

بررسی توان تفکیک و دقت مدل‌های رقومی ارتفاعی در پهنه‌بندی مناطق سیل‌خیز به روش یادگیری ماشینی

پژوهشگر: حسین باقری
کد پروژه:

چکیده

خطاهای انسانی این مطالعه، چارچوبی برای پهنه‌بندی نقاط سیل‌خیز و تهیه نقشه خطر احتمال سیلاب با استفاده از الگوریتم‌های یادگیری ماشین در بالادست رودخانه زاینده‌رود ارائه می‌دهد. در این مطالعه هدف اصلی بررسی تأثیر کیفیت مدل رقومی ارتفاعی در پهنه‌بندی سیلاب است. برای این منظور از DEM ۱۰ متری (حاصل از نقشه‌های توپوگرافی پایه کشوری) و SRTM ۳۰ متری استفاده گردید. در ابتدا پارامترهای موثر بر پهنه‌بندی سیلاب از مدل‌های رقومی مذکور استخراج گردید و پس از افزودن سایر پارامترهای موثر نظیر کاربری زمین، سنگ‌شناسی، خاک و غیره، از مدل‌های یادگیری ماشین شامل مدل جنگل تصادفی، مدل ماشین بردار پشتیبان و رگرسیون لجستیک بمنظور طبقه‌بندی و پهنه‌بندی استفاده گردید. نتایج نشان که افزایش توان تفکیک مکانی و دقت ارتفاعی مدل‌های رقومی ارتفاعی به نتایج دقیق‌تری می‌انجامد. بهترین نتایج با استفاده از دو مدل جنگل تصادفی و ماشین بردار پشتیبان حاصل گردید. این مدل‌ها، با دقت مشابه، عملیات طبقه‌بندی نقاط سیل‌خیز و غیر سیل‌خیز را انجام دادند. از نظر شاخص‌های متعدد، عملکرد هر دو مدل در این فرآیند طبقه‌بندی بهتر از ۹۸٪ گردید در حالی که در صورت استفاده از DEM ۳۰ متر دقت طبقه‌بندی از نظر شاخص‌ها کاهش یافت. نتایج بخوبی نشان داد که توان تفکیک مدل رقومی ارتفاع تأثیر بالایی در بهبود عملکرد یا کاهش دقت طبقه‌بندی دارد. همچنین ارزیابی اهمیت ویژگی‌های مختلف برای هر دو داده مستخرج از DEM ۱۰ متری و ۳۰ متری نشان داد که در صورت افزایش توان تفکیک، نقش پارامترهای مشتق شده از ارتفاعات (DEM) افزایش می‌یابد، در حالی که در DEM با رزولوشن کمتر، از اهمیت برخی پارامترهای مرتبط با توپوگرافی کاسته می‌شود. در نهایت پس از آموزش مدل‌های یادگیری ماشین و ارزیابی آن‌ها و انتخاب بهترین مدل، نقشه احتمال خطر سیلاب برای هر دو DEM ۱۰ متر و ۳۰ متر محاسبه گردید. ارزیابی کیفی نشان داد که DEM ۱۰ متر به نقشه دقیق‌تری جهت تعیین احتمال خطر منجر می‌شود و نسبت به نقشه حاصل از DEM ۳۰ متر با واقعیت همخوانی بیشتری دارد.

کلیدواژه: پهنه‌بندی سیلاب، مدل رقومی ارتفاع، یادگیری ماشینی، احتمال سیلاب

مقدمه

امروزه، سیلاب به عنوان یکی از مهم ترین پدیده های طبیعی مخرب برای انسان و محیط زیست شناخته شده است. سیلاب ها معمولاً به دلیل بارش شدید و یا ذوب برف در زمستان ایجاد می شوند. این پدیده باعث بروز خسارات جدی به زیرساخت ها و همچنین محیط زیست می گردد. با توجه به خسارت های شدید احتمالی سیلاب، ایجاد برنامه های پیشگیری و مدیریت مناسب برای کاهش خطرات سیلاب، امری بسیار مهم است. پهنه بندی سیلاب، یکی از راهکارهای مؤثر برای شناسایی مناطق با میزان خطر سیلاب متفاوت است. پهنه بندی سیلاب به معنای تقسیم نواحی مختلف به دسته های گوناگون بر اساس خطر سیلاب و میزان خسارتی که احتمال وقوع سیلاب برای آن ها دارد، می باشد [۱].

تاکنون روش های مختلفی برای پهنه بندی سیلاب توسعه داده شده است که عبارت اند از روش های آماری، روش های مدل سازی هیدرولوژیکی، روش های مبتنی بر یادگیری ماشینی و روش های ترکیبی. با استفاده از یادگیری ماشین، می توان ویژگی های مختلفی را که در تعیین پهنه های سیلابی مؤثر هستند، شناسایی کرد و با استفاده از آن ها، پهنه های سیلابی را با دقت بالاتری تعیین کرد. برای پهنه بندی سیلاب با استفاده از یادگیری ماشین، داده های هیدرولوژیکی و داده های مشاهداتی به عنوان ورودی استفاده می شوند و پهنه های سیلابی به صورت خودکار با استفاده از الگوریتم های یادگیری ماشین تعیین می شوند.

با توجه به نیاز به داده های کافی در روش های یادگیری ماشینی، استفاده از داده ها و محصولات حاصل از مشاهدات ماهواره ای، می تواند به صورت دقیق تری بارش، توپوگرافی و پوشش گیاهی، کاربری زمین و سایر پارامترهای مهم مرتبط با سیل را برآورد کرد. یکی از ورودی های مؤثر در پهنه بندی سیلاب مدل های رقومی ارتفاع است. مدل های رقومی ارتفاع بیانگر توپوگرافی منطقه هستند و در مدل سازی سیلاب نقش محوری دارند. با استفاده از داده های ماهواره ای و سنجش از دور، می توان به صورت دقیق تری توپوگرافی منطقه را برآورد کرد و از آن برای تعیین مسیر جریان آب و برآورد جریان سطحی و جریان رودخانه استفاده کرد.

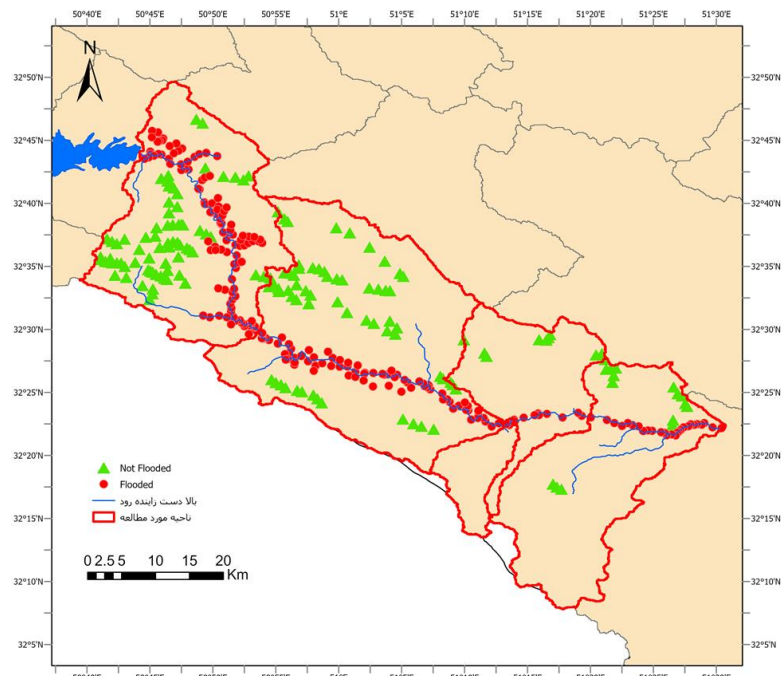
یکی از اهداف اصلی این پژوهش، بررسی قابلیت روش های یادگیری ماشین در پهنه بندی سیلاب است. هدف دیگر این مطالعه بررسی توان تفکیک و دقت ارتفاعی مدل های رقومی مختلف در تخمین و مدل سازی سیلاب با استفاده از الگوریتم های مختلف یادگیری ماشینی است. برای این منظور پارامترهای مختلف مؤثر در تخمین سیلاب نظیر ارتفاعات، میزان بارش، پوشش گیاهی، کاربری زمین و غیره به عنوان پارامتر ورودی در نظر گرفته می شوند و با استفاده از روش های یادگیری ماشین پهنه بندی سیلاب صورت می گیرد. برای ارزیابی اثر دقت مدل های رقومی ارتفاع در پهنه بندی سیلاب، مدل های مختلف ارتفاعی به دست آمده از داده ها و مشاهدات ماهواره ای به روش های مختلف نظیر مدل های رقومی حاصل از اینترفرومتری نظیر SRTM، مدل های رقومی حاصل از فتوگرامتری نظیر مدل رقومی ۱۰ متری تولید شده توسط سازمان نقشه برداری مورداستفاده قرار خواهد گرفت. همچنین در نهایت، اثر پارامترهای مختلف استخراج شده از مدل های رقومی زمین در پهنه بندی سیلاب مورد بررسی و ارزیابی قرار خواهد گرفت.

مواد و روش

منطقه مطالعاتی

حوضه آبریز زاینده رود واقع در منطقه مرکزی ایران (شکل ۱) حوضه کاملاً بسته ای است که هیچ راه خروجی به دریا ندارد. رودخانه زاینده به طول ۳۵۰ کیلومتر، در امتداد تقریبی غرب به شرق در جریان است و از کوه های زاگرس استان چهارمحال بختیاری سرچشمه گرفته و به باتلاق گاوخونی در شرق استان اصفهان می رسد. این رودخانه آب آبیاری، شرب و صنعت استان را که یکی از مهم ترین مناطق اقتصادی ایران است، تأمین می کند. مساحت کل حوضه زاینده رود حدود ۴۱۵۰۰ کیلومتر مربع است. البته فقط اراضی بالادست دریاچه سد زاینده رود، در تأمین جریان های آبی مؤثر بوده و محدوده

پایین دست سد، جریان ورودی به رودخانه نداشته و اگر هم جریانی بوده بسیار نادر و کم است. جریان طبیعی رودخانه زاینده رود با انحراف آب از تونل های انحرافی یک و دو کوهرنگ که از رودخانه کوهرنگ در استان چهارمحال بختیاری سرچشمه می گیرد، افزایش می یابد. در این پژوهش ارزیابی خطرپذیری سیلاب در حوزه بالادست رودخانه زاینده رود صورت گرفته است.



شکل ۱. نمایی از محدوده منطقه مورد مطالعه به همراه موقعیت قرار گیری نقاط سیل خیز و نقاط غیر سیل خیز نمونه برای آموزش مدل های یادگیری ماشین بمنظور تعیین خطر سیلاب

داده ها

سیلاب یکی از پدیده های طبیعی پیچیده ای است که می تواند واکنش های هیدروژئومورفولوژیک مختلفی را در مناطق مختلف ایجاد کند؛ بنابراین، برای درک اثرات این پدیده، ضروری است که عوامل مؤثر بر آن را به خوبی شناخت. باتوجه به ویژگی های هیدروژئومورفولوژیک حوزه زاینده رود و پیشنهادهای مطالعات قبلی [۲]، ۱۴ عامل که به طور قابل توجهی بر میزان سیلاب تأثیر می گذارند، انتخاب شدند. این عوامل شامل شیب، ارتفاع، انحنا، فاصله از رودخانه، تراکم زهکشی، کاربری اراضی، سنگ شناسی، جهت شیب، میزان بارش، طول شیب، شاخص توپوگرافی رطوبت (TWI)، شاخص موقعیت توپوگرافی (TPI)، خاک، و شاخص نرمال شده پوشش گیاهی (NDVI) هستند. نقشه پایه برای استخراج برخی از این عوامل مانند شیب، جهت شیب، ارتفاع، انحنا، رودخانه، TWI و TPI داده های ماهواره ای مرتبط با مدل رقومی زمین (DEM) است.

یکی از مدل های رقومی ارتفاعی، مدل رقومی ارتفاع ده متری است که توسط سازمان نقشه برداری کشور از روی نقشه های توپوگرافی و با استفاده از عکس برداری هوایی تهیه شده است، این مدل برای منطقه زاینده رود موجود و در دسترس است. با انتخاب مدل رقومی ۱۰ متری فراهم شده توسط سازمان نقشه برداری، برای مقایسه، مدل SRTM انتخاب شد که یک مدل راداری به دست آمده از طریق اینترفرومتری با توان تفکیک ۳۰ متر است که نسخه ۳ آن به دلیل تلفیق با ASTER دارای کیفیت و دقت بالاتر می باشد. بررسی ها نشان داد که مدل SRTM نسخه ۳ به طور کلی دقت بالاتری نسبت به مدل های رقومی دیگر دارد. در نتیجه استفاده از مدل رقومی ۱۰ متری حاصل از تصاویر هوایی اپتیک به روش فتوگرامتری و SRTM حاصل از اینترفرومتری و بررسی اثر دقت و توان تفکیک این مدل ها در پهنه بندی سیلاب جذاب تر و چالشی تر از سایر مدل های رقومی ارتفاع می باشد.

این انتخاب ها بر اساس سه شاخص، ماهیت منبع داده مورد استفاده در تولید مدل رقومی، توان تفکیک مکانی، و در دسترسی بودن مدل صورت گرفت. جدول ۴-۱، داده‌های مورد استفاده در این مطالعه را لیست کرده است.

جدول ۱. لیست داده‌های مورد استفاده در این پژوهش

داده	فرمت داده	توان تفکیک مکانی	منبع	پارامترهای استخراج شده
DEM	رستر	۱۰ متر	سازمان نقشه‌برداری کشور	شیب، جهت، انحنا، ارتفاع، فاصله از رودخانه، TPI, TWI, SPI
		۳۰ متر	NASA: Google Earth Engine	
کاربری زمین	وکتور	۱:۱۰۰۰۰۰	شرکت آب منطقه‌ای اصفهان	کاربری زمین
بارش	رستر	۱۱ کیلومتر	NASA: Google Earth Engine	بارش
سنگ‌شناسی	وکتور	۱:۲۵۰۰۰	سازمان زمین‌شناسی و اکتشافات معدنی ایران	سنگ‌شناسی
خاک‌شناسی	وکتور	۱:۱۰۰۰۰۰	شرکت آب منطقه‌ای اصفهان	نوع خاک
Landsat8	رستر	۳۰ متر	NASA: Google Earth Engine	NDVI

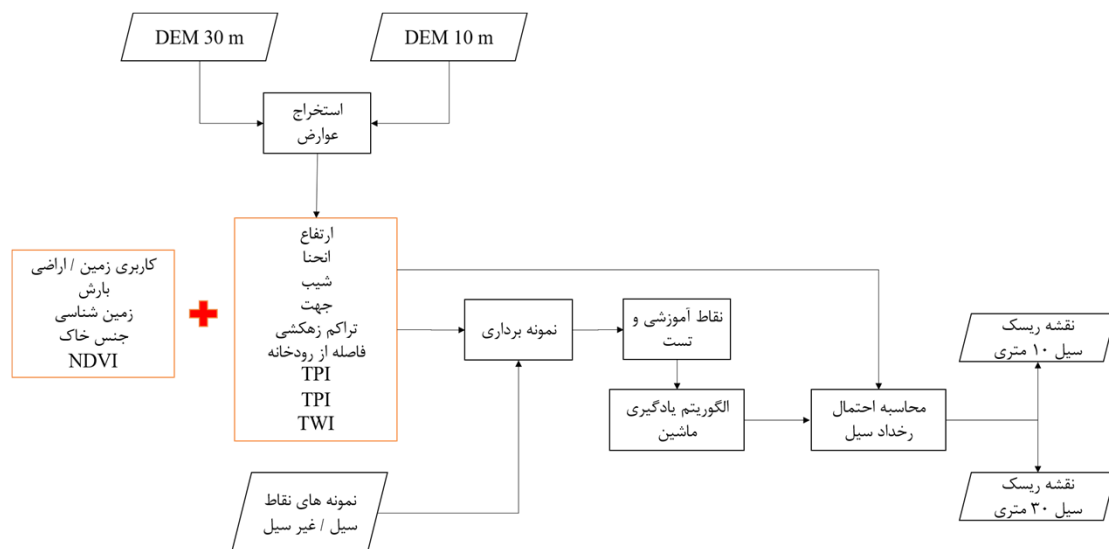
روش و مراحل انجام پژوهش

شکل ۲ فرایند تهیه نقشه خطرپذیری سیلاب را در منطقه مورد مطالعه به تصویر می‌کشد. به‌طور کلی مطابق با فلوچارت فرایند پهنه‌بندی خطر سیلاب را می‌توان در گام‌های ذیل خلاصه نمود:

- تعیین نقاط منفرد سیل خیز و غیر سیل خیز با استفاده از اطلاعات ماهواره‌ای و شاخص‌های مرتبط
- استخراج شاخص‌های مؤثر در تخمین سیلاب نظیر کاربری زمین، NDVI، بارش از تصاویر ماهواره‌ای و داده‌های ماهواره‌ای در محیط Google Earth Engine
- تهیه و آماده‌سازی مدل‌های رقومی مختلف بادقت و رزولوشن‌های مختلف نظیر SRTM با توان تفکیک ۳۰ متر و مدل رقومی سطح ۱۰ متری فراهم شده توسط سازمان نقشه‌برداری کل کشور
- استخراج پارامترهای مرتبط با مشخصات فیزیکی حوزه آبریز از مدل‌های رقومی زمین نظیر تراکم زهکشی، شیب و جهت شیب و انحنا و غیره.

- استخراج پارامترهای مختلف زمین‌شناسی و تیپ‌بندی خاک و کاربری زمین از لایه‌های خاک و سنگ‌شناسی و کاربری زمین

- استخراج پارامترهای مؤثر در تخمین سیلاب در نقاط نمونه سیل خیز و غیر سیل خیز و آماده‌سازی و پیش‌پردازش داده‌ها
- تعیین نقاط آموزش، اعتبارسنجی و تست
- پیاده‌سازی مدل‌های یادگیری ماشینی با استفاده از نمونه آموزشی و مقایسه و ارزیابی دقت آن‌ها با استفاده از داده‌های تست و انتخاب بهترین مدل با بالاترین دقت
- ایجاد مدل نهایی بر اساس پارامترهای بهینه
- بررسی تأثیر پارامترهای توپوگرافی، دقت و توان تفکیک مدل‌های رقومی ارتفاعی مختلف در تخمین سیلاب
- تهیه نقشه‌های پهنه‌بندی خطر سیلاب با استفاده از مدل آموزش دیده در حوزه آبریز



شکل ۲. چارچوب طراحی شده برای مدلسازی خطر سیلاب در منطقه مطالعات

استخراج پارامترهای مؤثر در سیلاب

در ابتدا پارامترهای مؤثر در تخمین سیلاب از هر دو DEM مورد مقایسه (۱۰ متری و ۳۰ متری)، استخراج گشتند. ۱۴ پارامتر شامل شیب، ارتفاع، انحنا، فاصله از رودخانه، تراکم زهکشی، کاربری اراضی، سنگ شناسی، جهت شیب، میزان بارش، طول شیب، شاخص توپوگرافی رطوبت (TWI)، شاخص موقعیت توپوگرافی (TPI)، شاخص قدرت جریان (SPI)، خاک و شاخص نرمال شده پوشش گیاهی (NDVI) بودند که در توان تفکیک های مکانی ۱۰ متر و ۳۰ متر تهیه گشتند. این پارامترها تاثیر قابل توجهی در رخداد سیلاب دارند که باعث کاهش یا افزایش ریسک سیل می گردند.

مدلسازی خطر سیلاب

در این پژوهش، به دلیل عدم دسترسی به تاریخچه سیلاب و عدم دسترسی به نقاط سیل خیز و غیر سیل خیز، از خروجی پهنه بندی مدل SWAT با دوره بازگشت ۱۰۰ ساله [۳] برای تعیین داده های آموزشی و تست استفاده گردید. در نهایت، ۲۲۹ مکان مستعد سیلاب و ۱۶۹ مکان غیر مستعد با استفاده از داده های موجود و نقشه ها انتخاب شدند. این مکان ها به طور تصادفی به دو بخش آزمایش (۳۰٪) و آموزش (۷۰٪) تقسیم شدند. ۱۶۰ مکان سیلاب به همراه ۱۱۸ مکان غیر مستعد به طور تصادفی برای آموزش مدل های یادگیری ماشین انتخاب شدند. برچسب ۱ وقوع سیلاب را نشان می دهد و مقدار ۰ نشان دهنده عدم وقوع سیلاب در منطقه است. در نتیجه دو دسته نمونه متشکل از پارامترهای پیش بینی کننده مستخرج از DEM ۱۰ متری و ۳۰ متری به همراه برچسب تهیه شد. شکل ۱ محل قرارگیری نقاط آموزشی و تست برای هر دو کلاس سیل خیز و غیر سیل خیز را نشان می دهد. از داده های ایجاد شده به منظور آموزش مدل های طبقه بندی یادگیری ماشین استفاده گردید. این مدل ها با استفاده از داده های آموزشی آموزش داده شدند و پس از تنظیم های پارامترها دقت طبقه بندی آن ها مورد ارزیابی قرار گرفت. با توجه به عملکرد مدل ها بر روی داده های تست در نظر گرفته شده، بهترین مدل طبقه بندی برای تهیه نقشه خطر احتمال سیلاب مشخص گردید. نقشه های احتمال سیلاب منطقه مورد مطالعه با استفاده از بهترین مدل برای دو نوع داده با توان تفکیک مکانی ۱۰ متر و ۳۰ متر تولید شد و با استفاده از نرم افزار ArcGIS در ۴ دسته دسته بندی شد. برای ایجاد قابلیت مقایسه و تولید نقشه ها با شرایط یکسان، دسته بندی نقشه های مرتبط با هر دو نوع داده با الگوریتم شکست طبیعی انجام شد که یکی از الگوریتم های دسته بندی نقشه های رستری است. در نهایت، بر اساس مطالعات قبلی و نوع نقشه های موجود، این نقشه ها به ۴ کلاس تقسیم شدند: بی خطر، خطر کم، متوسط و بالا.

یافته‌ها و نتایج تحقیق

جدول ۲ نتایج عملکرد مدل‌های مختلف یادگیری ماشین در طبقه بندی نقاط سیل خیز و غیر سیل خیز را برای پارامترهای مستخرج شده از DEM ۱۰متری نشان می‌دهد. همانگونه که در جدول ۲ نشان داده شده است الگوریتم‌های جنگل تصادفی و ماشین بردار پشتیبان عملکرد یکسانی داشته‌اند و بالاترین دقت را در طبقه بندی بدست داده‌اند. با توجه به توپوگرافی منطقه مورد مطالعه و کیفیت بالاتر داده‌های جمع شده (از DEM ۱۰ متر)، دقت بسیار بالایی توسط دو مدل حاصل شده است. مدل رگرسیون لجستیک به دلیل مبانی ساده تر نسبت به دو مدل قبلی، ظرفیت کمتری برای مدلسازی دارد. در نتیجه دقت عملکرد آن نسبت به مدل ماشین بردار و جنگل تصادفی کاهش می‌یابد. با این حال این مدل از نظر پردازشی و هزینه محاسباتی مقرون بصرفه ترین مدل می‌باشد.

جدول ۲ نتایج حاصل از ارزیابی دقت الگوریتم‌های مختلف یادگیری ماشین در طبقه بندی نقاط سیل خیز و غیر سیل خیز با استفاده از داده‌های مشتق شده از DEM ۱۰ متر

الگوریتم	Accuracy	Recall	F1 Score	Precision	AUC
جنگل تصادفی	۹۹/۲۲	۱۰۰	۹۹/۳۲	۹۸/۶۵	۹۹/۱۱
ماشین بردار پشتیبان	۹۹/۲۲	۱۰۰	۹۹/۳۲	۹۸/۶۵	۹۹/۱۱
رگرسیون لجستیک	۹۶/۹۰	۱۰۰	۹۷/۳۳	۹۴/۸۱	۹۶/۴۳

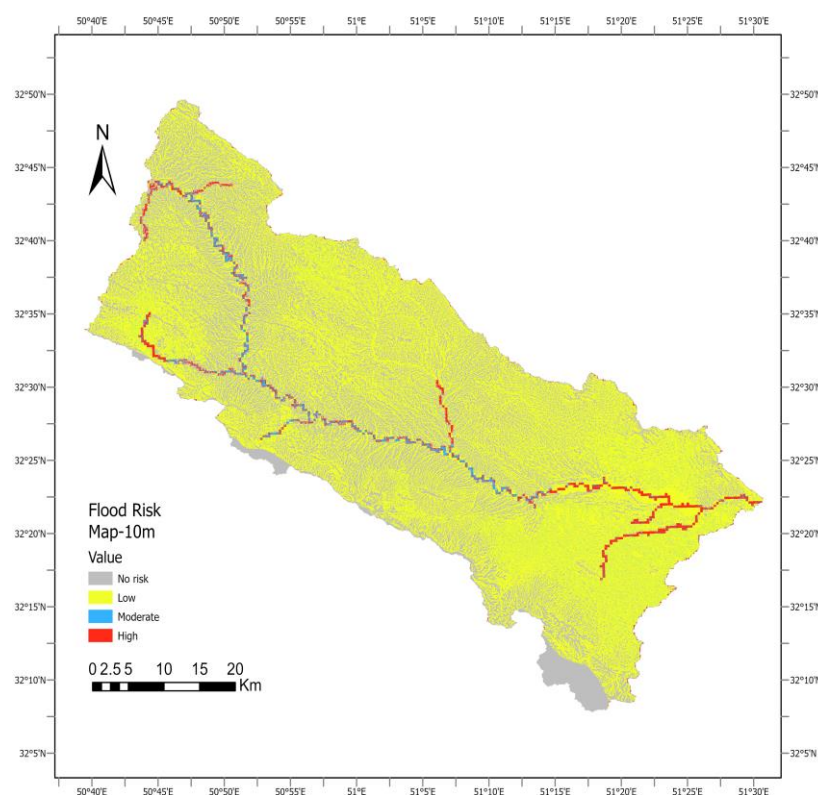
نتایج طبقه بندی نقاط سیل خیز و غیر سیل خیز با استفاده از داده‌های آموزشی استخراج شده از DEM ۳۰متر نیز در جدول ۳ برای هر سه مدل یادگیری ماشین نشان داده شده است. جدول ۳ بخوبی نشان می‌دهد که مدل ماشین بردار پشتیبان بالاترین دقت را بدست داده است. در رتبه بعدی مدل جنگل تصادفی قرار دارد اگرچه مدل رگرسیون لجستیک از نظر معیار Recall به دقت بالاتری (۹۷.۲۶) در مقابل جنگل تصادفی (۹۴.۵۲) دست یافته است.

جدول ۳ نتایج حاصل از ارزیابی دقت الگوریتم‌های مختلف یادگیری ماشین در طبقه بندی نقاط سیل خیز و غیر سیل خیز با استفاده از داده‌های مشتق شده از DEM ۳۰ متر

الگوریتم	Accuracy	Recall	F1 Score	Precision	AUC
جنگل تصادفی	۹۶/۹۰	۹۴/۵۲	۹۷/۱۸	۱۰۰	۹۷/۲۶
ماشین بردار پشتیبان	۹۸/۴۵	۹۷/۲۶	۹۸/۶۱	۱۰۰	۹۸/۶۳
رگرسیون لجستیک	۹۶/۱۲	۹۷/۲۶	۹۶/۶۰	۹۵/۹۵	۹۵/۹۵

پس از انتخاب آموزش و ارزیابی مدل‌های یادگیری ماشین و انتخاب بهترین مدل، نقشه احتمال سیلاب یا نقشه خطر سیلاب برای هر دو DEM ۱۰متری و ۳۰ متری در منطقه مورد مطالعه بدست آمد. شکل ۳ نقشه خطر سیلاب با توان تفکیک مکانی ۱۰ متری را برای بالادست زاینده رود با استفاده از DEM ۱۰متری نشان می‌دهد. همان‌طور که در نقشه ۳ نشان داده شده است نقاط با احتمال خطر بالا در حاشیه زاینده و زیرشاخه‌های آن قرار گرفته‌اند. در قسمت‌های بالادست اندکی از میزان خطر کاهش یافته و در بخش‌هایی از رودخانه احتمال خطر سیلاب، متوسط نمایش داده شده است. در این قسمت از رودخانه بدلیل وجود ناهمواری با ارتفاع زیاد و گذر رودخانه از میان این ناهمواری‌ها (به دلیل ارتفاع بالا و شیب زیاد) خطر سیلاب کاهش می‌یابد. در قسمت‌های

دشت و مناطق هموار میزان خطر کاهش یافته و به درجه کم رسیده است که به دلیل دوری از رودخانه خطر سیلاب بطور قابل توجهی کاهش می یابد اما به دلیل هموار بودن احتمال اندکی از رخداد سیلاب بود. در ارتفاعات بالا (رنگ خاکستری) به دلیل عدم دسترسی آب، بدرستی احتمال رخداد سیلاب نزدیک صفر تشخیص داده شده است و خطری در این نواحی وجود ندارد.



شکل ۳ نقشه خطر سیل حاصل از DEM ۱۰متر

متناظرا، نقشه احتمال رخداد سیل برای DEM ۳۰ متری نیز با استفاده از الگوریتم طبقه‌بند آموزش دیده محاسبه گردید که در شکل ۴ نمایش داده شده است. اگرچه در DEM ۳۰ متر نیز احتمال رخداد سیل در حاشیه رودخانه‌ها به درستی تشخیص داده شده است، اما بسیاری از ناهمواری‌ها که در آن احتمال دسترسی به آب به دلیل شدت توپوگرافی نیست، خطر احتمال سیلاب همچنان (ولو اندک) وجود دارد. به عبارت دیگر در مدل رقومی ۳۰ متری به دلیل رزولوشن و دقت ارتفاعی پایین‌تر، تخمین پارامترهای مؤثر در سیلاب با خطای بیشتری صورت می‌گیرد و نتیجه آن، کاهش دقت نقشه خطر احتمال سیلاب است. درحالی‌که نقشه حاصل از به کارگیری DEM ۱۰ متر با واقعیت منطقه همخوانی بیشتری دارد.

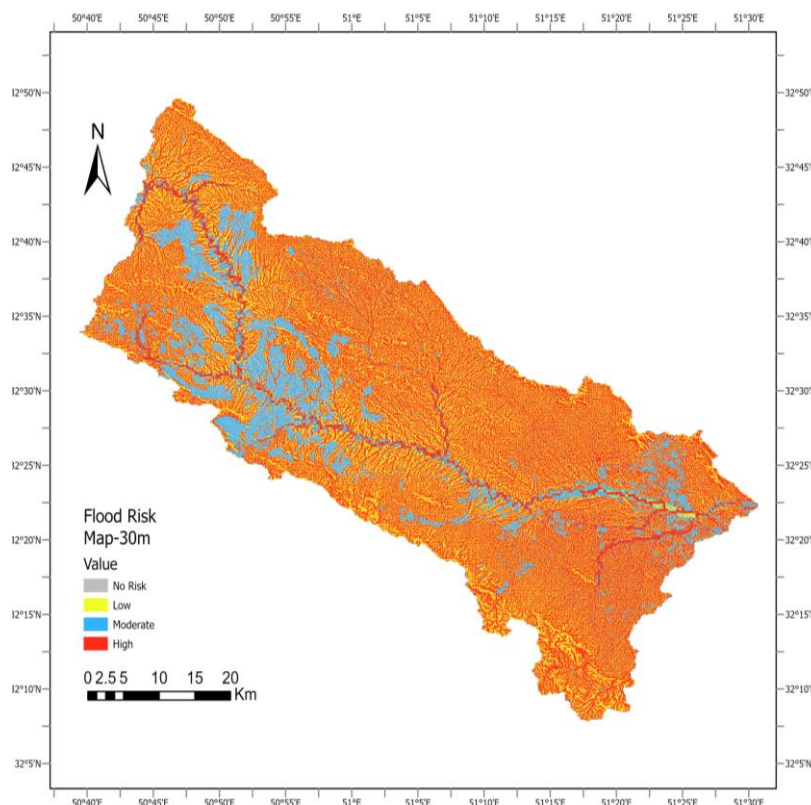
جمع‌بندی و نتیجه‌گیری

به‌طور کلی، نتایج به‌دست‌آمده مشخص کرد که توان تفکیک مکانی نقش مهمی در کیفیت و دقت پهنه‌بندی سیلاب با تکنیک‌های یادگیری ماشین دارد. ویژگی‌های متعددی از مدل‌های رقومی ارتفاع استخراج می‌شوند که با تغییر توان تفکیک مکانی، دقت استخراج این ویژگی‌ها تغییر خواهد کرد. به طور دقیق‌تر نتایج نشان داد که استفاده از DEM ۱۰متری به نتایج بهتری در طبقه‌بندی نقاط سیل‌خیز و غیر سیل‌خیز می‌انجامد. همچنین نقشه‌های خطر احتمال سیلاب حاصل از مدل رقومی ده‌متری کیفیت بالاتری نسبت به نقشه‌های حاصل از DEM ۳۰متری داشتند.

ذیل این تحقیق به پرسش‌های مختلفی پاسخ داده شد که به شرح ذیل است:

- الگوریتم‌های یادگیری ماشین بادقت بالایی توانایی طبقه‌بندی نقاط سیل‌خیز و غیر سیل‌خیز را دارند. این دقت برای مدل‌های جنگل تصادفی و ماشین بردار پشتیبان به بیش از ۹۸٪ در صورت استفاده از DEM ۱۰ متری می‌رسد.

- یکی از مزیت‌های رگرسیون لجستیک کم‌هزینه‌تر بودن، سریع‌تر اجرا شدن، تفسیرپذیری بالاتر و سادگی پیاده‌سازی آن است. با این حال، در صورتی که تعداد داده‌ها بسیار زیاد باشد، ممکن است دقت آن تحت تأثیر قرار بگیرد. در مقابل، مدل‌های جنگل تصادفی و مدل‌های ماشین بردار پشتیبان سناریوهای پیچیده‌تر را نیز در نظر می‌گیرند و در صورت تنوع داده‌ها از ثبات بالاتری برخوردار هستند. هر سه این مدل‌ها قابل استفاده هستند و هر کدام مزایای خود را دارند.
- دقت طبقه‌بندی مناطق سیل‌خیز و غیر سیل‌خیز با تغییر توان تفکیک مکانی، دقت ارتفاعی و ماهیت مدل رقومی ارتفاع تغییر می‌کند. به‌طور کلی مدل‌های رقومی بادقت پایین‌تر در تعیین پهنه‌های ریسک سیلاب خصوصاً مناطق با خطر متوسط و بالا دارای عدم قطعیت هستند.
- پارامترهای مرتبط با توپوگرافی نقش بسیار مهمی در تعیین دقت طبقه‌بندی دارند. این مسئله، به‌ویژه در صورت استفاده از DEM‌های با کیفیت و توان تفکیک مکانی بالا به‌خوبی نمایان می‌گردد.



شکل ۴ نقشه خطر سیل حاصل از DEM ۳۰ متر

منابع

[1] T. Grahm and L. Nyberg, 'Assessment of pluvial flood exposure and vulnerability of residential areas', International Journal of Disaster Risk Reduction, vol. 21, pp. 367–375, 2017.

[2] C. M. Kaya and L. Derin, 'Parameters and methods used in flood susceptibility mapping: a review', Journal of Water and Climate Change, vol. 14, no. 6, pp. 1935–1960, 2023.

[۳] امینی، محمدامین، ترکان، غزاله، اسلامیان، سعید، زارعیان، محمد جواد، بسالت پور، علی اصغر. (۱۳۹۷). ارزیابی مدل هیدرولوژیک SWAT در شبیه‌سازی بیلان آب در حوضه‌های آبریز مناطق نیمه-خشک (مطالعه موردی: حوضه آبریز زاینده‌رود). آب و خاک، ۳۳(۵).

Abstract

This study presents a framework for flood-prone area zoning and the development of a flood hazard probability map using machine learning algorithms in the upstream region of the Zayandeh-Rud River. The primary objective of the research is to examine the impact of Digital Elevation Model (DEM) quality on flood zoning. For this purpose, a 10-meter DEM (derived from national base topographic maps) and a 30-meter SRTM DEM were used. Initially, flood zoning parameters were extracted from these DEMs, and after adding other relevant factors such as land use, geology, soil, etc., machine learning models—including Random Forest, Support Vector Machine, and Logistic Regression—were employed for classification and zoning. The results indicate that increasing the spatial resolution and vertical accuracy of the DEMs leads to more precise outcomes. The best results were obtained using the Random Forest and Support Vector Machine models, both of which performed flood-prone and non-flood-prone area classification with similar accuracy. Both models achieved an accuracy of over 98% in this classification process, whereas the use of the 30-meter DEM resulted in reduced accuracy according to various indices. The findings clearly show that the resolution of the DEM has a significant impact on improving or reducing classification accuracy. Moreover, the evaluation of the importance of various features extracted from both the 10-meter and 30-meter DEMs revealed that with higher resolution, the role of elevation-derived parameters increases, while in lower-resolution DEMs, the importance of some topography-related parameters decreases. Finally, after training and evaluating the machine learning models and selecting the best one, flood hazard probability maps were generated for both the 10-meter and 30-meter DEMs. The qualitative assessment showed that the 10-meter DEM led to a more accurate map for determining flood risk probability and aligned more closely with reality compared to the map derived from the 30-meter DEM.

Key words: Flood Zoning, Digital Elevation Model (DEM), Machine Learning, Flood Probability