaboli, Mohsen Rashidian, Hossien Hedayat. Hazard and vulnerability in urban flood risk mapping: Machine learning techniques and considering the role of urban districts. *International Journal of Disaster Risk Reduction.* 2020; 50: 101687.
- [11]. Mohammad Khalid Hossain, Qingmin Meng. A fine-scale spatial analytics of the assessment and mapping of buildings and population at different risk levels of urban flood. *Land Use Policy.* 2020; 99: 104829.
- [12]. Chengguang Lai, Quanxi shao, Xiaohong chen, Zhaoli Wang, Xiaowen Zhou, Bing Yang, Lilan Zhang. Flood risk zoning using a rule mining based on ant colony algorithm. *Journal of Hydrology.* 2016; 542: 268-280.
- [13]. Hossein Mojaddadi, Biswajeet Pradhan, Haleh Nampak, Noordin Ahmad & Abdul Halim bin Ghazali. Ensemble machine-learning-based geospatial approach for flood risk assessment using multi-sensor remote-sensing data and GIS. *Geomatics, Natural Hazards and Risk.* 2017. To link to this article: <http://dx.doi.org/10.1080/19475705.2017.1294113>.

بهتری را اتخاذ کند. در الگوریتم جنگل تصادفی مجموعه‌ای از درخت‌های تصمیم وجود دارند که به هر یک از آن‌ها یک زیرمجموعه از داده‌ها تزریق می‌شود و هنگام پیش‌بینی هر یک از این الگوریتم‌های یادگرفته‌شده یک نتیجه را پیش‌بینی می‌کنند و در نهایت با استفاده از رأی‌گیری، آن طبقه‌ای که بیشترین رأی را آورده است انتخاب می‌شود و به عنوان خروجی نهایی معرفی می‌شود. اوند و همکاران (۱۳۹۹)، توانایی مدل RF را در تحقیق خود با عنوان تهیه نقشه حساسیت سیل در حوضه آبریز تجن با مقدار AUC ۹۱ درصد نشان دادند [۴۰]. نتایج پژوهش حاضر با یافته‌های قبلی پشتیبانی می‌شود که نشان‌دهنده قابلیت زیاد و اطمینان زیاد روش‌های یادگیری ماشین است. نتایج این تحقیق نشان داد طبقات حساسیت از وقوع سیل در منطقه مورد مطالعه با آن بخش‌هایی از منطقه که فعالیت انسانی مانند مناطق مسکونی همراه است تطابق دارد و به این ترتیب نشان‌دهنده نیاز به طراحی سیستم‌های سیل‌بند برای جلوگیری از اثر سیل و ذخیره آب سیلاب در این منطقه برای کاهش خسارت‌های احتمالی است.

از نتایج پژوهش حاضر می‌توان به منظور شناسایی مناطق مناسب برای توسعه‌های عمرانی استفاده کرد. با توجه به نتایج قابل قبول که توسط مدل جنگل تصادفی ارائه شد، پیشنهاد می‌شود از این روش به منظور تهیه نقشه حساسیت و خطر سیل‌گیری در مناطق مشابه دیگر نیز استفاده شود.

## منابع

- [1]. Vitalea C, Meijerinka S, Mocchiab, F D, Ache P. Urban flood resilience, a discursive-institutional analysis of planning practices in the Metropolitan City of Milan. *Land Use Policy.* 2020; (95): 104575.
- [2]. Lin J, He X, Lu S, Liu D, He P. Investigating the influence of three-dimensional building configuration on urban pluvial flooding using random forest algorithm, *Environmental Research.* 2020. <https://doi.org/10.1016/j.envres.2020.110438>.
- [3]. UNDP. The Human Cost of Weather Related Disasters. 2015. Retrieved from. [http://www.unisdr.org/files/46796\\_cop21weatherdisastersreport2015.pdf](http://www.unisdr.org/files/46796_cop21weatherdisastersreport2015.pdf).



- [14]. Youssef A M, Pradhan B, Hassan A M. Flash flood risk estimation along the St. Katherine road, southern Sinai, Egypt using GIS based morphometry and satellite imagery. *Environ. Earth Sci.* 2011; 62 (3): 611–623. <https://doi.org/10.1007/s12665-010-0551-1>.
- [15]. Darabi H, Choubin B, Rahmati O, Torabi Haghighi A, Pradhan B, Kløve B. Urban flood risk mapping using the GARP and QUEST models: a comparative study of machine learning techniques. *J. Hydrol.* 2019; 142–154. <https://doi.org/10.1016/j.jhydrol.2018.12.002>
- [16]. Fernández D S, Lutz M A. Urban flood hazard zoning in Tucumán Province, Argentina, using GIS and multicriteria decision analysis. *Eng. Geol.* 2010; 111 (1–4), 90–98. <https://doi.org/10.1016/j.enggeo.2009.12.006>.
- [17]. Lee MJ, Kang J, Jeon S. Application of frequency ratio model and validation for predictive flooded area susceptibility mapping using GIS. 2012 IEEE International Geoscience and Remote Sensing Symposium. 2012. doi:10.1109/igarss.2012; 6351414.
- [18]. Pradhan B. Flood susceptible mapping and risk area delineation using logistic regression, GIS and remote sensing. *J Spatial Hydrol.* 2010; 9:1–18.
- [19]. Chau KW, Wu CL, Li YS. Comparison of several flood forecasting models in Yangtze River. *J Hydrol Eng.* 10: 485–491. doi:10.1061/(asce)1084-0699.2005; 10:6(485).
- [20]. Zhou J, Deng W, Zou Q, Xiao J, Zhang Y, Hua W. Flood disaster evaluation model based on kernel dual optimization support vector machine. *Information Technol J.* 2013; 12: 2412–2418. doi:10.3923/itj.2013.2412.2418.
- [21]. Tehrany MS, Pradhan B, Jebur MN. Flood susceptibility mapping using a novel ensemble weights-of-evidence and support vector machine models in GIS. *J Hydrol.* 2014; 512: 332–343. Doi:10.1016/j.jhydrol.2014.03.008.
- [22]. Lecca G, Petitdidier M, Hluchy L, Ivanovic M, Kussul N, Ray N, Thieron V. Grid computing technology for hydrological applications. *J Hydrol.* 2011; 403:186–199.
- [23]. Gholamali Fard M, Fatehi M, Bidel H, Ghafouri B. Analysis of dependence of land use changes in Mashhad city with topographical parameters in GIS environment. *Environmental science and technology.* 2016. 19th period Number 5: 299-315. In Persian.
- [24]. USGS. 2019. The National Map. Retrieved August 17. from. <https://viewer.nationalmap.gov/basic/>.
- [25]. Tehrany M S, Pradhan, B, Jebur, M N. Spatial prediction of flood susceptible areas using rule based decision tree (DT) and a novel ensemble bivariate and multivariate statistical models in GIS. *J. Hydrol.* 2013. (Amst) 504: 69–79. <https://doi.org/10.1016/j.jhydrol.2013.09.034>.
- [26]. Mukerji A, Chatterjee C, Raghuvanshi, N S. Flood forecasting using ANN, Neuro-Fuzzy, and Neuro-GA models. *J. Hydrol. Eng.* 2009; 14 (6): 647–652. [https://doi.org/10.1061/\(ASCE\)HE.1943-5584.0000040](https://doi.org/10.1061/(ASCE)HE.1943-5584.0000040).
- [27]. Pradhan B, Lee S. Landslide susceptibility assessment and factor effect analysis: backpropagation artificial neural networks and their comparison with frequency ratio and bivariate logistic regression modelling. *Environ. Model. Softw.* 2010; 25 (6): 747–759. <https://doi.org/10.1016/j.envsoft.2009.10.016>.
- [28]. Khosravi K, Nohani E, Maroufinia E, Pourghasemi H R. A GIS-based flood susceptibility assessment and its mapping in Iran: a comparison between frequency ratio and weights-of-evidence bivariate statistical models with multi-criteria decision- making technique. *Nat. Hazards.* 2016; 83 (2), 947–987. <https://doi.org/10.1007/s11069-016-2357-2>.
- [29]. Chu H, Wu W, Wang Q J, Nathan R. & Wei J. An ANN-based emulation modelling framework for flood inundation modelling: Application, challenges and future directions. *Environmental Modelling & Software.* 2020; 124: 104587.
- [30]. Gokceoglu C, Sonmez H, Nefeslioglu H A, Duman TY, Can T. The 17 March 2005 Kuzulu landslide (Sivas, Turkey) and landslide-susceptibility map of its near vicinity. *Eng Geol.* 2005; 81:65–83. Doi:10.1016/j.enggeo.2005.07.011.
- [31]. Rizeei HM, Saharkhiz M A, Pradhan B, Ahmad N. Soil erosion prediction based on land cover dynamics at the Semenyih watershed in Malaysia using LTM and USLE models. *Geocarto Int.* 2016; 31:1158–1177. Doi: 10.1080/10106049.2015.1120354.
- [32]. Rostami Khalaj M. Hasami D. Salmani H, Teymourian T. Urban flood risk zoning using multi-variable decision-making method (case study: Imam Ali town, Mashhad city). *Environmental science and technology.* The twenty-second period. 2018; Number (11). February 1998. In Persian.
- [33]. Netra R. Regmi John R. Giardino John D. Vitek. Assessing susceptibility to landslides: Using models to understand observed changes in

- slopes. *Geomorphology*. 2010. Volume 122, Issues 1–2, 1 October 2010, Pages 25-38.
- [34]. Breiman L. Random forests. *Mach learn*. 2001; 45:5–32.
- [35]. Ho TK. Random decision forests. *Proceedings of 3rd International Conference on Document Analysis and Recognition*. New York (NY): IEEE. 1995.
- [36]. Kleinberg EM. On the algorithmic implementation of stochastic discrimination. *IEEE Trans Pattern Anal Mach Int*. 2000; 22:473–490.
- [37]. Ellis K, Kerr J, Godbole S, Lanckriet G, Wing D, Marshall S. A random forest classifier for the prediction of energy expenditure and type of physical activity from wrist and hip accelerometers. *Physiol Meas*. 2014; 35:2191.
- [38]. Sunmin Lee, Jeong-Cheol Kim, Hyung-Sup Jung, Moungh Jin Lee & Saro Lee. Spatial prediction of flood susceptibility using random-forest and boostedtree models in Seoul metropolitan city, Korea, *Geomatics, Natural Hazards and Risk*. 2017. DOI: 10.1080/19475705.2017.1308971
- [39]. Pradhan B, Hagemann U, Tehrany, MS, Prechtel N. An easy to use ArcMap based texture analysis program for extraction of flooded areas from TerraSAR-X satellite image. *Comput Geosci*. 2014; 63:34–43. doi:10.1016/j.cageo.2013.10.011.
- [40]. Avand M T, Moradi H R, Ramzanzadeh Lesboi m. Preparation of flood sensitivity map using two random forest machine learning models and Bayesian generalized linear model. *Environment and water engineering*. 2019; Volume 6. Number 1. 83-95. In Persian.