Kharazmi University

Spatial Analysis Environmental Hazards

Online ISSN: 2588-5146
Homepage: https://khu.ac.ir

Detection of flood zones using machine learning algorithms and remote sensing to determine the area of land under flood damage (case study: March 2019 flood in Aqqla city)

Omid Ashkriz[1]✉ | Fatemeh Falahati [2] | Seyyed Amir Hossein Garakani[3]

1. Corresponding author, Natural Disasters Research Institute, Tehran, Iran. E-mail: omidashkriz7@mail.com
2. Department of Hydrological Hazards Research, Natural Disasters Research Institute, Tehran, Iran. E-mail: falahati_fp@yahoo.com
3. Department of Architecture, Faculty of Architecture and Urban Planning, Central Tehran Branch, Islamic Azad University, Tehran, Iran. E-mail: garakani@ndri.ac.ir

Article Info

Article type:
Research Article

Article history:
Received 27 February 2023
Accepted 02 November 2024
Published online 23 December 2024

Keywords:
Detection of the flood zone,
Spatial Cross Validation Method,
Remote Sensing,
Random Forest,
Multilayer Perceptron Neural Network,
Support Vector Machine.

ABSTRACT

Objective: The growth of settlements and the increase in human activities in floodplains, especially riverbanks and flood-prone areas, have increased the amount of damage caused by this hazard. Therefore, determining the extent of flooding in line with risk reduction, preparedness and response, and recovery and rehabilitation planning after this hazard is of great importance.

Methods: Research Method: In the present study, common machine learning algorithms and Sentinel 2 image classification were used to produce land cover maps, detect flood zones, and determine the area of land damaged by the April 2019 flood in Aqqala County. Also, in order to examine and increase the accuracy of the algorithms, three normalized spectral indices of vegetation (NDVI), water areas (MNDWI), and built-up areas (NDBI) were combined with the images used. Different parameters of each algorithm were evaluated using spatial cross-validation to determine their effect on classification accuracy and to prevent optimistic results due to spatial correlation between the training and experimental samples.

Results: The results showed that the combination of spectral indices led to an increase in the overall accuracy of the algorithms and in order to produce land cover maps, the random forest algorithm with an accuracy of 83.08% due to the use of the ensemble learning method has higher accuracy and generalizability than other support vector machine and neural network algorithms with an accuracy of 11.79 and 44.75%, respectively. After determining the most accurate algorithm, the floodplain map was produced using the random forest algorithm in two classes of irrigated and non-irrigated lands, and the overall accuracy of the algorithm was 40.93% in the most optimal parameters and with the combination of the spectral index (MNDWI). Then, by overlapping the land cover and floodplain maps, the area of built-up land, cropland, and green spaces under flood damage was estimated to be 2008.4 and 0772.41 square kilometers, respectively.

Conclusions: Based on the results obtained, the generated land cover and floodplain maps provide reliable results for estimating and understanding the effects of flood risk.

Cite this article: Ashkriz, O., Falahati, F., & Garakani, S.A.H. (2024). Detection of flood zones using machine learning algorithms and remote sensing to determine the area of land under flood damage (case study: March 2019 flood in Aqqla city). *Spatial Analysis Environmental Hazards*, 11 (3), 35-54. https://doi.org/10.61186/jsaeh.11.3.1

© The Author(s)
Publisher: Kharazmi University
DOI: https://doi.org/10.61186/jsaeh.11.3.1

[Downloaded from system.khu.ac.ir on 2025-12-05]
[DOI: 10.61186/jsaeh.11.3.1]

EXTENDED ABSTRACT

Introduction

People around the world face events every day that result in death, injury, property damage, and disruption of daily activities. Floods, as one of the natural hazards, have threatened and invaded rural and urban communities for a long time and have caused human and financial losses. The growth of settlements and the increase of human activities in flood plains, especially along the banks of rivers and flood-prone places, have increased the amount of damage caused by this hazard. Therefore, flood is considered as one of the events that causes damages to human societies. According to the report of the World Bank, after the earthquake, the flood is the second natural disaster in Iran, whose destructive effects have affected people's lives more than the disaster. Therefore, it is very important to determine the flood extent in line with risk reduction planning, preparedness and response, and recovery and rehabilitation after this hazard.

Methods

The use of remote sensing and GIS techniques in the shortest possible time provides accurate mapping of the area affected by floods and provides the possibility of detailed investigation of cases related to this hazard. Among the different types of satellite data, Sentinel 2 satellite images are one of the sources of information that provide appropriate data with spatial and temporal resolution in any situation, especially during floods. For this purpose, Sentinel 2 satellite images were used in this research. These images are related to the closest time before the flood in order to produce land cover maps of the study area, as well as the peak time of the flood in order to produce a map of flood zones. There are various methods for extracting information from satellite data, in the current research, from the popular and advanced algorithms of random forest, support vector machine and multilayer perceptron neural network due to their high accuracy and flexibility compared to other algorithms. Were used to produce land cover maps, reveal flood zones and determine the area of lands damaged by the flood of March 2019 in Aqqla city. Also, in order to check and increase the accuracy of the algorithms, three normalized spectral indices of vegetation (NDVI), water areas (MNDWI) and built-up land (NDBI) were combined with the used images. The different setting parameters of each of the algorithms were evaluated by spatial cross-validation method in order to determine their effect on classification accuracy and prevent optimistic results due to spatial correlation between training and experimental samples.

Results

The results showed that the random forest algorithm with an overall accuracy of 82.21% compared to the support vector machine and artificial neural network algorithms with the values of 78.96% and 73.58%, respectively, has higher accuracy and generalizability regarding the production of land cover maps. has Also, the implementation of these algorithms with the combination of spectral indices has led to a significant increase in the overall accuracy of the algorithms, and the random forest algorithm is the most accurate algorithm with 83.08% compared to other support vector machine and network algorithms. Artificial neural networks are 79.11% and 75.44% respectively. Also, the spatial accuracy of the land cover maps obtained from the aforementioned algorithms was calculated using 30 percent of the samples and the Kappa index was evaluated and compared. The results of this index showed that there is a direct relationship between the overall accuracy of the algorithms and the Kappa index. The random forest algorithm combined with spectral indices showed a

[Downloaded from system.khu.ac.ir on 2025-12-05]

[DOI: 10.61186/jsaeh.11.3.1]

coefficient of 82.38%, which is more accurate than the support vector machine algorithm with 75.52% and the neural network algorithm with 69.12% for the separation of land cover classes. The random forest algorithm uses the collective learning method to separate the classes, which leads to the reduction of variance, therefore it is one of the most accurate machine learning algorithms. After determining the most accurate machine learning algorithm, the map of flood zones was produced using the random forest algorithm in two classes of water and non-water lands, and the overall accuracy of the algorithm in the most optimal parameters and with the combination of spectral index (MNDWI) 93.40% and the spatial accuracy of the flood zones map was obtained with a kappa index of 91.30%. With the cell-by-cell measurement of the most accurate land cover map with flood zone map, the area covered by flood in built-up lands was 4.008 square kilometers and in agricultural lands and green spaces was calculated as 41.0772 square kilometers. Based on the obtained results, the produced land cover maps and flood zones provide a reliable result in order to estimate and understand the effects of flood risk.

Author Contributions

All authors participated in data collection, preparation of the research report, and data analysis.

Data Availability Statement

Data available on request from the authors.

Acknowledgements

The aforementioned article is extracted from the research project titled “Detecting Flood Zones Using Machine Learning Algorithms and Remote Sensing to Determine the Area of Land Under Flood Damage and Optimal Location of Temporary Shelters (Case Study: April 2019 Flood in Aqqala County)”. Therefore, the Natural Disasters Research Institute is gratefully acknowledged for all its material and moral support.

Ethical considerations

The authors avoided data fabrication, falsification, plagiarism, and misconduct.

Funding

This article was supported by the Natural Disasters Research Institute.

Conflict of interest

The authors declare that there is no conflict of interest in this article

[Downloaded from system.khu.ac.ir on 2025-12-05]

[DOI: 10.61186/jsaeh.11.3.1]

دانشگاه خوارزمی

تحلیل فضایی مخاطرات محیطی

شاپا الکترونیکی: **۵۱۴۶- ۲۵۸۸**

Homepage: https://khu.ac.ir

آشکارسازی پهنه‌های سیلابی و تعیین مساحت اراضی تحت خسارت سیلاب با استفاده از الگوریتم‌های یادگیری ماشین و سنجش از دور (مطالعه موردی: سیل فروردین ۹۸ شهرستان آق‌قلا)

امید اشکریز [۱] ✉| فاطمه فلاحتی [۲] | سیدامیرحسین گرکانی [۳]

۱. نویسنده مسئول، پژوهشکده سوانح طبیعی، تهران، ایران. رایانامه: omidashkriz7@mail.com

۲. گروه پژوهشی مخاطرات هیدرولوژیکی، پژوهشکده سوانح طبیعی، تهران، ایران. رایانامه: falahati_fp@yahoo.com

۳. گروه معماری، دانشکده معماری و شهرسازی، واحد تهران مرکزی، دانشگاه آزاد اسلامی، تهران، ایران. رایانامه: garakani@ndri.ac.ir

اطلاعات مقاله

نوع مقاله:
مقاله پژوهشی

تاریخ دریافت: ۱۴۰۱/۱۲/۰۸
تاریخ پذیرش: ۱۴۰۳/۰۸/۱۲
تاریخ انتشار: ۱۴۰۳/۱۰/۰۳

کلیدواژه‌ها:
آشکارسازی پهنه سیلاب،
روش اعتبارسنجی متقابل مکانی،
سنجش از دور،
الگوریتم جنگل تصادفی،
الگوریتم شبکه عصبی پرسپترون چند لایه،
الگوریتم ماشین بردار پشتیبان.

چکیده

هدف: رشد سکونتگاه‌ها و افزایش فعالیت‌های انسانی در دشت‌های سیلابی به ویژه حاشیه رودخانه‌ها و مکان‌های مستعد سیل، میزان خسارت ناشی از این مخاطره را افزایش داده است. از این رو تعیین گستره سیلاب در راستای برنامه‌ریزی‌های کاهش خطر، آمادگی و پاسخ و بازیابی و بازتوانی پس از این مخاطره از اهمیت بسزایی برخوردار است.

روش پژوهش: در مطالعه حاضر از الگوریتم‌های متداول یادگیری ماشین و طبقه‌بندی تصاویر سنتینل۲ جهت تولید نقشه‌های پوشش اراضی، آشکارسازی پهنه‌های سیلابی و تعیین مساحت اراضی تحت خسارت سیلاب فروردین۱۳۹۸ شهرستان آق‌قلا استفاده شد. همچنین به منظور بررسی و افزایش دقت الگوریتم‌ها، سه شاخص طیفی نرمال شده پوشش گیاهی (NDVI)، پهنه‌های آبی (MNDWI) و اراضی ساخته شده (NDBI) با تصاویر مورد استفاده ترکیب شدند. پارامترهای مختلف تنظیم هر یک از الگوریتم‌ها به منظور تعیین تأثیر آن‌ها بر دقت طبقه‌بندی و جلوگیری از کسب نتایج خوشبینانه ناشی از همبستگی مکانی میان نمونه‌های آموزشی و آزمایشی، با روش اعتبارسنجی متقابل مکانی ارزیابی شدند.

یافته‌ها: نتایج نشان داد که ترکیب شاخص‌های طیفی منجر به افزایش دقت کلی الگوریتم‌ها شده و به منظور تولید نقشه‌های پوشش اراضی، الگوریتم جنگل تصادفی با دقت ۸۳/۰۸ درصد بدلیل استفاده از روش یادگیری جمعی از دقت و تعمیم‌پذیری بالاتری نسبت به سایر الگوریتم‌های ماشین بردار پشتیبان و شبکه عصبی با دقت به ترتیب ۷۹/۱۱ و ۷۵/۴۴ درصد برخوردار است. پس از مشخص شدن دقیق‌ترین الگوریتم، نقشه پهنه‌های سیلابی با استفاده از الگوریتم جنگل تصادفی در دو کلاس اراضی آبی و غیر آبی تولید گردید و دقت کلی الگوریتم در بهینه‌ترین پارامترها و با ترکیب شاخص طیفی(MNDWI) ۹۳/۴۰ درصد بدست آمد. سپس با همپوشانی نقشه‌های پوشش اراضی و پهنه‌های سیلابی، سطح اراضی ساخته شده و اراضی زراعی و فضاهای سبز تحت خسارت سیلاب به ترتیب ۴/۲۰۰۸ و ۴۱/۰۷۷۲ کیلومتر مربع برآورد گردید.

نتیجه‌گیری: بر اساس نتایج بدست آمده نقشه‌های پوشش اراضی و پهنه‌های سیلابی تولید شده نتیجه قابل اعتمادی به منظور تخمین و درک اثرات مخاطره سیلاب ارائه می‌دهند.

استناد: اشکریز، امید؛ فلاحتی، فاطمه؛ و گرکانی، سیدامیرحسین (۱۴۰۳). آشکارسازی پهنه‌های سیلابی و تعیین مساحت اراضی تحت خسارت سیلاب با استفاده از الگوریتم‌های یادگیری ماشین و سنجش از دور (مطالعه موردی: سیل فروردین ۹۸ شهرستان آق‌قلا). *تحلیل فضایی مخاطرات محیطی*، ۱۱ (۳)، ۳۵-۵۴. https://doi.org/10.61186/jsaeh.11.3.1

ناشر: دانشگاه خوارزمی

© نویسندگان

[Downloaded from system.khu.ac.ir on 2025-12-05]

[DOI: 10.61186/jsaeh.11.3.1]

مقدمه

امروزه سیل به عنوان یکی از شایع‌ترین مخاطرات طبیعی است که در سراسر جهان، محصولات کشاورزی، اموال و زیرساخت‌ها را به طور جدی تحت تأثیر قرار می‌دهد و با خسارت‌های جانی و تأثیر بر اقتصاد ملی، منجر به زیان‌های جبران ناپذیر می‌گردد (لاکشمی[۱]، ۲۰۱۶). بطوری که در سه دهه گذشته، مخاطره سیلاب حدود ۲/۸ میلیارد نفر را در جهان تحت تاثیر قرار داده و تلفات انسانی و خسارات قابل توجهی به اموال و اقتصاد را در پی داشته است (رحمان و خان، ۲۰۱۱). طبق آمار، سیلاب بیش از ۴۷ درصد از کل مخاطرات جوی را در سراسر جهان تشکیل می‌دهد (هالگاموژ و نیرمالاتاس[۲]، ۲۰۱۷). در ایران نیز بدلیل وسعت زیاد، اقلیم‌های متعدد، تغییر پوشش اراضی، تراکم زمانی و مکانی بارش‌ها در اکثر حوضه‌های آبخیز همه ساله شاهد سیلاب‌های عظیمی هستیم (پوراکرمی، ۱۳۹۶) و بر اساس گزارش بانک جهانی، بعد از زلزله، وقوع سیلاب دومین مخاطره طبیعی فراگیر در ایران می‌باشد که آثار مخرب آن زندگی مردم را تحت تأثیر قرار داده است (بانک[۳]، ۲۰۱۹). بنابراین برنامه‌ریزی در راستای مدیریت بحران و نیاز به پایش و ارزیابی سیلاب، اطلاعات مکانی و زمانی مورد اعتماد و به‌روزی را می‌طلبد که آشکارسازی و نمایش دقیق گستره مکانی مناطق آبگرفته از جمله این اطلاعات است. این اطلاعات به منظور ارزیابی و محاسبه خسارات وارده، برنامه‌ریزی در راستای کاهش خطر، آمادگی پاسخ و بازیابی و بازتوانی پس از سانحه حائز اهمیت است (موناسینگ[۴] و همکاران، ۲۰۱۸).

با توجه به توسعه روزافزون علوم و قابلیت استفاده از فناوری‌های نوین در مراحل مختلف مدیریت بحران، ابزارهای متنوعی جهت پایش و ارزیابی مخاطره سیلاب در یک منطقه در دسترس است که یکی از این ابزارها، داده‌های سنجش از دور می‌باشد (فیروزآبادی، ۲۰۰۳). استفاده از تکنیک‌های سنجش از دور و GIS در کمترین زمان، نقشه‌برداری دقیقی از گستره تحت تأثیر سیلاب ارائه و امکان بررسی دقیق موارد مرتبط با این مخاطره را فراهم می‌کند (خالید[۵] و همکاران، ۲۰۱۸). در بین انواع مختلف داده‌های ماهواره‌ای، تصاویر ماهواره سنتینل[۶] یکی از منابع اطلاعاتی است که در هر شرایطی، به ویژه در سیلاب‌ها، داده‌های مناسبی با قدرت تفکیک مکانی و زمانی ارائه دهند (یانگ[۷] و همکاران، ۲۰۱۷). روش‌های مختلفی جهت استخراج اطلاعات از داده‌های ماهواره‌ای وجود دارد (توکلی و همکاران، ۲۰۱۸). در یک تقسیم‌بندی کلی روش‌های استخراج اطلاعات از تصاویر ماهواره‌ای به دو دسته‌ی روش‌های طبقه‌بندی با نظارت و روش‌های طبقه‌بندی بدون نظارت مبتنی بر تشخیص تغییرات طیفی و خوشه‌بندی تقسیم می‌شوند (حسینی و همکاران، ۲۰۱۵).

در طبقه‌بندی نظارت‌شده نظر کارشناس به طور مستقیم در تصمیم‌گیری در مورد داده‌های آموزشی دخیل است و بر نتایج نهایی تأثیر می‌گذارد (غیور و همکاران، ۲۰۲۱). الگوریتم‌های طبقه‌بندی تحت نظارت شامل الگوریتم حداقل فاصله (آل احمدی و حامس[۸]، ۲۰۰۹)، الگوریتم حداکثر احتمال (کانتاکومار و نیلمستتی[۹]، ۲۰۱۵)، الگوریتم شبکه عصبی مصنوعی (ژانگ[۱۰] و همکاران، ۲۰۱۸؛ عمر[۱۱] و همکاران، ۲۰۱۵)، الگوریتم جنگل تصادفی (تلوگونتلا[۱۲] و همکاران، ۲۰۱۸؛ آدام[۱۳] و همکاران، ۲۰۱۴) و الگوریتم

---

1 - Lakshmi
2 - Halgamuge & Nirmalathas
3 - Bank
4 - Munasinghe
5 - Khalid
6 - Sentinel
7 - Yang
8 - Al-Ahmadi & Hames
9 - Kantakumar and Neelamsetti
10 - Zhang
11 - Omer
12 - Teluguntla
13 - Adam

[Downloaded from system.khu.ac.ir on 2025-12-05]
[DOI: 10.61186/jsaeh.11.3.1]

ماشین بردار پشتیبان (تالوکدار[1] و همکاران، ۲۰۲۰) می‌باشد. در مقایسه با الگوریتم‌های بدون نظارت، الگوریتم‌های تحت نظارت در تمایز کلاس بهتر عمل می‌کنند و از این رو دقت بهتری را ارائه می‌دهند (بوری[2] و همکاران، ۲۰۱۸). اگرچه در مطالعات مختلف از انواع الگوریتم‌های طبقه‌بندی استفاده می‌شود، اما در سال‌های اخیر الگوریتم‌های پیشرفته یادگیری ماشین به ویژه در ارزیابی و پیش‌بینی خطرات طبیعی مانند سیل، به دلیل دقت و انعطاف‌پذیری بالا بیشتر از آن‌ها نسبت به سایر الگوریتم‌ها استفاده می‌شود (غیور و همکاران، ۲۰۲۱).

در پژوهش حاضر به منظور آشکارسازی و ارزیابی پهنه خسارات ناشی سیل از تصاویر ماهواره‌ای سنتینل۲ و با استفاده از الگوریتم‌های متداول یادگیری ماشین مانند: شبکه عصبی مصنوعی، ماشین بردار پشتیبان و جنگل تصادفی به همراه شاخص‌های طیفی نرمال شده تفاوت پوشش گیاهی (NDVI)، اراضی ساخته شده (NDBI) و پهنه‌های آبی (MNDWI) استفاده شد و سپس از روش ارزیابی متقابل مکانی دقیق‌ترین الگوریتم در میان سایر الگوریتم‌ها جهت بر آورد خسارات مورد بررسی قرار گرفت. هدف از این پژوهش ترسیم یک نقشه دقیق از گستره مکانی سیلاب و احتمال خسارت وارده به پوشش اراضی در کمترین زمان ممکن بوده که می‌تواند در امدادرسانی، مدیریت بحران و همچنین برنامه‌ریزی‌های آتی مورد توجه مدیران و برنامه‌ریزان قرار گیرد.

پیشینه پژوهش

۱. پیشینه تجربی

از میان تحقیق‌های انجام شده با استفاده از الگوریتم‌های یادگیری ماشین می‌توان به موارد ذیل اشاره کرد:

اپیکی (۱۳۹۶) با استفاده از الگوریتم شاخص تفاضل پوشش گیاهی و طبقه‌بندی نظارت شده سلول‌ها در تصاویر ماهواره‌ای مربوط به سنجده‌های OLI و MODIS، سیل مربوط به آبان ۱۳۹۴ استان ایلام را ارزیابی کرد. نتایج نشان داد ۸۵ درصد از پوشش گیاهی شهرستان دره شهر از بین رفته است.

تراهی و حسنی مقدم (۱۳۹۷) نیز به تعیین گستره سیل سال ۱۳۹۵ دزفول با استفاده از داده‌های سنجنده OLI و الگوریتم ماشین بردار پشتیبان پرداختند. در این پژوهش نتایج نشان داد بیشترین میزان گستره سیل مربوط به زمین‌های کشاورزی با گسترش ۸۴۶۷/۴۵ هکتار و مناطق تفریحی و گردشگری حاشیه رودخانه با گسترش ۲۶۵۹/۱۴ هکتار بوده است.

بی‌غم سرشکه و همکاران (۱۳۹۹) به ارزیابی صحت روش‌های طبقه‌بندی تصویر ماهواره‌ای سنتینل ۲ با رویکرد پیکسل‌مبنا و شی‌مبنا و با به کارگیری دو الگوریتم جنگل تصادفی و ماشین بردار پشتیبان در پهنه‌بندی مناطق سیل‌گیر رودخانه طالقان پرداختند و نتایج نشان داد هر دو روش در رویکرد شی‌مبنا اغلب بیش‌ترین همپوشانی را با دوره‌های بازگشت مختلف داشته‌اند و بالاترین میزان هم‌پوشانی با پهنه‌های سیلاب با دوره بازگشت دو ساله بترتیب با مقادیر ۶۸ و ۶۶ درصد می‌باشد.

هوانگ و جین[3] (۲۰۲۰) به نقشه‌برداری و ارزیابی سریع سیلاب با طبقه‌بندی تحت نظارت و تشخیص تغییر با استفاده از داده‌های نوری سنتینل۲ و راداری سنتینل۱ پرداختند. در این پژوهش نقشه سیل استخراج شده با روش‌های آستانه‌گذاری (تصاویر سنتینل۱) و شاخص‌های آب معمولی (تصاویر سنتینل ۲) مقایسه شده است. نتایج نشان داد این رویکرد می‌تواند سیل در مناطق پوشش گیاهی را به خوبی شناسایی کند. علاوه بر این، همه داده‌های مورد نیاز و پردازش داده‌ها ساده هستند، بنابراین می‌توان آن‌ها را در نقشه‌برداری سریع سیلاب در امدادرسانی اولیه از بلایا رواج داد.

همچنین سجاد و همکاران (۲۰۲۰) با استفاده از تکنیک‌های سنجش از دور و GIS به پایش عملیاتی و ارزیابی خسارت سیل رودخانه‌ای ۲۰۱۴ در دشت چناب پاکستان پرداختند. در این پژوهش تصاویر لندست ۸ برای دوره‌های قبل از سیل، همزمان سیل و پس از سیل برای تجزیه و تحلیل جامع و ترسیم وسعت سیل، ارزیابی خسارت و مدت زمان آن مورد استفاده قرار گرفت. از

[1] - Talukdar
[2] - Boori
[3] - Huang & Jin

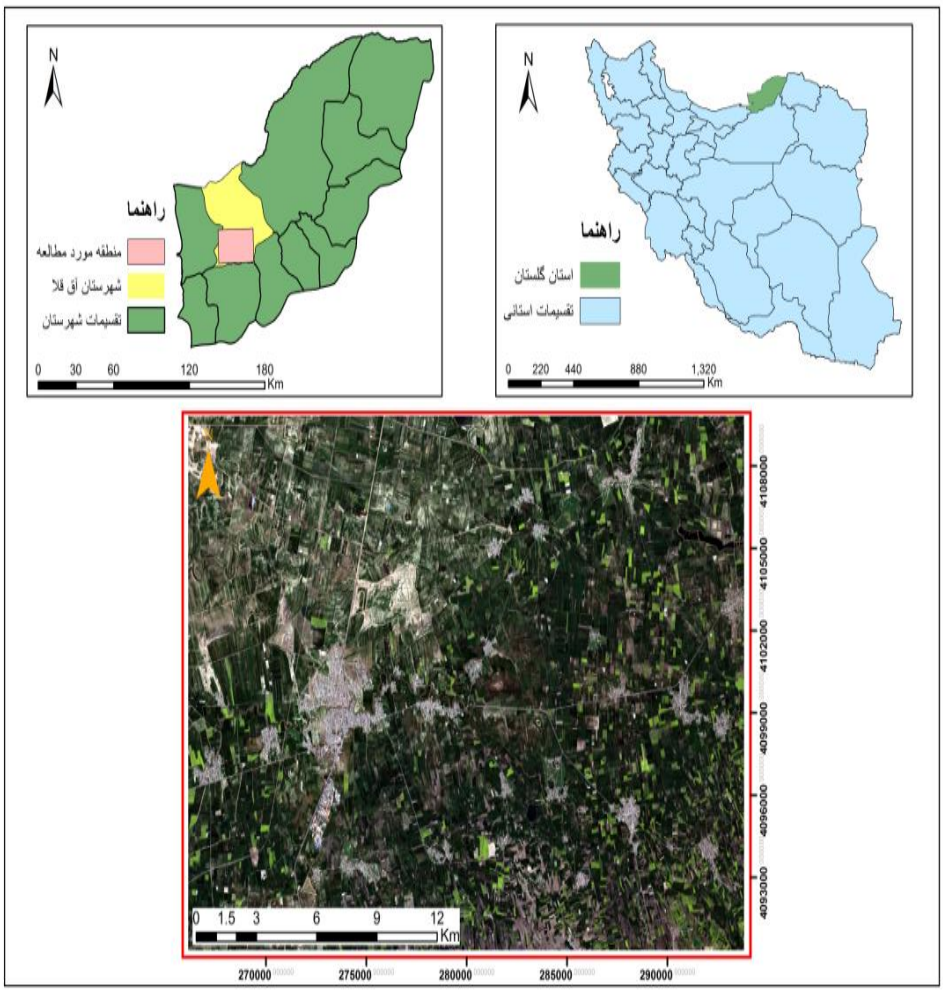
[Downloaded from system.khu.ac.ir on 2025-12-05]
[DOI: 10.61186/jsaeh.11.3.1]

روش طبقه‌بندی نظارت‌شده برای تعیین تغییرات کاربری اراضی، و شاخص تفاوت نرمال‌شده آب (MNDWI) به دست آمده از ماهواره‌ای برای شناسایی مناطق سیل‌زده استفاده کردند. تجزیه و تحلیل داده‌ها اجازه داد تا مدت زمان طغیان سیل، خسارات وارده به مناطق مسکونی و کشاورزی محاسبه کنند.

روش شناسی پژوهش

۱. قلمرو جغرافیایی مورد مطالعه

استان گلستان از مناطق سیل‌خیز ایران است و طی سال‌های اخیر با چندین سیلاب مخرب مواجه گردیده که تلفات انسانی و خسارات اقتصادی فراوانی را در پی داشته است. شهرستان آق‌قلا در شمال این استان و در ۵۴ درجه و۱۴ دقیقه تا ۵۴ درجه و ۵۱ دقیقه طول شرقی و ۳۶ درجه و ۵۵ دقیقه تا ۳۰ درجه و ۳۷ دقیقه عرض شمالی و در دو طرف رودخانه گرگان‌رود واقع شده است. جمعیت این شهرستان در حدود ۱۰۹ هزار نفر و وسعت آن معادل یک هزار و ۷۶۳ کیلومترمربع می‌باشد (اداره آب منطقه ای استان گلستان). رودخانه گرگان‌رود یکی از مهم‌ترین رودخانه‌های سیل‌خیز این استان می‌باشد و با طول حدود ۳۰۰ کیلومتر و حوضه آبریزی به وسعت ۱۰۲۵۰کیلومتر مربع از دامنه‌های شمال شرقی البرز و دامنه‌های غربی ارتفاعات استان خراسان شمالی سرچشمه می‌گیرد (کاردان و همکاران، ۱۳۹۶). در شکل ۱ موقعیت جغرافیایی منطقه مورد مطالعه نشان داده شده است.



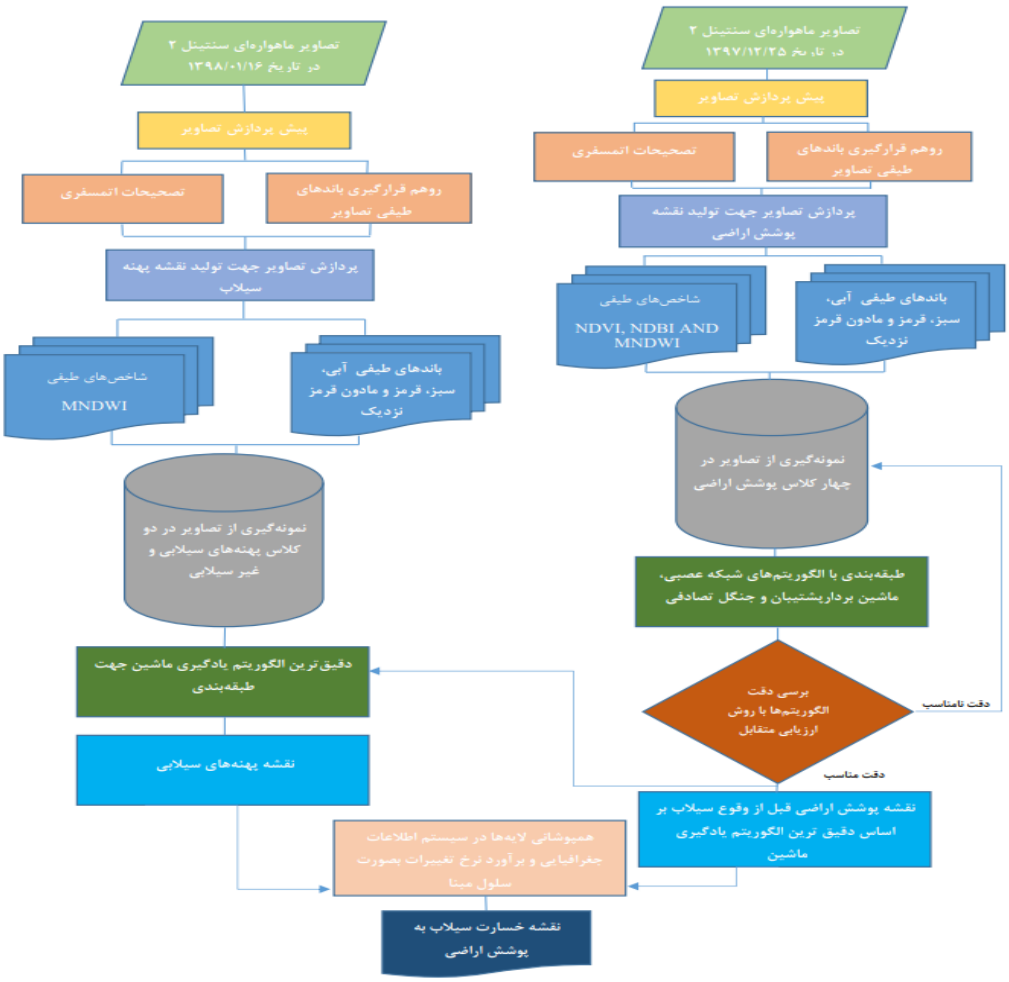
شکل ۱. موقعیت جغرافیایی منطقه مورد مطالعه

۲. داده‌ها و روش کار

در ابتدای این بخش جهت توصیف بهتر روند انجام تحقیق، روش کار بصورت فلوچارت در شکل ۲، نمایش داده شده است.

[Downloaded from system.khu.ac.ir on 2025-12-05]

[DOI: 10.61186/jsaeh.11.3.1]



شکل ۲. فرایند انجام پژوهش

۲-۱. تهیه نقشه‌های پوشش اراضی منطقه مورد مطالعه

ابتدا به منظور تولید نقشه‌های پوشش اراضی و نقشه پهنه‌های سیلابی تصاویر ماهواره‌ای سنتینل۲ با مشخصات ذیل (جدول ۱) از وب‌سایت مرکز داده علمی کوپرنیک[1] (https://scihub.copernicus.eu) اخذ گردید و چهار باند داده سنتینل ۲ شامل: آبی (باند۲ در ۴۹۰ نانومتر) سبز (باند ۳ در۵۶۰ نانومتر)، قرمز (باند ۴ در ۶۶۵ نانومتر) و مادون قرمز نزدیک (باند ۸ در ۸۴۲ نانومتر) به کار گرفته شد.

جدول ۱. مشخصات تصاویر ماهواره‌ای سنتینل۲

شماره تصویر	ماهواره	نوع سنجنده	تاریخ اخذ تصویر	شماره ردیف و گذر	قدرت تفکیک مکانی	قدرت تفکیک رادیومتریک
۱	سنتینل ۲	MSI	۱۳۹۷/۱۲/۲۵	T40SBG	۱۰متر	۱۶ بیت
۲	سنتینل ۲	MSI	۱۳۹۸/۱/۱۶	T40SBG	۱۰متر	۱۶ بیت

با بررسی‌های بصری و تهیه تصاویر متعدد با توجه به عدم وجود ابر در منطقه مورد مطالعه دو تصویر جهت تولید نقشه پوشش اراضی و نقشه پهنه‌های سیلابی انتخاب گردید، که به ترتیب مشخصات آن در جدول ۱ نمایش داده شده است. تصویر اول مشخصات نزدیک‌ترین تصویر قبل وقوع سیلاب و تصویر دوم اوج سیلاب در منطقه مورد مطالعه می‌باشد. تصاویر در سطح

[1] - Copernicus

[Downloaded from system.khu.ac.ir on 2025-12-05]
[DOI: 10.61186/jsaeh.11.3.1]

L2A تهیه گردید، که تصحیحات رادیومتریک و هندسی توسط آژانس فضایی اروپا انجام شده است. تصحیح خطای اتمسفری تصاویر نیز با استفاده از جعبه ابزار Sen2Cor نرم افزار اسنپ[۱] انجام شد. این ابزار تصحیح خطای اتمسفری را از طریق فرآیند (TOA[۲]) انجام می‌دهد (اسلام و همکاران، ۲۰۱۹).

۲-۱-۱. استخراج شاخص‌های طیفی

به منظور طبقه‌بندی تصاویر ماهواره‌ای سنتینل ۲ و کاهش اختلاط کلاس‌ها و افزایش صحت طبقه‌بندی، سه شاخص گیاهی تفاضلی نرمال شده[۳](NDVI)، شاخص اراضی ساخته شده تفاضلی نرمال شده[۴] (NDBI) و شاخص آب تفاضلی نرمال شده اصلاحی[۵](MNDWI) مورد استفاده قرار گرفت (ریمال[۶] و همکاران، ۲۰۲۰).

الف- شاخص طیفی پوشش گیاهی

شاخص گیاهی تفاضلی نرمال شده بر حسب دو باند قرمز و مادون قرمز نزدیک به صورت زیر تعریف می‌شود (فاطمی و همکاران، ۱۳۹۲):

$$NDVI = \frac{NIR - R}{NIR + R}$$ رابطه ۱

NDVI یک شاخص طیفی مهم برای تشخیص پوشش گیاهی است و همچنین به طور گسترده برای تجزیه و تحلیل ثبات سنجنده چند طیفی MSI استفاده شده است (روی[۷] و همکاران، ۲۰۱۶).

ب- شاخص طیفی اراضی ساخته شده

از این شاخص برای نقشه‌برداری خودکار مناطق ساخته شده شهری و با استفاده از پاسخ‌های طیفی منحصر به فرد مناطق ساخته‌شده و سایر انواع پوشش زمین مورد استفاده قرار می‌گیرد (دلیری[۸] و همکاران، ۲۰۲۰). این شاخص با استفاده از رابطه ۲ محاسبه می‌شود:

$$NDBI = \frac{SWLR - NIR}{SWIR + NIR}$$ رابطه ۲

در این روش از باندهای مادون قرمز میانی (SWIR) و مادون قرمز نزدیک (NIR) برای شکل‌دهی مناطق ساخته شده استفاده شد.

ج- شاخص طیفی پهنه آبی

شاخص آب تفاضلی نرمال شده اصلاحی (MNDWI) یک شاخص بهبود یافته از شاخص آب تفاضلی نرمال شده (NDWI) برای تشخیص پهنه‌های آبی است (سو[۹]، ۲۰۰۶). این شاخص به طور گسترده‌ای برای تعیین سریع آب سیلاب برای پایش و ارزیابی خسارات سیل استفاده (ریویلا[۱۰] و همکاران، ۲۰۱۵) و بر اساس رابطه ۳ محاسبه می‌شود:

$$MNDWI = \frac{Green - SWIR}{Green + SWIR}$$ رابطه ۳ :

این شاخص با استفاده از طول موج های سبز (Green) و مادون قرمز موج نزدیک میانی (SWLR) برای ترسیم پهنه‌های آبی مورد استفاده قرار می‌گیرد (سجاد و همکاران، ۲۰۲۰).

[1] - SNAP
[2] - Top-of-Atmosphere
[3] - Normalized Difference Vegetation Index
[4] - Normalized Difference Built-up Index
[5] - Modification of Normalised Difference Water index
[6] - Rimal
[7] - Roy
[8] - Deliry
[9] - Xu
[10] - Revilla

[Downloaded from system.khu.ac.ir on 2025-12-05]
[DOI: 10.61186/jsaeh.11.3.1]

۲-۲. الگوریتم‌های یادگیری ماشین استفاده شده در این پژوهش به منظور مقایسه خروجی‌ها و تعیین دقیق ترین الگوریتم

الف- الگوریتم شبکه عصبی مصنوعی

شبکه عصبی پرسپترون چند لایه[۱] (MLP) یکی از موثرترین تکنیک‌های شبکه عصبی مصنوعی است، که با تقریب دقیق‌تر به طور گسترده برای مدل‌سازی فرآیندهای غیرخطی و پیچیده دنیای واقعی مورد استفاده قرار می‌گیرد (بوی[۲] و همکاران، ۲۰۲۰؛ ساندوسکی[۳] و همکاران، ۲۰۱۸). اساساً، MLP یک مدل پیشخور است که شامل یک لایه ورودی، یک یا چند لایه پنهان و یک لایه خروجی است (کاوزوغلو و ماتر[۴]، ۲۰۰۳). در این الگوریتم تعداد لایه و نورون‌های پنهان بر اساس مجموعه داده‌های آموزشی باید کمی‌سازی شوند (غیور و همکاران، ۲۰۲۱).

ب- الگوریتم ماشین بردار پشتیبان

الگوریتم ماشین‌بردار پشتیبان مرز تصمیم‌گیری بهینه را در یک فضای ویژگی تعیین می‌کند. این مرز تصمیم به عنوان ابر صفحه[۵] نامیده می‌شود، و مسئله طبقه‌بندی را به مجموعه‌ای از کلاس‌های از پیش تعریف شده که با داده‌های آموزشی سازگار است متمایز می‌کند. یک فرآیند تکراری برای یافتن مرز ابرصفحه بهینه در فضای طبقه‌بندی چند بعدی برای تمایز کلاس‌ها در داده‌های آموزشی اعمال می‌کند (محمدعبدی، ۲۰۱۹). توابع کرنل مختلفی وجود دارد که از طریق آن‌ها می‌توان مرزهای ابر صفحه را تعریف کرد. برای باندهای تصاویر سنتینل ۲ که به صورت خطی قابل جداسازی نیستند، از هسته تابع پایه شعاعی به عنوان بهترین تابع استفاده گردید (یو[۶] و همکاران، ۲۰۰۶). پارامترهای این تابع پهنای باند (σ) می‌باشد که مقدار بهینه آن طی فرایند کالیبراسیون به‌همراه عدد ثابت C (ضریب جریمه) تعیین می‌گردد.

ج- الگوریتم جنگل تصادفی

الگوریتم جنگل تصادفی یک الگوریتم یادگیری جمعی[۷] است، این الگوریتم بر اساس ترکیبی از درختان تصمیم که از طریق نمونه برداری بوت استرپ[۸] تولید شده، بهتر از یک طبقه‌بندی منفرد عمل می‌کنند (بریمن[۹]، ۲۰۰۱). مؤلفه بوت استرپ به این معنی است که هر درخت تصمیم جهت طبقه‌بندی با استفاده از مجموعه‌ای از ویژگی‌های نمونه تصادفی با جایگزینی از داده‌های آموزشی انتخاب می‌شود، که این امر باعث عدم همبستگی بین درختان تصمیم می‌گردد (هستی[۱۰] و همکاران، ۲۰۰۹). هر یک از درختان با توجه به ویژگی خود به یک کلاس رأی می‌دهد و در نهایت تخصیص هر نمونه آموزشی به یک کلاس بر مبنای بیشترین تعداد آرایی که از طبقه‌بندی کننده‌ها دریافت می‌کنند، انجام می‌گیرد (گیسلاسون[۱۱]، ۲۰۰۶). در این الگوریتم نیاز به تعیین مقادیر بهینه دو پارامتر تعداد درخت‌های تصمیم و تعداد زیر مجموعه تصادفی از ویژگی‌ها می‌باشد، که روی خطای نهایی مؤثر است (هوانگ و بوتروس[۱۲]، ۲۰۱۶).

۲-۳. طبقه‌بندی و انتخاب نمونه‌های آموزشی و آزمایشی

[1] - Multilayer Perceptron Neural Network
[2] - Bui
[3] - Sadowski
[4] - Kavzoglu and Mather
[5] - Hyperplane
[6] - Yu
[7] - Ensemble Learning
[8] - Bootstrap
[9] - Breiman
[10] - Hastie
[11] - Gislason
[12] - Huang and Boutros

[Downloaded from system.khu.ac.ir on 2025-12-05]
[DOI: 10.61186/jsaeh.11.3.1]

پارامترهای زیادی بر نحوه طبقه‌بندی تصاویر ماهواره‌ای تأثیر می‌گذارند که ازجمله می‌توان به داده‌هـای ورودی و ارائه پیکسل هایی به عنوان نمونه‌های آموزشی اشاره کرد (برینک[1] و همکاران، ۲۰۱۷). لذا در این پژوهش به منظور تهیه نمونه‌های آموزشی از تصاویر ماهواره‌ای سنتینل۲، تصاویر گوگل ارث[2]و نقشه ۱:۲۰۰۰ منطقه مورد مطالعه استفاده گردید. این نوع نمونه‌ها با بررسی‌های بصری در چهار پوشش اراضی عمده مانند: اراضی ساخته شده، اراضی کشاورزی و فضاهای سبز، پهنه‌های آبی و اراضی بایر انتخاب شدند. نمونه‌گیری تصادفی ساده به منظور تولید ۲۶۹۹ نمونه در هر کلاس انجام شد. این مقدار نمونه به اندازه کافی بزرگ است که در منطقه مورد مطالعه پراکنده شود و توان محاسباتی الگوریتم‌ها را تخلیه نمی‌کند (محمدعبدی، ۲۰۱۹). همچنین در این پژوهش جهت عدم تعادل در نمونه‌های آموزشی تعداد نمونه‌ها در هر کلاس به صورت یکسان در نظر گرفته شد. جهت بررسی و ارزیابی تاثیر شاخص‌های طیفی در بهبود دقت نقشه پوشش اراضی الگوریتم‌های مذکور داده‌ها را یکبار بدون شاخص با ۵ باند طیفی (آبی، سبز، قرمز، مادون قرمز و مادون قرمز میانی) و بار دیگر به همراه سه شاخص طیفی متداول طبقه‌بندی کردند.

داده‌های مرجع می‌توانند به طور قابل اعتمادی روی زمین تکرار شوند و راهنمای خوبی برای ارزیابی دقت نقشه‌های تولیدی می‌باشند (اولوفسون[3] و همکاران، ۲۰۱۴). ۲۶۹۹ نمونه به دو بخش تقسیم شدند: مجموعه داده آموزشی شامل ۷۰ درصد (۱۸۹۰) از نمونه‌ها و مجموعه داده آزمایشی شامل۳۰ درصد باقی مانده (۸۰۹) که به هر کلاس نقشه پوشش اراضی به همان تعداد نمونه آموزشی و آزمایشی اختصاص داده شد (غیور و همکاران، ۲۰۲۱؛ محمدعبدی، ۲۰۱۹). معمولاً از یک روش آماری برای ارزیابی بهبود نسبی طبقه‌بندی‌کننده نسبت به نمونه‌های واقعی زمین و کلاس‌های پوشش زمین استفاده می‌شود (سجاد، ۲۰۲۰).

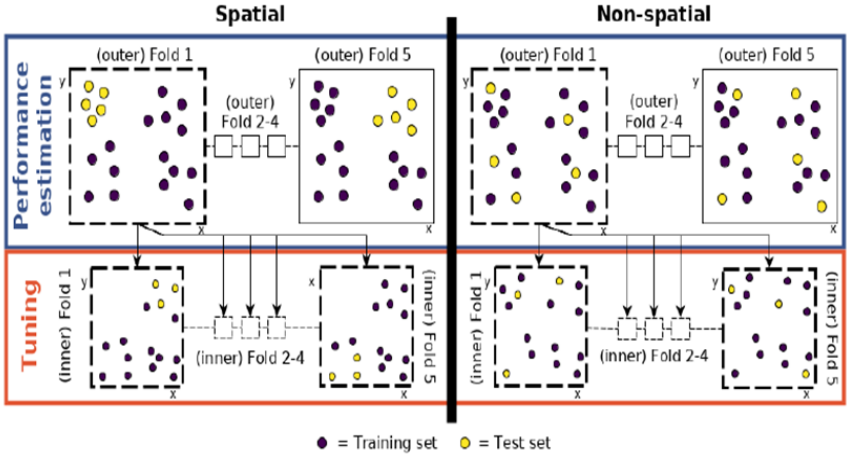
ماتریس سردرگمی برای ارزیابی دقت طبقه‌بندی کاربری/پوشش زمین استفاده می‌شود که بدین منظور نتایج حاصل از نقشه های خروجی با داده‌های آزمایشی مقایسه گردید. همچنین از ضریب کاپا[4] به‌عنوان شاخصی برای تأیید توافق کیفی، مثبت یا منفی، بین نمونه‌های طبقه‌بندی‌شده و نقاط آزمایشی (مرجع) استفاده شد (کوهن[5]، ۱۹۶۸).

الف- بدست آوردن پارامترهای بهینه الگوریتم‌های یادگیری ماشین به روش ارزیابی متقابل مکانی

هر الگوریتم یادگیری ماشینی مراحل تنظیم و پارامترهای تنظیم متفاوتی دارد (نوی و کاپاس[6]، ۲۰۱۸)، که تنظیم آن‌ها یکی از مراحل کلیدی طبقه‌بندی برای بهبود دقت است (باقری و همکاران، ۲۰۱۹). در صورت وجود همبستگی مکانی بین نمونه‌های آموزشی و آزمایشی الگوریتم یادگیری ماشین، ممکن است برآوردهای به دست آمده با استفاده از روش اعتبارسنجی متقابل[7] بر اساس قانون اول جغرافیا- نقاط نزدیک به یکدیگر شباهت بیشتری با نقاط دورتر دارند- خوشبینانه و در حالت بیش برازش[8] باشند (میلر[9]، ۲۰۰۴). لذا در این پژوهش ملزم به اتخاذ روش ارزیابی متقاطع مکانی با ایده تقسیم‌بندی مکانی برابر نمونه‌ها به دسته‌های بر اساس مختصات مکانی آن‌ها برای کاهش همبستگی بین نمونه‌های آموزش و آزمایش شده است (برنینگ[10]، ۲۰۱۲). این تقسیم‌بندی‌ها با حالت‌های مختلف تکرار می‌شوند تا تضمین کنند که هر نمونه حداقل یک بار در یکی ازدسته‌های آزمایش موردارزیابی قرار گرفته باشد (شراتز[11] و همکاران، ۲۰۱۸). شکل ۳ نحوه‌ی عملکرد روش ارزیابی متقابل مکانی را نمایش می‌دهد.

[1] - Brink
[2] - Google Earth
[3] - Olofsson
[4] - Kappa Coefficient
[5] - Cohen
[6] - Noi & Kappas
[7] - Cross-Validatio
[8] - Over Fit
[9] - Miller
[10] - Brenning
[11] - Schratz

[Downloaded from system.khu.ac.ir on 2025-12-05]
[DOI: 10.61186/jsaeh.11.3.1]



شکل ۳. نحوه عملکرد روش ارزیابی متقابل مکانی (Schratz et al., 2018)

در این پژوهش برای آموزش الگوریتم‌های شبکه عصبی پرسپترون چندلایه، ماشین بردار پشتیبان و جنگل تصادفی جهت طبقه‌بندی تصاویر ماهواره‌ای از نرم افزار آماری متن باز R و بسته mlr (بیشل[۱] و همکاران، ۲۰۱۶) استفاده شد.۷۰ درصد از کل نمونه‌های جمع آوری شده در تصاویر پیش از وقوع سیلاب بعنوان نمونه آموزشی بصورت تصادفی انتخاب گردید، و سپس از روش اعتبارسنجی متقابل مکانی به تعداد ۵ دسته و با جستجوی تصادفی[۲] ۵۰۰ تکرار برای یافتن پارامترهای بهینه هر یک از الگوریتم‌ها استفاده شد. جستجوی تصادفی دارای ویژگی مناسب در ابعاد بالا است و در مقایسه با جستجوی شبکه[۳]، از دقت مناسبی برخوردار است (برگسترا و بنژیو[۴]، ۲۰۱۲). تعداد تکرار جستوجوی تصادفی و دامنه مقادیر پارامترها یکی از عوامل مهم در کارایی الگوریتم‌های یادگیری ماشین در برآورد مقدار بهینه می‌باشد، که این مقادیر بر اساس تحقیقات (لاولیس[۵] و همکاران، ۲۰۱۹؛ شراتز[۶] و همکاران، ۲۰۱۸؛ بیشل[۷] و همکاران، ۲۰۱۶) نتخاب شدند. همچنین در این پژوهش برای مقایسه دقت الگوریتم ها از دقت کلی به عنوان معیار ارزیابی استفاده کردیم (غیور و همکاران، ۲۲۰۲۱). در نهایت نقشه‌های پوشش اراضی برای دوره قبل از وقوع سیلاب، بر اساس الگوریتم‌های آموزش داده شده با دو حالت (به‌همراه شاخص‌های طیفی به عنوان ویژگی و بدون شاخص‌های طیفی) جهت بررسی افزایش دقت تولید گردید.

۲-۴. تهیه نقشه پهنه‌های سیلابی

تکنیک‌های مختلفی جهت طبقه‌بندی بر روی تصاویر مناطق سیل‌زده برای تعیین بهترین عملکرد در شناسایی پیکسل‌های سیلاب استفاده می‌شود (موناسینگ[۸]، ۲۰۱۸) و در این میان، طبقه‌بندی نظارت شده، روشی قابل اعتماد جهت طبقه‌بندی مناطق آبی و غیر آبی است (سجاد و همکاران، ۲۰۲۰؛ الفیصل[۹] و همکاران، ۲۰۱۸؛ خالید[۱۰] و همکاران، ۲۰۱۸). در این پژوهش جهت تولید نقشه پهنه‌های سیلابی از تصاویر ماهواره‌ای سنتینل ۲ در تاریخ ۹۸/۱/۱۶ استفاده گردید، در این بخش نیز مانند تولید نقشه پوشش اراضی قبل از سیلاب از ۰/۲ درصد پیکسل‌های تصویر به عنوان نمونه‌های آموزشی و آزمایشی استفاده شد. ۷۰ درصد از کل نمونه‌ها در مناطق پهنه‌های آبی و غیر آبی، برای آموزش الگوریتم طبقه‌بندی کننده دوکلاسه، بصورت تصادفی انتخاب شدند

[1] - Bischl
[2] - Random Search
[3] - Grid Search
[4] - Bergstra & Bengio
[5] - Lovelace
[6] - Schratz
[7] - Bischl
[8] - Munasinghe
[9] - AlFaisal
[10] - Khalid

[Downloaded from system.khu.ac.ir on 2025-12-05]
[DOI: 10.61186/jsaeh.11.3.1]

و بقیه نمونه‌ها جهت ارزیابی دقت مورد استفاده قرار گرفت. بهترین الگوریتم مورد استفاده جهت طبقه‌بندی مناطق سیلابی و غیر سیلابی، از بررسی دقت الگوریتم‌های مذکور به روش ارزیابی متقابل مکانی انتخاب گردید. برای بررسی و ارزیابی تأثیر شاخص طیفی در بهبود دقت نقشه پهنه‌های سیلابی، الگوریتم داده‌ها را یکبار بدون شاخص با ۵ باند طیفی و بار دیگر به همراه شاخص طیفی MNDWI به دو کلاس اراضی سیلابی و غیر سیلابی طبقه‌بندی کرده و در نهایت جهت نمایش و محاسبه مساحت سیل‌زده در محدوده مورد مطالعه و بررسی و آشکارسازی پهنه تحت خسارت سیلاب در مناطق ساخته شده و اراضی زراعتی و فضاهای سبز، نقشه پوشش اراضی به عنوان مرجع و نقشه پهنه‌های سیلابی روی هم قرار گرفته و به صورت سلول به سلول مقایسه شدند.

یافته‌های پژوهش

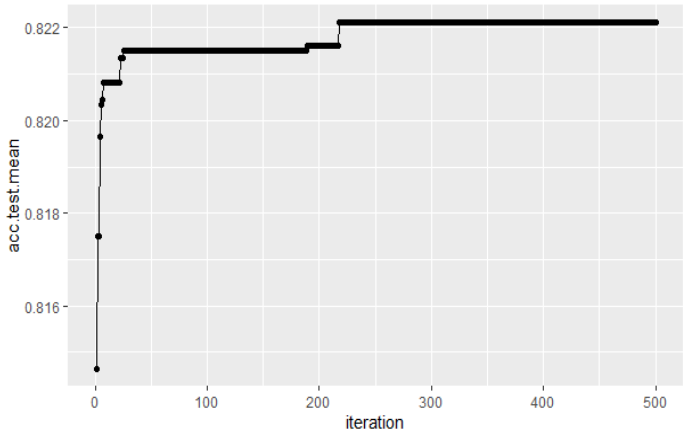
۱. ارزیابی دقت الگوریتم‌های یادگیری ماشین

مقادیر بهینه پارامترهای مدل‌های یادگیری ماشین و دقت کلی برای الگوریتم‌های یادگیری ماشین با ترکیب شاخص‌های طیفی و بدون ترکیب شاخص طیفی در این مطالعه در جدول ۲ ارائه گردیده است.

جدول ۲. مقادیر بهینه پارامترهای الگوریتم‌ها با ترکیب و بدون ترکیب شاخص‌های طیفی

الگوریتم	پارامترها	دامنه مقادیر	مقدار بهینه با ترکیب	مقدار بهینه بدون ترکیب	دقت کلی با ترکیب	دقت کلی بدون ترکیب
جنگل تصادفی	mtry	۰ تا ۸	۱	۳	۸۳/۸۰	۸۲/۲۱
	Number of trees	۰ تا ۵۰۰۰	۸۵۱	۷۰۳		
ماشین بردار پشتیبان	C	2^{-12} تا 2^{15}	۳۶۷	۳۶۷	۷۹/۱۱	۷۸/۹۶
	σ	2^{-15} تا 2^{6}	۰/۰۰۰۲۳۸	۰/۰۰۰۲۳۸		
شبکه عصبی	Number of neurons	۰ تا ۱۰۰۰	۳۴	۳۳	۷۵/۴۴	۷۳/۵۸

بر اساس نتایج بدست آمده بالاترین مقادیر دقت کلی در بهینه‌ترین حالت برای الگوریتم‌های جنگل تصادفی، ماشین‌بردار پشتیبان و شبکه عصبی چند لایه با ترکیب شاخص‌های طیفی برابر ۸۳/۰۸ درصد، ۷۹/۱۱ درصد و ۷۵/۴۴ درصد بدست آمد که نشان دهنده دقت عالی برای الگوریتم جنگل تصادفی و بسیار خوب برای الگوریتم‌های ماشین بردار پشتیبان و شبکه عصبی پرسپترون چندلایه است. همچنین برای داده‌های بدون ترکیب شاخص‌های طیفی مذکور (پنج باند طیفی) الگوریتم جنگل تصادفی با دقت کلی ۸۲/۲۱ درصد در جایگاه اول، الگوریتم ماشین بردار پشتیبان با ۷۸/۹۶ درصد در رتبه دوم قرار گرفت و در نهایت الگوریتم شبکه عصبی چند لایه ضعیف‌ترین عملکرد را در بین سایر الگویتم با دقت کلی ۷۳/۵۸ کسب کرد. شکل ۴ نمودار دقت کلی اجرای[1] الگوریتم جنگل تصادفی بهمراه شاخص های طیفی را در ۵۰۰ تکرار نشان می‌دهد.



شکل ٤. نمودار دقت کلی الگوریتم جنگل تصادفی با ترکیب شاخص‌های طیفی

[1] - Run

[Downloaded from system.khu.ac.ir on 2025-12-05]

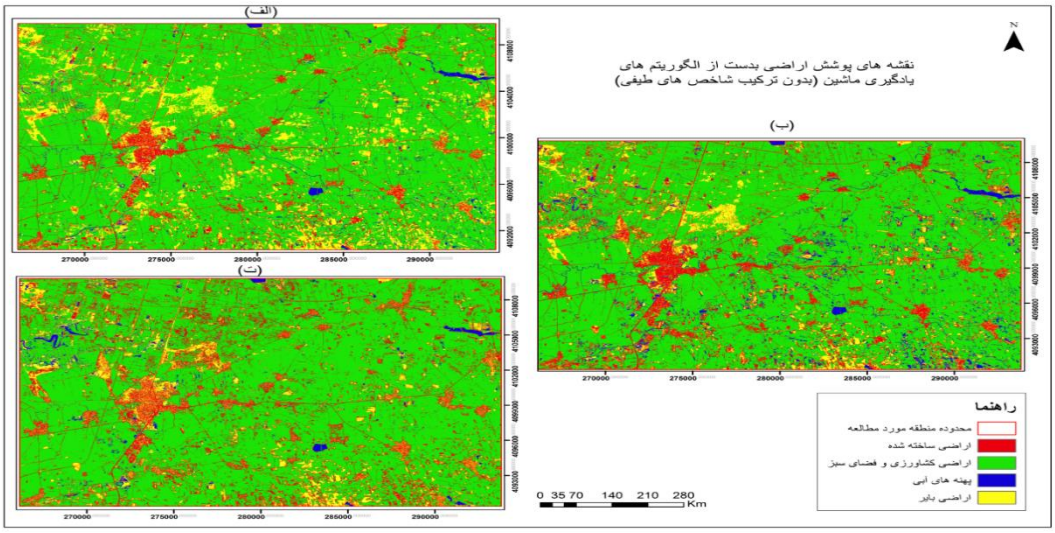
[DOI: 10.61186/jsaeh.11.3.1]

این نتایج نشان می‌دهد الگوریتم جنگل تصادفی تعمیم‌پذیری بالاتر و حساسیت کمتری به تنظیم ابرپارامترها نسبت به سایر الگوریتم‌ها در روش اعتبارسنجی متقابل مکانی دارد. به طور کلی، الگوریتم جنگل تصادفی دقت بهتری در تولید نقشه‌های پوشش اراضی برای هر دو مجموعه داده (به همراه و بدون ترکیب شاخص‌های طیفی) ارائه می‌دهد. این الگوریتم از روش یادگیری جمعی برای تفکیک کلاس‌ها استفاده می‌کند که منجر به کاهش واریانس می‌گردد، بدین دلیل یکی از دقیق‌ترین الگوریتم‌های یادگیری ماشین می‌باشد.

جهت بررسی و ارزیابی تفاوت روش ارزیابی متقابل معمولی و مکانی در بدست آوردن پارامترهای بهینه، الگوریتم جنگل تصادفی در شرایط یکسان با روش ارزیابی متقابل مکانی بصورت معمولی اجرا شد که مقادیر دقت کلی ۹۴/۸۸ درصد بدست آمد. دلیل این امر این است که در روش ارزیابی متقابل معمولی به علت توزیع تصادفی نمونه‌های آموزشی و آزمایشی و نزدیکی قرارگیری این نمونه‌ها به یکدیگر، به علت همبستگی مکانی باعث بیش برازش دقت الگوریتم‌ها می‌شوند.

۲. ارزیابی نقشه‌های پوشش اراضی

پس از محاسبه ماتریس سردرگمی، از دقت کاربر برای ارزیابی تمایز کلاس‌ها استفاده شد و همچنین دقت کلی و شاخص کاپا برای هر یک از نقشه‌ها تولید گردید. شکل ۵ به ترتیب نقشه پوشش اراضی تولید شده توسط الگوریتم‌های جنگل تصادفی و ماشین بردار پشتیبان و شبکه عصبی مصنوعی بدون ترکیب شاخص‌های طیفی را نشان می‌دهد. در جدول ۳ دقت کلی و ضریب کاپای محاسبه شده برای هر یک از نقشه‌های پوشش اراضی بدون ترکیب شاخص‌های طیفی نشان داده شده است.



شکل ۵. نقشه‌های پوشش اراضی بدست آمده از الگوریتم‌های جنگل تصادفی (الف)، ماشین بردار پشتیبان (ب) و شبکه عصبی چند لایه (ت) بدون ترکیب شاخص‌های طیفی

جدول۳. مشخصات نقشه‌های پوشش اراضی حاصل از طبقه‌بندی تصاویر ماهواره‌ای بدون ترکیب شاخص‌های طیفی

کلاس‌های پوشش اراضی	الگوریتم جنگل تصادفی		الگوریتم ماشین بردار پشتیبان		الگوریتم شبکه عصبی (MLP)	
	مساحت (km^2)	دقت کاربر[1]	مساحت (km^2)	دقت کاربر	مساحت (km^2)	دقت کاربر
اراضی ساخته شده	۹۶۶۰/۴۸	۸۳/۶۰	۶۷/۳۱۵۹	۸۳/۰۶	۷۶/۹۹۹۸	۶۴/۰۰
اراضی زراعی فضای سبز	۴۱۱/۳۰۴۵	۹۱/۸۶	۴۰۹/۸۵۱۶	۸۵/۲۰	۴۱۲/۹۴۵۷	۹۶/۰۰
پهنه‌های آبی	۱۴/۲۱۸۶	۹۳/۳۳	۲۳/۶۴۸۹	۹۲/۹۳	۱۲/۰۲۶۹	۶۵/۳۳
اراضی بایر	۶۵/۴۶۵۹	۶۰/۰۰	۳۹/۱۳۸۶	۵۷/۶۰	۳۷/۹۸۲۶	۵۷/۸۶
دقت کلی[2]	۸۲/۲۰		۷۹/۷۰		۷۰/۸۰	
ضریب کاپا	۷۷/۳۰		۷۴/۶۵		۶۴/۵۷	

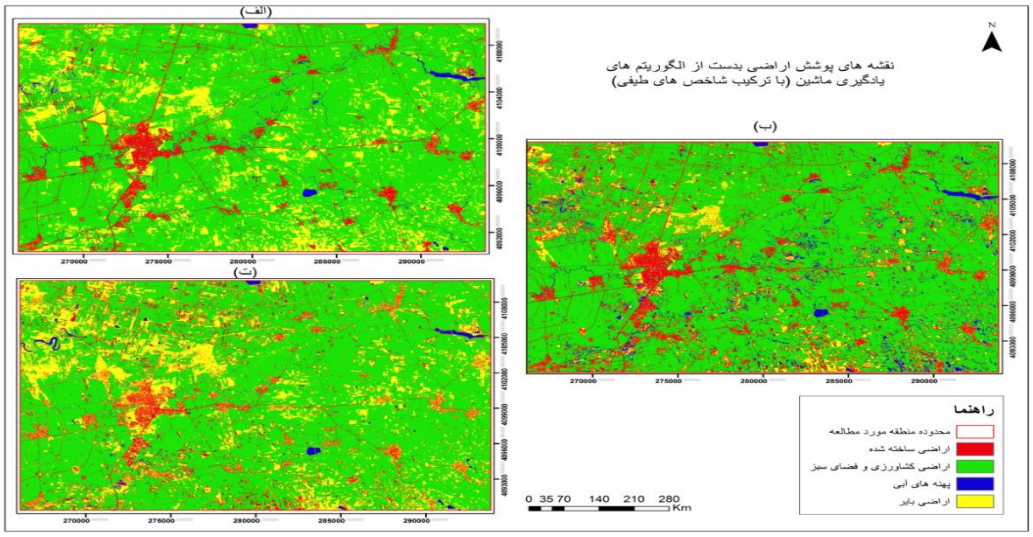
[1] - User accuracy

[2] - Overall accuracy

[Downloaded from system.khu.ac.ir on 2025-12-05]

[DOI: 10.61186/jsaeh.11.3.1]

همچنین شکل ۶ به ترتیب نقشه پوشش اراضی تولیدشده توسط الگوریتم‌های جنگل تصادفی و ماشین بردار پشتیبان و شبکه عصبی مصنوعی با ترکیب شاخص‌های طیفی را نشان می‌دهد. در جدول ۴ نیز دقت کلی و ضریب کاپای محاسبه شده برای هر یک از نقشه‌های پوشش اراضی با ترکیب شاخص‌های طیفی نشان داده شده است.



شکل ٦. نقشه‌های پوشش اراضی بدست آمده از الگوریتم‌های جنگل تصادفی (الف)، ماشین بردار پشتیبان (ب) و شبکه عصبی چند لایه (ت) با ترکیب شاخص‌های طیفی

جدول۴. مشخصات مربوط به نقشه‌های پوشش اراضی حاصل از طبقه‌بندی تصاویر ماهواره‌ای با ترکیب شاخص‌های طیفی

کلاس‌های پوشش اراضی	الگوریتم جنگل تصادفی		الگوریتم ماشین بردار پشتیبان		الگوریتم شبکه عصبی (MLP)	
	مساحت (km^2)	دقت کاربر	مساحت (km^2)	دقت کاربر	مساحت (km^2)	دقت کاربر
اراضی ساخته شده	۳۵/۴۱۸۵	۸۴/۱۳	۶۷/۲۰۴۳	۸۳/۶۰	۴۵/۸۰۲۴	۵۶/۸۰
اراضی زراعی و فضای سبز	۴۱۱/۷۷۵۱	۹۰/۶۶	۴۰۹/۵۹۸۳	۸۵/۲۰	۳۸۴/۲۹۸۸	۹۴/۸۰
پهنه‌های آبی	۶/۳۹۵۰	۹۶/۰۰	۲۳/۸۷۶۷	۹۴/۲۶	۱۰/۹۶۴۷	۷۰/۸۰
اراضی بایر	۸۵/۶۱۷۷	۸۲/۸۰	۳۹/۲۳۹۷	۵۹/۰۶	۹۸/۸۹۱	۸۳/۲۰
دقت کلی	۸۸/۴۰		۸۰/۵۳		۷۶/۴۰	
ضریب کاپا	۸۲/۳۸		۷۵/۵۲		۶۹/۱۲	

نتایج نشان می دهد که نقشه پوشش اراضی تهیه شده توسط الگوریتم جنگل تصادفی با ضریب کاپای ۷۷/۳۰ درصد، بالاترین دقت را نسبت به نقشه‌های پوشش اراضی تهیه شده با استفاده از الگوریتم‌های ماشین بردار پشتیان و شبکه عصبی مصنوعی با ضریب کاپای به ترتیب ۷۴/۶۵ و ۶۴/۵۷ درصد دارد. همچنین درخصوص دقت نقشه‌های پوشش اراضی تولید شده توسط الگوریتم‌های یادگیری ماشین (جدول۴)، الگوریتم جنگل تصادفی با ضریب کاپای ۸۲/۳۸ درصد، بالاترین دقت را در تهیه نقشه پوشش اراضی با ترکیب شاخص‌های طیفی مذکور دارد و الگوریتم‌های ماشین بردار پشتیبان و شبکه عصبی مصنوعی چند لایه با ضریب کاپای به ترتیب ۷۵/۵۲ و ۶۹/۱۲ درصد، همانند نقشه‌های پوشش اراضی بدون ترکیب شاخص‌های طیفی در رتبه‌های بعدی دقت قرار دارد. بطور کلی رابطه مثبت و معنی داری بین دقت الگوریتم‌های یادگیری ماشین و ضریب کاپای تولید شده از نقشه‌های پوشش اراضی حاصل از این الگوریتم‌ها وجود دارد.

نقشه پوشش اراضی که توسط الگوریتم جنگل تصادفی بر اساس تصویر سنتینل۲ تهیه شده است، طبقه‌بندی پوشش اراضی را با دقت بالایی به تصویر می‌کشد (شکل۴، الف)، که بالاترین دقت کاربر مربوط به پهنه‌های آبی با ۹۶/۰۰ درصد و کمترین مطابقت با واقعیت زمینی مربوط به کلاس اراضی بایر با۸۲/۸۰ درصد می‌باشد. با این حال، نقشه پوشش اراضی تولید شده با استفاده از الگوریتم شبکه عصبی مصنوعی چندلایه در شناسایی اراضی ساخته شده با دقت کاربر ۵۶/۸۰ درصد بسیار ضعیف عمل

[Downloaded from system.khu.ac.ir on 2025-12-05]
[DOI: 10.61186/jsaeh.11.3.1]

کرده است. به دلیل شباهت طیفی بالای اراضی ساخته شده با سایر کاربریها از قبیل زمین برهنه، تمایز آن‌ها یکی از چالش‌برانگیزترین بخش‌های پردازش است که در این بین الگوریتم جنگل تصادفی با دقت کاربر ۸۴/۱۳ درصد دقت قابل قبولی کسب کرده است و نسبت به نقشه بدون ترکیب شاخص‌های طیفی، ۰/۵۳ درصد بهبود دقت داشته است. نتایج نشان می‌دهند که اگرچه روش‌های پیشرفته‌تر یادگیری ماشین مانند الگوریتم جنگل تصادفی، طبقه‌بندی دقیق‌تری را در خصوص کلاس‌های مختلف ارائه می‌دهد، اما شبکه عصبی مصنوعی درخصوص طبقه‌بندی تصاویر در طیف‌های خاص مانند اراضی زراعتی و فضای-های سبز عملکرد بهتری دارد. مشاهدات نشان می‌دهد تا زمانی که عملکرد یک الگوریتم با روش‌های مشابه مقایسه نشود، نمی‌توان در مورد انتخاب الگوریتم اطمینان حاصل کرد؛ بنابراین اجرای همزمان چندین روش و مقایسه نتایج آن‌ها دقیق‌ترین نقشه را ایجاد می‌کند.

شایان ذکر است که نتایج این بخش از مطالعه با خروجی مطالعات مشابه مورد مقایسه قرار گرفت. به عنوان نمونه آدام[1] و همکاران (۲۰۱۴) از دو الگوریتم جنگل تصادفی و ماشین بردار پشتیبان به منظور تولید نقشه پوشش اراضی یک منطقه در سواحل شرق آفریقا استفاده کردند و دریافتند که الگوریتم جنگل تصادفی با دقت کلی ۹۳/۰۷ درصد در مقایسه با ماشین بردار پشتیبان با دقت کلی ۹۱/۸۰ درصد دقیق‌تر است. در مطالعه دیگری کانتاکومار و نیلمستتی[2] (۲۰۱۵) از سه الگوریتم شبکه عصبی مصنوعی، ماشین بردار پشتیبان و نقشه‌بردار زاویهٔ طیفی[3] برای طبقه‌بندی در واراناسی[4] هند استفاده کردند و نتایج نشان داد که الگوریتم‌های ماشین بردار پشتیبان با دقت ۴۵/۹۳ درصد و شبکه عصبی مصنوعی با دقت ۳۲/۹۲ درصد بهتر از الگوریتم نقشه-بردار زاویهٔ طیفی در طبقه‌بندی منطقه مورد مطالعه عمل کرد. در پژوهش محمدعبدی و همکاران (۲۰۱۹) مشخص گردید الگوریتم جنگل تصادفی نسبت به الگوریتم ماشین بردار پشتیبان دقیق‌تر و الگوریتم قابل اعتماد در میان الگوریتم‌های یادگیری ماشین است. همچنین غیور و همکاران (۲۰۲۱) نیز در پژوهش خود به این نتیجه رسیدند که الگوریتم ماشین بردار پشتیبان نسبت به الگوریتم شبکه عصبی مصنوعی در طبقه‌بندی تصاویر ماهواره‌ای دقیق‌تر است. نتایج این مطالعه نشان داد که در حالی که از مجموعه نمونه‌های آموزشی و آزمایشی یکسان استفاده می‌شود، ترکیب شاخص‌های طیفی با باندهای تصاویر بر نتایج نهایی تأثیر می‌گذارد و بنابراین یکی از مراحل کلیدی در تولید نقشه‌های پوشش اراضی است. علاوه بر این تنظیم پارامترهای هر کدام از الگوریتم‌ها دقت نقشه تولید شده را افزایش می‌دهد.

۳. بهبود نتایج الگوریتم جنگل تصادفی با استفاده از شاخص طیفی

پس از تشخیص الگوریتم جنگل تصادفی به عنوان دقیق‌ترین الگوریتم به منظور تهیه نقشه‌های پوشش اراضی و تشخیص کلاس‌های طبقه‌بندی، از آن در تهیه نقشه پهنه‌های سیلابی منطقه آق قلا در تاریخ ۱۳۹۸/۱/۱۶ با استفاده از تصویر سنتینل ۲ استفاده گردید. برای بررسی تأثیر شاخص طیفی در افزایش دقت تفکیک کلاس‌ها، همانند شرایط تولید نقشه پوشش اراضی اقدام شد. یکبار داده‌ها با ۵ باند طیفی تصویر ماهواره‌ای اجرا شدند و بار دوم جهت بررسی افزایش دقت از ترکیب شاخص طیفی (MNDWI) استفاده گردید. جدول ۵ دقیق‌ترین پارامترهای بهینه الگوریتم جنگل تصادفی را در۵۰۰ تکرار و جدول ۶ مشخصات کلاس‌ها، دقت کلی و ضریب کاپای نقشه پهنه‌های سیلابی را در هر دو حالت نمایش می‌دهد.

جدول ۵. مقادیر بهینه پارامترها برای الگوریتم جنگل تصادفی

الگوریتم جنگل تصادفی	پارامترها	دامنه مقادیر	مقدار بهینه
با ترکیب شاخص طیفی (MNDWI)	Mtry	۰ تا ۶	۲
	Number of trees	۰ تا ۵۰۰۰	۱۸۴
بدون ترکیب شاخص طیفی (MNDWI)	Mtry	۰ تا ۵	۲
	Number of trees	۰ تا ۵۰۰۰	۱۷۶

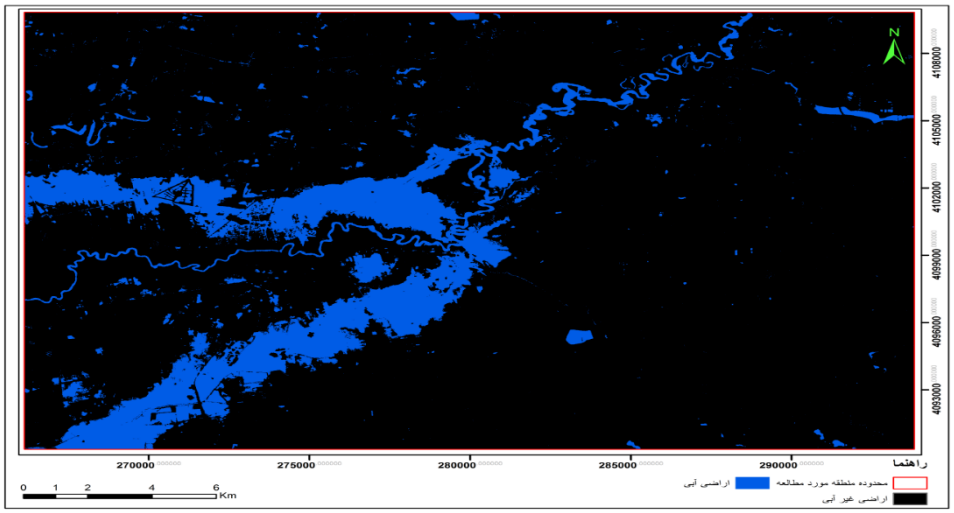
[1] - Adam
[2] - Kantakumar and Neelamsetti
[3] - Spectral Angle Mapper (SAM)
[4] - Varanasi

[Downloaded from system.khu.ac.ir on 2025-12-05]
[DOI: 10.61186/jsaeh.11.3.1]

جدول٦. مشخصات مربوط به نقشه پهنه‌های سیلاب حاصل از طبقه‌بندی تصاویر ماهواره‌ای

الگوریتم جنگل تصادفی	با ترکیب شاخص طیفی		بدون ترکیب شاخص طیفی	
	مساحت (km^2)	دقت کاربر	مساحت (km^2)	دقت کاربر
مناطق آبی	۶۴/۹۵۰۶	۹۶/۸۰	۶۴/۸۷۹۸	۹۵/۹۳
مناطق غیر آبی	۴۷۵/۰۰۴۴	۹۰/۰۰	۴۷۵/۰۷۵۲	۸۹/۶۰
دقت کلی	۹۳/۴۰		۹۲/۷۶	
ضریب کاپا	۹۱/۳۰		۹۰/۴۶	

بر اساس نتایج، دقت کلی و ضریب کاپای نقشه پهنه‌های سیلابی حاصل از الگوریتم جنگل تصادفی، بالاتر از ۹۰ درصد به دست آمد (جدول ۶). بالاترین دقت کلی۹۳/۴۰ درصد، مربوط به ترکیب شاخص طیفی MNDWI با باندهای تصویر اوج دبی سیلاب و کمترین دقت کلی ۹۲/۷۶ درصد، بدون ترکیب شاخص طیفی MNDWI بدست آمده که ۰/۶۴ درصد دقت کلی را بهبود داده است و مناطق آبی نسبت به غیر آبی را با افزایش دقت کاربری ۰/۸۷درصد دقیق‌تر تشخیص داده است. به طور مشابه، میانگین دقت کاربر، دقت کلی و ضریب کاپا در هر دو نقشه بالای ۹۰ درصد است، که نشان می‌دهد طبقه‌بندی کننده باینری نسبت به طبقه‌بندی چند کلاسه دقیق‌تر عمل کرده است. ضریب کاپای طبقه‌بندی نشان می‌دهد که تصاویر طبقه‌بندی شده برای تجزیه و تحلیل قابل اعتماد هستند. شکل ۷ نقشه پهنه‌های سیلابی تولید شده توسط الگوریتم جنگل تصادفی را با ترکیب شاخص طیفی MNDWI نمایش می‌دهد.



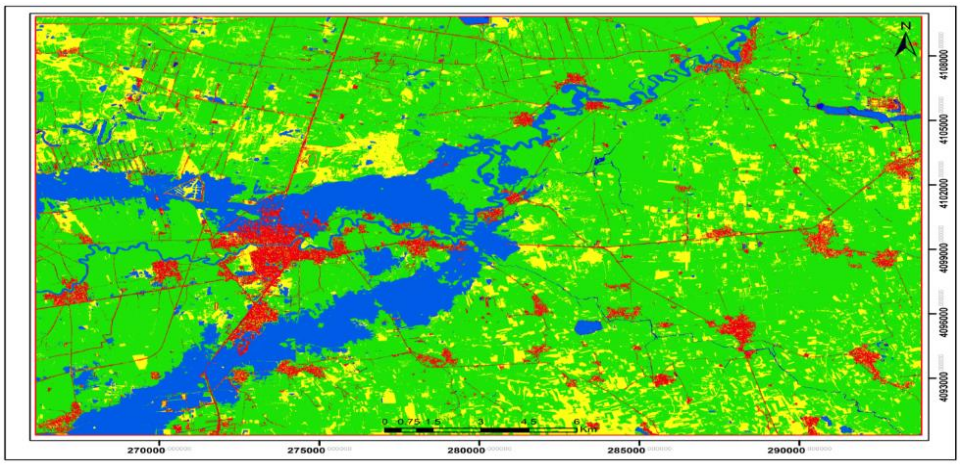
شکل ۷. نقشه پهنه‌های سیلابی تولید شده توسط الگوریتم جنگل تصادفی را با ترکیب شاخص طیفی MNDWI

در نقشه پهنه‌های سیلابی با ترکیب شاخص طیفی MNDWI مساحت سطح سیلابی را ۶۴/۹۵۰۶ کیلومتر مربع نشان می دهد، در حالی که در تصویر طبقه‌بندی شده بدون ترکیب شاخص طیفی MNDWI مساحت ۶۴/۸۷۹۸ کیلومتر مربع از منطقه زیر آب رفته است که نشان دهنده تغییرات اندکی می‌باشد. نتایج نشان می‌دهد که هم با ترکیب شاخص طیفی MNDWI و هم بدون ترکیب آن، مساحت مناطق سیلابی تقریباً مشابه بوده است، با این حال ترکیب شاخص طیفی باعث افزایش جزئی دقت طبقه‌بندی شده است.

۴. محاسبه مساحت اراضی تحت خسارت سیلاب

همپوشانی نقشه طبقه‌بندی پوشش اراضی با پهنه‌های سیلابی گام اساسی در برآورد خسارات ناشی از سیلاب است که در این پژوهش به آن پرداخته شده و می‌تواند در برآورد هزینه‌های اقتصادی و اجتماعی ناشی از سیلاب مورد استفاده قرار گیرد. در این مطالعه همپوشانی نقشه طبقه‌بندی پوشش اراضی با پهنه‌های سیلابی در شکل ۸ ارائه گردیده است.

[Downloaded from system.khu.ac.ir on 2025-12-05]
[DOI: 10.61186/jsaeh.11.3.1]



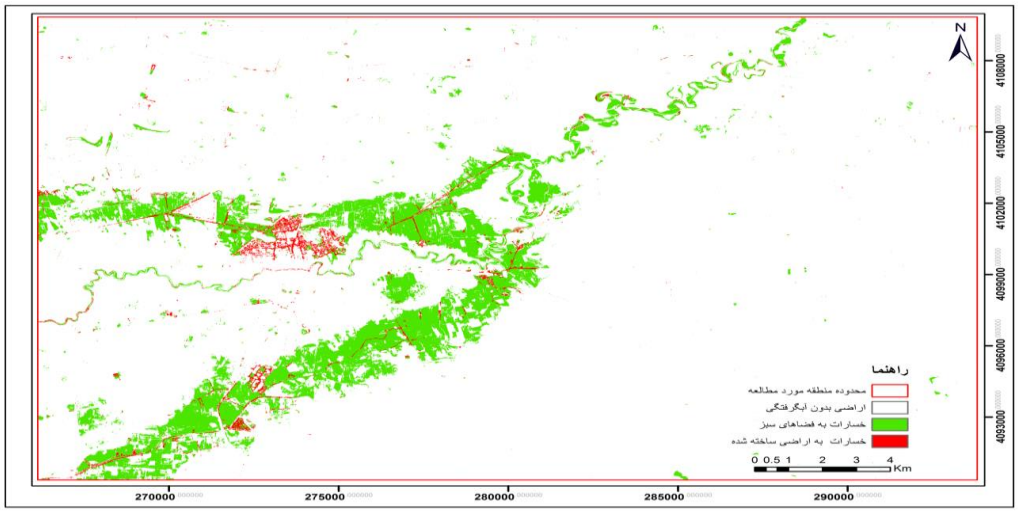
شکل ۸. نقشه پوشش سیلاب

همانطور که در شکل ۸ نشان داده شده حجم آب در شرایط پیش از سیلاب تنها ۱/۱۸ درصد از کل منطقه را به خود اختصاص داده است، با این حال، در اوج دبی سیلاب، یک تغییر ناگهانی در پهنه‌های آبی ظاهر شده که نشان دهنده افزایش گسترده به ۱۲/۰۴درصد از کل مساحت منطقه است. نتایج نشان می‌دهد که آبگرفتگی قابل ملاحظه‌ای در تمام کاربری‌های طبقه‌بندی شده رخ داده و سیلاب باعث آبگرفتگی ۴/۲۰۰۸ کیلومتر مربع از اراضی ساخته شده، ۴۱/۰۷۷۲ کیلومتر مربع اراضی پوشش گیاهی و کشاورزی و ۱۵/۹۸۷۵کیلومتر مترمربع از زمین‌های بایر شده است. بنابراین، سیلاب باعث گردیده که بخش بزرگی از مناطق کشاورزی و پوشش گیاهی زیر آب رفته و محصولات کشاورزی با آسیب جدی مواجه گردند. علاوه بر این مناطق شهری، روستایی و راه‌های ارتباطی موجود در منطقه مورد مطالعه درگیر سیلاب و خسارات ناشی از سیلاب گردیده است. جدول ۷ مساحت آبگرفتگی پوشش اراضی بعد از وقوع سیلاب را نمایش می‌دهد.

جدول ۷. مساحت آبگرفتگی پوشش اراضی

پوشش اراضی	ساخته شده	پوشش گیاهی/کشاورزی	آبی	بایر
قبل از وقوع سیلاب (KM)	۳۵/۴۱۸۵	۴۱۱/۷۷۵۱	۶/۳۹۵	۸۵/۶۱۷۷
بعد از وقوع سیلاب(KM)	۳۱/۲۱۷۷	۳۷۰/۶۹۸	۶۱/۲۶۵۴	۶۹/۶۳۰۲

جزئیات بیشتر در مورد زمین‌های کشاورزی آسیب دیده و مناطق ساخته شده درگیر سیلاب در شکل ۹ ارائه گردیده است.



شکل ۹. نقشه خسارات به پوشش اراضی

[Downloaded from system.khu.ac.ir on 2025-12-05]

[DOI: 10.61186/jsaeh.11.3.1]

مناطق شمالی شهر آق‌قلا نسبت به سایر اراضی ساخته شده بیشتر دچار سیلاب شده و همچنین زمین‌های کشاورزی اطراف شهر آق‌قلا از شمال، شرق و جنوب به دلیل توپوگرافی حاکم بر منطقه، به صورت کمانی شکل به زیر آب رفته است.

نتیجه‌گیری و پیشنهادها

در دهه اخیر به دلیل تغییرات اقلیمی و فعالیت‌ها و دستکاری‌های انسان در طبیعت، فراوانی وقوع مخاطره سیلاب افزایش یافته و خسارت‌های جانی و مالی گسترده‌ای را نیز برجای گذاشته است. با توجه به اهمیت ارزیابی خسارات ناشی از سیلاب در مدیریت بحران، آشکارسازی پهنه خسارات ناشی از سیل در کمترین زمان، گامی موثر در بهبود نتایج برنامه‌ریزی بویژه در مراحل پاسخ، بازسازی و بازتوانی و موارد مرتبط با بیمه سیلاب به شمار می‌رود. از این رو در این پژوهش آشکارسازی پهنه‌های سیلابی با استفاده از الگوریتم‌های یادگیری ماشین و سنجش از دور با هدف تعیین مساحت اراضی تحت خسارت سیلاب مدنظر قرار گرفت و با توجه به بارش‌های بهار ۱۳۹۸ و درگیری بسیاری از استان‌ها در سطح کشور با مخاطره سیلاب، شهرستان آق قلا در استان مازندران با توجه به سطح درگیری بالا به عنوان منطقه مورد مطالعه انتخاب گردید. در این بارش‌ها این منطقه با افزایش دبی رودخانه گرگان‌رود از سرشاخه‌های بالادست و افزایش حجم آب بیش از حد ظرفیت جاری رودخانه، با طغیان گسترده آب مواجه گردید.

بدین منظور تصاویر ماهواره‌ای سنتیل ۲ مورد استفاده قرار گرفت. این تصاویر مربوط به نزدیک‌ترین زمان قبل وقوع سیلاب به منظور تولید نقشه‌های پوشش اراضی منطقه مورد مطالعه و همچنین زمان اوج وقوع سیلاب با هدف تولید نقشه پهنه‌های سیلابی می‌باشند. برای تولید و ارزیابی دقت نقشه‌های پوشش اراضی از الگوریتم‌های جنگل تصادفی، الگوریتم ماشین بردار پشتیان و الگوریتم شبکه عصبی مصنوعی بصورت طبقه‌بندی نظارت شده استفاده گردید. همچنین جهت بررسی و افزایش دقت الگوریتم‌ها، سه شاخص طیفی نرمال شده پوشش گیاهی (NDVI)، پهنه آبی (MNDWI) و اراضی ساخته شده (NDBI) با پنج باند طیفی تصاویر سنتینل ۲ ترکیب شدند. پارامترهای تنظیم متفاوت نیز برای هر کدام از الگوریتم‌ها به منظور تعیین تأثیر آن‌ها بر دقت طبقه‌بندی و جلوگیری از بیش برازش مدل‌ها و کسب نتایج خوشبینانه به علت همبستگی مکانی میان نمونه‌های آموزشی و آزمایشی استفاده و این الگوریتم‌ها با روش اعتبارسنجی متقابل مکانی ارزیابی شدند. نتایج حاصل نشان داد الگوریتم جنگل تصادفی با دقت کلی۸۲/۲۱ درصد نسبت به الگوریتم‌های ماشین‌بردار پشتیبان و شبکه عصبی مصنوعی بترتیب با مقادیر۷۸/۹۶ درصد و ۷۳/۵۸ درصد، از دقت و تعمیم‌پذیری بالاتری درخصوص تولید نقشه‌های پوشش اراضی برخوردار است. همچنین اجرای این الگوریتم‌ها با ترکیب شاخص‌های طیفی، منجر به افزایش چشم‌گیر دقت کلی الگوریتم‌ها گردیده و به ترتیب الگوریتم جنگل تصادفی با ۸۳/۰۸ درصد، دقیق‌ترین الگوریتم نسبت به سایر الگوریتم‌های ماشین بردار پشتیبان و شبکه عصبی مصنوعی با مقادیر ۷۹/۱۱ و ۷۵/۴۴ درصد می‌باشد. همچنین دقت مکانی نقشه‌های پوشش اراضی بدست آمده از الگوریتم‌های مذکور، با استفاده از ۳۰ درصد نمونه‌ها محاسبه و شاخص کاپا مورد ارزیابی و مقایسه قرار گرفت. نتایج این شاخص نشان داد رابطه مستقیم بین دقت کلی الگوریتم‌ها و شاخص کاپا وجود دارد. الگوریتم جنگل تصادفی در حالت ترکیب با شاخص‌های طیفی، ضریب ۸۲/۳۸ درصد را نشان داده که نسبت به الگوریتم ماشین بردار پشتیبان با ۷۵/۵۲ درصد و الگوریتم شبکه عصبی با ۶۹/۱۲ درصد از دقت بالای جهت تفکیک کلاس‌های پوشش اراضی برخودار است. پس از مشخص شدن دقیق‌ترین الگوریتم یادگیری ماشین، نقشه پهنه‌های سیلابی با استفاده از الگوریتم جنگل تصادفی در دو کلاس اراضی آبی و غیر آبی تولید گردید، که دقت کلی الگوریتم در بهینه‌ترین پارامترها و با ترکیب شاخص طیفی (MNDWI) ۹۳/۴۰ درصد و دقت مکانی نقشه پهنه های سیلاب با شاخص کاپای ۹۱/۳۰ درصد بدست آمد. با مقاسیه سلول به سلول دقیق‌ترین نقشه پوشش اراضی با نقشه پهنه-های سیلابی، سطح پوشیده شده از سیلاب در اراضی ساخته شده ۴/۲۰۰۸ کیلومتر مربع و در اراضی زراعی و فضاهای سبز ۴۱/۰۷۷۲ کیلومتر مربع محاسبه گردید. بر اساس نتایج بدست آمده نقشه‌های پوشش اراضی و پهنه‌های سیلابی تولید شده نتیجه قابل اعتمادی به منظور تخمین و درک اثرات مخاطره سیلاب ارائه می‌دهند.

ملاحظات اخلاقی

[Downloaded from system.khu.ac.ir on 2025-12-05]

[DOI: 10.61186/jsaeh.11.3.1]

نویسندگان اصول اخلاقی را در انجام و انتشار این پژوهش علمی رعایت نموده‌اند و این موضوع مورد تأیید همه آنهاست.

مشارکت نویسندگان

همه نویسندگان در جمع‌آوری داده‌ها، تهیه گزارش پژوهش و تحلیل داده‌ها مشارکت داشتند.

تعارض منافع

بنا بر اظهار نویسندگان این مقاله تعارض منافع ندارد.

حامی مالی

مقاله حاضر با حمایت مالی پژوهشکده سوانح طبیعی انجام شد.

سپاسگزاری

مقاله مذکور از پروژه تحقیقاتی با عنوان «آشکارسازی پهنه‌های سیلابی با استفاده از الگوریتم‌های یادگیری ماشین و سنجش از دور به منظور تعیین مساحت اراضی تحت خسارت سیلاب و مکان یابی بهینه محل اسکان موقت (مطالعه موردی:سیل فروردین ۹۸ شهرستان آق‌قلا)» استخراج گردیده است. لذا از پژوهشکده سوانح طبیعی به جهت تمامی حمایت‌های مادی و معنوی سپاسگزاری و قدردانی می‌گردد.

منابع

اداره آب منطقه‌ای استان گلستان، آدرس سایت: http://www.gsrw.ir

ایپکی، احسان (۱۳۹۶). ارزیابی و بررسی تغییرات سطحی ناشی از سیل با استفاده از تصاویر ماهواره ای (مطالعه موردی: استان ایلام). *موسسه آموزش عالی عمران و توسعه*، همدان.

بیغم، مریم؛ خیرخواه، میر مسعود؛ و قرمزچشمه، باقر (۱۳۹۹). ارزیابی صحت روش های طبقه بندی تصویر با رویکرد پیکسل مبنا و ماهواره ای سنتینل ۲ شی مبنا در پهنه بندی مناطق سیل گیر رودخانه طالقان. *انجمن ابخیر داری ایران*، ۴، ۱-۱۰.

تراهی، علی اصغر؛ و حسنی، مقدم (۱۳۹۷). تعیین گستره سیل با استفاده از داده‌های سنجنده (OLI مطالعه موردی: سیل سال ۱۳۹۵ دزفول). *محیط زیست و مهندسی آب*، ۱، ۲۴-۳۵.

پوراکرمی، سعید؛ توکلی، محمد؛ محمدی، ذوافقار؛ و قاسم زاده، امیر (۱۳۹۶). استخراج سریع محدوده سیل زده با استفاده از تصاویر SAR. *همایش ملی ژئوماتیک*، تهران.

کاردان، نازیلا؛ حسن زداده، یوسف؛ و ارزنلو، ابوالفضل (۱۳۹۶). شبیه‌سازی دوبعدی جریان‌های سیلابی شهری با مدل CCHE 2D (مطالعه موردی: شهر آق‌قلا). *دریا فنون*، ۴، ۲۵-۳۶.

فاطمی، سید باقر؛ و رضایی، یوسف (۱۳۹۱). *مبانی سنجش از دور*. انتشارات آزاده، ویرایش دوم.

[Downloaded from system.khu.ac.ir on 2025-12-05]

[DOI: 10.61186/jsaeh.11.3.1]

References

Al-Ahmadi, F., and Hames, A. (2009). Comparison of four classification methods to extract land use and land cover from raw satellite images for some remote arid areas, Kingdom of Saudi Arabia. *Earth*, 20, 167–191.

AlFaisal, A., Kafy, A. A., and Roy, S. (2018). Integration of Remote Sensing and GIS Techniques for Flood Monitoring and Damage Assessment: a Case Study of Naogaon District, Bangladesh. *Remote Sensing*, 7, 236.

Adam, E., Mutanga, O., Odindi, J., and Abdel-Rahman, E. M. (2014). Land-use/cover classification in a heterogeneous coastal landscape using RapidEye imagery: Evaluating the performance of random forest and support vector machines classifiers. *Remote Sensing*, 35, 3440–3458.

Bagheri, M., Akbari, A., and Mirbagheri, S. A. (2019). Advanced control of membrane fouling in filtration systems using artificial intelligence and machine learning techniques: A critical review. *Process. Saf. Environ*, 123, 229–252.

Bui, D. T. (2020). A novel ensemble learning based on Bayesian Belief Network coupled with an extreme learning machine for flash flood susceptibility mapping. *Engineering Applications of Artificial Intelligence*, 96, 103971.

Boori, M., Paringer, R., Choudhary, K., and Kupriyanov, A. (2018). Supervised and unsupervised classification for obtaining land use/cover classes from hyperspectral and multi-spectral imagery. *In Proceedings of the Sixth International Conference on Remote Sensing and Geoinformation of the Environment*, Paphos 107730L.

Bischl, B., Lang, M., Kotthoff, L., Schiffner, J., Richter, J., Studerus, G., Casalicchio, G., and Jones, Z. M. (2016). mlr: Machine Learning in R. *The Journal of Machine Learning Research*, 17, 5938-5942.

Bergstra, J., and Bengio, Y. (2012). Random search for hyper-parameter optimization. *Journal of Machine Learning Research*, 13, 281-305.

Brenning, A. (2012). Spatial cross-validation and bootstrap for the assessment of prediction rules in remote sensing: The R package sperrorest. *IEEE international geoscience and remote sensing symposium*, 10, 5372-5375.

Brink, H., Richards, J. W., and Fetherolf, M. (2017). *Real-world Machine Learning*. Shelter Island, NY: Manning.

Breiman, L. (2001). Random forests. *Machine Learning*, 45, 123-140.

Cohen, J. (1968). Weighted kappa: Nominal scale agreement provision for scaled disagreement or partial credit. *Psychological bulletin*, 70, 213.

Ghayour, L., Neshat, A., Paryani, S., Shahabi, H., Shirzadi, A., Chen, W., Al-Ansari, N., Geertsema, M., Amiri, M., Gholamnia, M., Dou. J., and Ahmad, A. (2021). Performance Evaluation of Sentinel-2 and Landsat 8 OLI Data for Land Cover/Use Classification Using a Comparison between Machine Learning Algorithms. *Remote Sensing*, 13, 1-23.

Gislason, P. O., Benediktsson, J. R., and Sveinsson J. A. (2006). Random forests for land cover classification. *Pattern Recognition Letters*. 27, 294-300.

Hastie, T., Tibshirani, R., and Friedman, J. (2009). *Random Forests*. In The Elements of Statistical Learning: Data Mining, Inference, and Prediction, 587–604. New York: Springer New York.

Huang, B., and Boutros, C. (2016). The parameter sensitivity of random forests. *BMC Bioinformatics*, 17, 331.

Huang, M., and Jin, S. (2020). Rapid Flood Mapping and Evaluation with a Supervised Classifier and Change Detection in Shouguang Using Sentinel-1 SAR and Sentinel-2 Optical Data. *Remote sensing*, 12, 2073.

Halgamuge, M. N., and Nirmalathas, A. (2017). Analysis of large flood events: Based on flood data during 1985–2016 in Australia and India. *Int. J. Disaster Risk Reduct*. 24, 1–11.

Islam, S., Tanim, A., and Mullick, A. (2019). Land Use and Land Cover Classification of Coastal Districts of Bangladesh in a 10m Resolution of Sentinel-2 Satellite Image. *Proceedings of International Conference on Planning, Architecture and Civil Engineering, 07 - 09 February 2019, Rajshahi University of Engineering & Technology, Rajshahi, Bangladesh.*

[Downloaded from system.khu.ac.ir on 2025-12-05]

[DOI: 10.61186/jsaeh.11.3.1]

James, G., Witten, D., Hastie, T., and Tibshirani, R. (2013). *An introduction to statistical learning*. New York, Springer.

Kantakumar, L. N., and Neelamsetti, B. (2015). Multi-temporal land use classification using hybrid approach. *Remote Sensing*, 18, 289–295.

Khalid, B., Cholaw, B., Alvim, D.S., Javeed, S., Khan, J.A., Javed, M.A., and Khan, A.H. (2018). Riverine flood assessment in Jhang district in connection with ENSO and summer monsoon rainfall over Upper Indus Basin for 2010. *Nat. Hazards*, 92, 971–993.

Kavzoglu, T., and Mather, P.M. (2003). The use of back-propagating artificial neural networks in land cover classification. *Int. J. Remote Sens*. 24, 4907–4938.

Lovelace, R., Nowosad, J., and Muenchow, J. (2019). Geocomputation with R: Statistical learning. CRC Press, Statistical learning, 1/24-24-24.

Lakshmi, V. ed. (2016). *Remote Sensing of Hydrological Extremes*. Switzerland: Springer International Publishing.

Mohamed Abdi, A. (2019). Land cover and land use classification performance of machine learning algorithms in a boreal landscape using Sentinel-2 data. *GIScience & Remote Sensing*, 57, 1-20.

Miller, J. (2004). *Tobler's First Law and Spatial Analysis*. Annals of the Association of American Geographers.

Munasinghe, D., Cohen, S., Huang, Y.F., Tsang, Y.P., Zhang, J., and Fang, Z.F. (2018). Intercomparison of Satellite Remote Sensing-Based Flood Inundation Mapping Techniques. *Water Resour*, 54, 834–846.

Omer, G., Mutanga, O., Abdel-Rahman, E.M., and Adam, E. (2015). Performance of support vector machines and artificial neural network for mapping endangered tree species usingWorldView-2 data in Dukuduku forest, South Africa. *Remote Sensing*, 8, 4825–4840.

Olofsson, P., Foody, G. M., Herold, M., Stehman, S. V., Woodcock, C. E., and Wulder, M. A. (2014). Good Practices for Estimating Area and Assessing Accuracy of Land Change. *Remote Sensing of Environment*, 148, 42–57.

Rahman, A., and Khan, A.N. (2011). Analysis of flood causes and associated socio-economic damages in the Hindu Kush region. *Nat. Hazards*. 59, 1239-1260.

Roy, D.P., Kovalskyy, V., Zhang, H.K., Vermote, E.F., Yan, L., Kumar, S.S., and Egorov, A. (2016). Characterization of Landsat-7 to Landsat-8 reflective wavelength and normalized di erence vegetation index continuity. *Remote Sensing*, 185, 57–70.

Sajjad, A., Lu, J., Chen, X., Chisenga, C., Saleem, N., and Hassan, H. (2020). Operational Monitoring and Damage Assessment of Riverine Flood-2014 in the Lower Chenab Plain, Punjab, Pakistan, Using Remote Sensing and GIS Techniques. *Remote Sensing*, 12, 714.

Schratz, P., Muenchow, J., Iturritxa, E., Richter, J., and Brenning, A. (2018). Performance Evaluation and Hyperparameter Tuning of Statistical and Machine-Learning Models Using Spatial Data. *Ecological Modelling*, 24, 109-120.

Tavakkoli, S., hasani, M., and Moghaddam, H. (2018). Evaluating the capability of multi temporal remote sensing data in estimation of flood extent. *2nd International congress of geographic information system and remote sensing innovation technologies in monitoring environmental changes*, Tehran, Iran.

Teluguntla, P., Thenkabail, P.S., Oliphant, A., Xiong, J., Gumma, M.K., Congalton, R.G., Yadav, K., and Huete, A. (2018). A 30-m landsatderived cropland extent product of Australia and China using random forest machine learning algorithm on Google Earth Engine cloud computing platform. *ISPRS J. Photogramm. Remote Sensing*, 144, 325–340.

Talukdar, S., Singha, P., Mahato, S., Swades Pal, S., Liou, Y., and Rahman, A. (2020). Land-Use Land-Cover Classification by Machine Learning Classifiers for Satellite Observations—A Review. *Remote Sensing*, 12, 1135.

Yang, Z. S., Qin, X., Zhao. N., and Liang, L. (2017). Mapping of Urban Surface Water Bodies from Sentinel-2 MSI Imagery at 10 m Resolution via NDWI-Based Image Sharpening. *Remote Sensing*, 9, 596.

Xu, H. (2006). Modification of normalised di_erence water index (NDWI) to enhance open water features in remotely sensed imagery. *Remote Sensing*, 27, 3025–3033.

[Downloaded from system.khu.ac.ir on 2025-12-05]

[DOI: 10.61186/jsaeh.11.3.1]

Zeaiean Firoozabadi P., Mousavi, A., Shakiba, A., and Naseri, H. (2003). Flood incident simulation using remote sensing data and automatic cell model (Case study part of Ghaemshahr Talar River Basin). *Geogr*, 1, 125-144.

Zha, Y., Gao, J., and Ni, S. (2003). Use of normalized di_erence built-up index in automatically mapping urban areas from TM imagery. *Int. J. Remote Sens*, 24, 583–594.

Zhang, C., Zhang, C., Sargent, I., Pan, X., Li, H., Gardiner, A., Hare, J. and Atkinson, P.M. (2018). An object-based convolutional neural network (OCNN) for urban land use classification. *Remote Sensing of Environment*, 216, 57–70.

Zhang, J., Huang, Y., Munasinghe, D., Fang, Z., Tsang, Y., and Cohen, S. (2018). Comparative Analysis of Inundation Mapping Approaches for the 2016 Flood in the Brazos River, Texas. *Journal of the American Water Resources Association*.13, 17-35.

[Downloaded from system.khu.ac.ir on 2025-12-05]

[DOI: 10.61186/jsaeh.11.3.1]