International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 13 Issue: 06 | June 2026 www.irjet.net p-ISSN: 2395-0072

Comparative Analysis of different RS indices for Flood Inundation Mapping in Arid Regions

Mostafa Mashal[1,2], Doaa Amin[2], Mona A. Hagras[1], Ashraf M. Elmoustafa[1]

[1]Irrigation and Hydraulics Department, Faculty of Engineering, Ain Shams University, Cairo 11517, Egypt.
[2]Water Resources Research Institute (WRRI), National Water Research Center (NWRC), Ministry of Water Resources and Irrigation (MWRI), Qalyubia 13621, Egypt.

Abstract-*Predicting flood inundation zones is essential for disaster management, especially in arid and semi-arid regions with limited ground observations. This study evaluates three multispectral indices Normalized Difference Water Index (NDWI), Modified Normalized Difference Water Index (MNDWI), and Surface Water Moisture Index (SMI) for flood delineation using Sentinel-2 imagery. Two case studies were analyzed: Wadi El-Allaqi in Egypt and a flood event in Oman on 14–15 April 2024. Satellite data were processed to extract inundation zones and assess index accuracy. SMI achieved the highest performance, with average precision, recall, and F1-score of 0.97, and overall accuracies of 94.66% in Wadi El-Allaqi and 95% in Oman. MNDWI showed high recall (0.97) but lower precision (0.79), with an average accuracy of 88%, while NDWI performed weakest, with precision, recall, and F1-scores of 0.65, 0.50, and 0.54. These results indicate SMI is the most robust index for flood mapping in data-scarce arid regions.*

Keywords: Sentinal-2; flood mapping; water detection; spectral indices; arid regions.

INTRODUCTION

A flash flood is a natural disaster in which a large areas of land like mainstreams or downstream areas of wadi are temporarily inundated by water, there are many different reasons for flooding, it may be due to heavy rainfall, snow melt, dam failures, soil erosion, and/or various human activities. Floods are one of the most devastating natural disasters, causing widespread destruction to property, infrastructure, and loss of lives each year across the globe (Danso-Amoako et al., 2012). They can occur very slowly or in a matter of few minutes. Accurate and timely mapping of flood inundation extent is crucial for effective disaster response, damage assessment, and long-term flood risk mitigation (Mosavi et al., 2018; Smith, 1997), also it is a very useful tool for better understanding the flood regime and contributes to reducing its negative effects (Dinh et al., 2019a). However, in many arid regions, particularly in developing countries, observed flood data and in-situ measurements are often scarce or lacking, hindering the development of reliable flood forecasting and modeling capabilities (Revilla-Romero et al., 2014).

Nowadays, the use of remote sensing technologies has become a very powerful and effective source in predicting flood boundaries, especially in cases of overcoming the gaps in data shortage and unreliability, which provides a very important vision from satellites about the dynamics of surface water and the extent of floods (Klemas, 2015). Because of the ability of satellite images to capture large areas of different fields, satellite images have become an important and essential data source in collecting different information for flood monitoring and mapping (Tsyganskaya et al., 2018), the use of spectral indices obtained from multispectral satellite data has garnered significant attention for its ability to effectively define and monitor flood inundation. This is one of the many remote sensing-based technologies that have earned substantial attention (McFeeters, 1996; Xu, 2006).

In addition to optical remote sensing data, synthetic aperture radar (SAR) satellite imagery has been widely used for flood mapping and monitoring due to its ability to operate independently of cloud cover and solar illumination conditions (Martinis et al., 2009). SAR sensors, such as those onboard the Sentinel-1 mission, are particularly effective for detecting inundated areas because smooth water surfaces produce low backscatter values, enabling clear discrimination between flooded and non-flooded regions ((Twele et al., 2016)). As a result, SAR-based approaches have been extensively applied for near-real-time flood detection, emergency response, and damage assessment, especially during extreme weather events when optical data availability is limited ((Schumann & Di Baldassarre, 2010)). However, the applicability of SAR data for event-based flood analysis remains strongly dependent on the temporal availability of satellite acquisitions coincident with the flood occurrence.

The Modified Normalized Difference Water Index (MNDWI) and the Normalized Difference Water Index (NDWI) are two of the most often used indices for mapping floods and water bodies (Feyisa et al., 2014; Ji et al., 2009), McFeeters (1996) firstly proposed the NDWI, which uses green and near-infrared (NIR) bands to improve the contrast between open water and other land cover features (McFEETERS, 1996), while the MNDWI, introduced by Xu (2006), further improves upon the NDWI by replacing the near infrared (NIR) band with the shortwave infrared (SWIR) band, which is more sensitive to the absorption characteristics of water and better suppresses the influence of built-up land and vegetation (Xu, 2006).

While NDWI and MNDWI have been widely used for flood mapping and monitoring, there is still a research gap in fully understanding their limitations and optimal application for predicting flood extent (Acharya et al., 2018; Rokni et al., 2014) specially in arid regions. Multiple factors such as soil moisture, vegetation cover, and weather conditions can affect the performance and accuracy of these indices, and can lead to over- or underestimation of flood locations and extent (Feyisa et al., 2014; Ji et al., 2009). In addition, studies comparing the relative effectiveness of NDWI, MNDWI, and other indices such as the Surface Water Moisture Index (SMI) in flood monitoring are very limited.

This study aims to address the research gap by evaluating the possibility usage of NDWI, MNDWI, and SMI in predicting flood inundation extent, with concentrating on understanding their limitations and the factors influencing their applicability. The results from this research will aid in creating more effective and reliable flood monitoring and prediction techniques using the available remote sensing data, which can support disaster management efforts and help in improving our understanding of flood dynamics in the locations with limited data.

Study Area

After selecting multiple case studies that experienced significant extreme rainfall over major wadis and resulting flash flood events, the corresponding satellite images of these events were investigated. It was very difficult to find suitable satellite images for the same dates as the flood events, except for the incident in Wadi Al-Allaqi, which was utilized for the analysis. To complement this study, Wadi Hodein, another major wadi in the Eastern Desert that drains directly into the Red Sea, was also considered as a second case study in order to assess different geomorphological and hydrological settings under extreme rainfall events.

Wadi Al-Allaqi basin is considered one of the largest valleys in the Eastern Desert of Egypt, in fact it is one of the greatest dry valleys that descends westward to the Nile River. Its course begins in Sudan at approximately latitude 20°00′ and ends at latitude 23°15′ north, while its basin is bounded between longitudes 32°25′ and 35°15′ east. It flows from the south to the northwest and discharges into Lake Nasser upstream of the High Aswan Dam (HAD), with a main stream length of approximately 400 km. The area of Wadi Al-Allaqi is around 70,000 km^2, extending partly in southern Egypt and partly in northern Sudan (Figure 1). Elevations range from 175 m at the downstream end near Lake Nasser to about 1,200 m in the upper reaches, with flat northern lands and hilly terrain in the eastern part of the basin. The valley is characterized in some parts by clear plant density, particularly near Lake Nasser, due to seasonal flooding and rising water levels in the lake. The area is also notable for migratory birds and small communities engaged in grazing and collection of medicinal and aromatic plants.

The Oman case study represents a distinctly different geomorphological and hydrological setting, as the investigated wadis drain eastward toward the Gulf of Oman (Figure 2). This region is part of the eastern flank of the Arabian Peninsula, characterized by highly rugged mountainous terrain influenced by the Al Hajar Mountains. Elevations in the upstream zones exceed 3,000 m, creating steep gradients that rapidly accelerate runoff generation and concentration during intense rainfall events. The catchments consist of narrow valleys, deeply incised channels, and short flow paths, all of which intensify the likelihood and severity of flash flooding. Their outlets discharge into the Gulf of Oman, where sudden flood pulses can affect coastal communities, transportation corridors, and critical infrastructure. These geomorphological characteristics make the Oman watersheds an ideal complementary case study, offering valuable contrast to basins in Egypt and enabling a broader understanding of flood dynamics across regions that drain into different hydrological sinks—Lake Nasser, the Red Sea, and the Gulf of Oman.

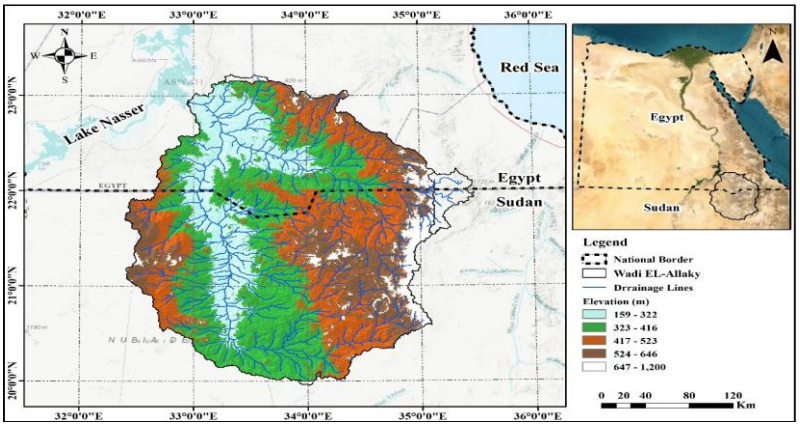

Figure 1: Location map of Wadi Al-Allaqi, and elevation range.

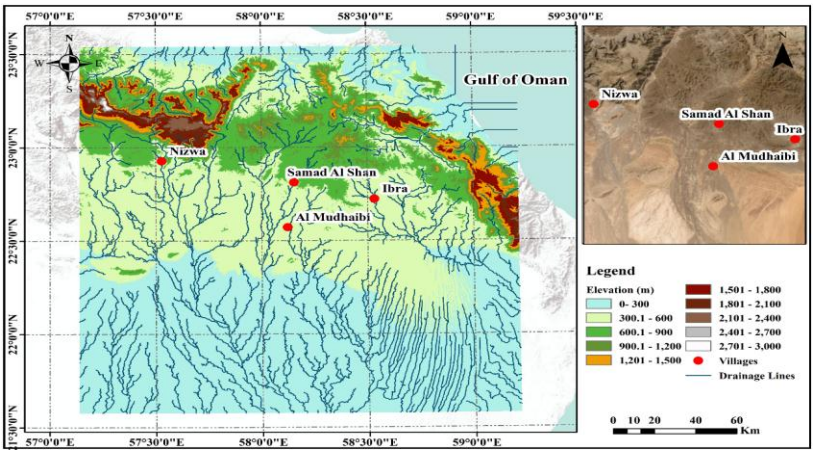

Figure 2: Location map of Oman case study, and elevation range.

Sentinel-2A Remote Sensing Data

The Sentinel-2A satellite, launched on June 23, 2015, captures high-resolution multispectral imagery across 13 spectral bands, spanning the visible, near-infrared (NIR), and shortwave infrared (SWIR) ranges. Among these, three bands (Band 1, Band 9, and Band 10) are specifically designed for atmospheric correction, enhancing data accuracy. Following its success, Sentinel-2B was deployed on March 7, 2017, further improving the frequency and coverage of Earth observation—particularly for agricultural monitoring. Even without the atmospheric bands, researchers can derive valuable spectral indices to analyze vegetation, water bodies, and land use changes.

As a key component of the European Union's Copernicus program (EOS, 2020), the Sentinel-2 mission provides freely accessible data through the European Space Agency's (ESA, 2015) open-access platform. The detailed specifications of its 13 spectral bands are summarized in Table 1, offering scientists and policymakers a powerful tool for environmental and agricultural assessments.

Table 1: Sentinel-2 satellite band specifications (ESA, 2015)

Band Number	Band name	Central Wavelength (nm)	Bandwidth (nm)	Resolution (m)
1	Coastal Aerosol	443	20	60
2	Blue	490	65	10
3	Green	560	35	10
4	Red	665	30	10
5	Vegetation Red Zone	705	15	20
6	Vegetation Red Zone	740	15	20
7	Vegetation Red Zone	783	20	20
8	Near Infrared-NIR	842	115	10
8 b	Narrow Close Infrared-NIR	865	20	20
9	Water vapor	945	20	60
10	Medium Infrared–SWIR–Cirrus	1375	30	60
11	Medium Infrared–SWIR	1610	90	20
12	Medium Infrared–SWIR	2190	180	20

In recent years, sentinel-2 satellite data has become a valuable resource in atmospheric and Earth sciences, supporting diverse research applications. Recent studies have utilized its capabilities for flood mapping (Dinh et al., 2019b; Rättich et al., 2020), agricultural monitoring (Kobayashi et al., 2020; Zhang et al., 2019), land use/land cover (LULC) assessments (Majidi Nezhad et al., 2019; Miranda et al., 2018), and wetland observation (Solovey, 2020). The Sentinel-2 data used in this study offers several key advantages:

- A 10-day revisit time at the equator, improving to 5 days at mid-latitudes due to the dual-satellite system (Sentinel-2A and -2B).
- High spatial resolution (10 m) for detailed environmental analysis.

These features make Sentinel-2 particularly effective for dynamic, large-scale monitoring tasks.

Although Sentinel-2 have the disadvantage of being affected by cloudiness, optical imaging satellites are preferred for flood studies to radar satellites, which are not affected by cloudiness but make it difficult to analyze and access data. Using free access satellite data is a low-cost method of analyzing flooded areas, while ground-based measurements and field surveys have high costs and are labor-intensive.

In this current study, multiple spectral indices, including the (NDWI), (MNDWI), and (SMI), were used to define and generate different maps for the flood inundation.

The indices based on remote sensing utilized the different spectral characteristics of water, vegetation, and soil moisture, which allowed assessment of the general flood scenario and the delineation of specific flooded areas. These calculations were carried out following appropriate pre-processing of the available satellite data to extract inundated zones corresponding to extreme rainfall events. In the case of Wadi Al-Allaqi, the analysis focused on the flood event that occurred in the southern part of Egypt during August 2024, enabling the mapping of floodwater extent within the basin. To broaden the scope of the study, some Wadis outside Egypt in Oman was also considered as a complementary case study. This basin drains eastward into the Gulf of Oman and experienced a significant extreme rainfall event on 14-15 April 2024, which triggered flash floods along its steep catchment. The flood mapping of these wadies provides valuable insights into flood hazards affecting coastal areas

directly connected to the Gulf of Oman, offering a different hydrological context compared to Wadi Al-Allaqi, which drains into Lake Nasser. Overall, these satellite-based flood mapping exercises provide critical information to support disaster response, damage assessment, and future flood risk management, particularly in the absence of field measurements but with the availability of reasonably high-resolution satellite imagery.

The aforementioned rainfall events led to a flash-flood that caused high amounts of surface runoff flowing through the main streams until it reached Lake Naser and the Gulf of Oman. Fortunately, satellite imagery was acquired on the same day as the flash flood for wadi Al-Allaqi, but after the event, providing a valuable opportunity to assess the extent of the flood using different spectral indices.

METHODOLOGY

Sentinel-2 is part of the Copernicus program, a large initiative led by the European Union and the European Space Agency (ESA) to provide free and open access to high-quality land surface data and information services. The advanced capabilities of the Copernicus platform, coupled with SNAP software, allowed reliable and efficient flood mapping of Wadi El-Allaqi basin using Sentinel-2 data.

For Wadi El-Allaqi, the Sentinel-2 images that covered part of the study area were combined into one seamless image. This mosaicking step automatically aligned and blended multiple Sentinel-2 tiles to create a smooth and consistent dataset for the analysis, covering a significant portion of the basin. After mosaicking, geometric resampling and subset operations were carried out to adjust the shapes and sizes of the images and to extract the relevant portion of the satellite data, focusing on the flood-affected region. Figure 3 illustrates the flood inundation extent in Wadi El-Allaqi extracted from Sentinel-2 images.

In contrast, the flood mapping for the Oman case study was conducted using Sentinel-2B imagery acquired during the intense rainfall event from 14 to early 15 April 2024. Sentinel-2B, with its Multispectral Instrument (MSI), provides 10–20 m resolution data well suited for detecting flood extent across the rugged mountainous catchments draining toward the Gulf of Oman. The imagery was pre-processed, atmospherically corrected, and subset to isolate the affected basins. These steps enabled accurate extraction of inundated areas and identification of flash-flood pathways within the steep valleys and deeply incised channels characteristic of the region. Figure 4 presents the flood inundation extent for the Oman watersheds derived from the Sentinel-2B dataset.

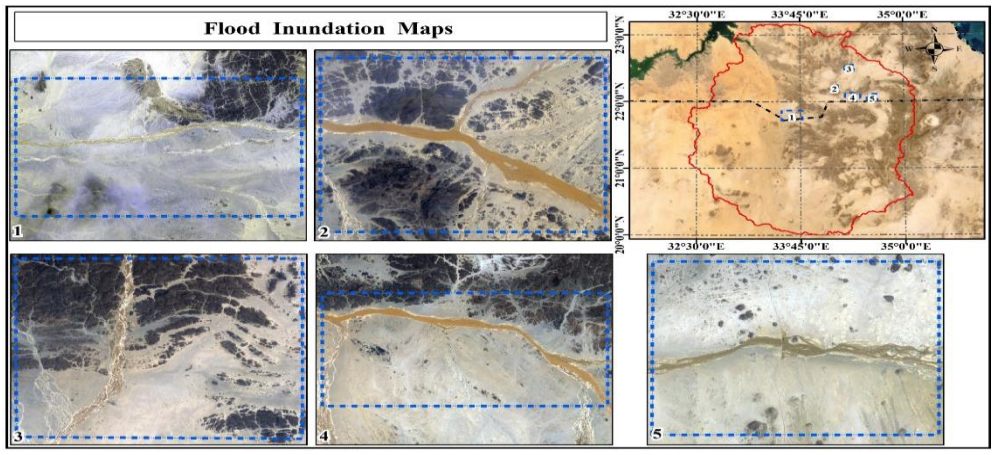

Figure 3: Flood inundation extent from sentinel-2 satellite images (Wadi El-Allaqi)

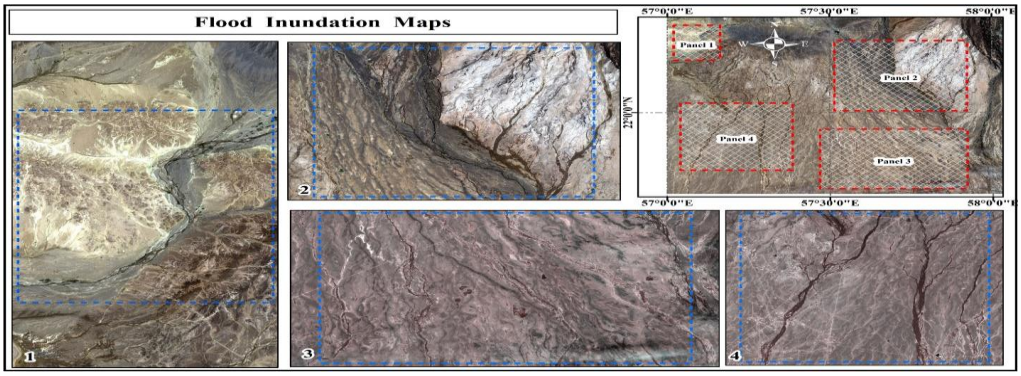

Figure 4: Flood inundation extent from sentinel-2 satellite images (Oman Case study)

NDWI was used to identify flooded areas in the satellite images. It was calculated using Equation (1).

$$\text{NDWI} = \frac{\text{Green} - \text{NIR}}{\text{Green} + \text{NIR}} \qquad (1)$$

Where:

Green (Green band):

- This is the amount of light reflected in the green part of the light spectrum, usually around 530-570 nm.
- The green band is used in detecting vegetation and helps tell the difference between water and land. Water usually reflects much less green light than land or plants
-

NIR (Near Infrared band):

- This is the amount of light reflected in the near-infrared part of the light spectrum, typically around 700-1300 nm.
- The NIR band is sensitive to vegetation and water, with water absorbing most of the infrared light, making it stand out in this spectral region.

The MNDWI was also used to identify the areas affected by the flood. The MNDWI was calculated using Equation (2). This method is better at reducing the impact of buildings and vegetation, which is especially useful in dry areas with little vegetation.

$$\text{MNDWI} = \frac{\text{Green} - \text{SWIR}}{\text{Green} + \text{SWIR}} \qquad (2)$$

Where:

SWIR (Short Wave Infrared Band):

- Reflectance in the short-wave infrared (SWIR) spectrum, which helps distinguish water from built-up areas and vegetation due to its strong absorption by water.

Additionally, SMI was looked at as a possible way to give a better overall picture of flood conditions. SMI was worked out using Equation (3).

$$\text{SMI} = \frac{\text{NIR} - \text{SWIR}}{\text{NIR} + \text{SWIR}} \qquad (3)$$

The Sentinel-2 and Landsat 8 satellite images, taken on the same day as the flash flood, gave a great chance to check how well NDWI, MNDWI, and SMI could map the actual flood area. By looking at the Sentinel-2 and Landsat 8 pictures and spotting the areas with surface water, a reference flood map representing a satellite image of a flash flood spreading across the watershed was made to compare with the flood areas shown by the indexes.

The study evaluated flood inundation mapping accuracy in the arid Wadi El-Allaqi and Wadi Hodein basin by systematically comparing flood extents derived from spectral indices and their optimal threshold values against high-resolution Sentinel-2 and Landsat 8 reference data. This approach identified the most effective index and threshold combinations for delineating flood boundaries in arid environments, where field-based validation data are often unavailable due many reasons especially like these regions.

In the final stage of the study, the different flood maps generated by the NDWI, MNDWI, and SMI methods were carefully evaluated against the actual flood extent captured in available Sentinel-2 and Landsat 8 satellite imageries from the peak of the event. Without reliable field measurements, the Sentinel-2 and Landsat 8 images served as the most authoritative reference to verify how effectively each method identified the flooded areas. By analyzing these results, the study could determine which technique NDWI, MNDWI, or SMI provided the closest match to real-world conditions, offering valuable insights for future flood monitoring in similar regions.

Statistical Evaluation Criteria

To assess the performance of the flood detection indices (NDWI, MNDWI, and SMI), a set of classification metrics was used to evaluate how well each method matched the reference flood inundation map derived from Sentinel-2 and Landsat 8 satellite imageries. The reference map was treated as ground truth, where each validation point was classified as either inundated (1) or non-inundated (0) based on visual interpretation of post-event satellite data.

In addition to overall accuracy, which measures the proportion of correctly classified points, the following statistical metrics were employed to provide a more comprehensive evaluation:

Precision (Positive Predictive Value)

Precision measures the proportion of points predicted as flooded that are actually flooded in the reference map. It helps assess how well the index avoids false flood detection.

$$\textbf{Precission} = \left(\frac{\mathrm{Tp}}{\mathrm{TP} + \mathrm{FP}}\right) \qquad (4)$$

Where:

- **TP**: True Positives — points correctly identified as flooded.
- **FP**: False Positives — points incorrectly identified as flooded (actually dry).

A high precision means the method reliably predicts flood zones with minimal overestimation.

Recall (Sensitivity or True Positive Rate)

Recall represents the proportion of actual flooded points in the reference map that were correctly identified by the index.

$$\textbf{Recall} = \left(\frac{\mathrm{Tp}}{\mathrm{TP} + \mathrm{FN}}\right) \qquad (5)$$

Where:

- **FN:** False Negatives points that were flooded in reality but not detected by the index.

High recall indicates the method is effective at capturing all or most flood-affected areas, minimizing underestimation.

F1 Score (Harmonic Mean of Precision and Recall)

The F1 Score combines both precision and recall into a single metric, providing a balance between the two. It is especially useful when the dataset is imbalanced (i.e., when the number of flooded and non-flooded points is not equal).

$$\textbf{F1 Score} = 2 * \left(\frac{\textbf{Precission} * \textbf{Recall}}{\textbf{Precission} + \textbf{Recall}}\right) \quad (6)$$

The F1 Score ranges between 0 and 1, where 1 indicates perfect precision and recall. It offers a more balanced view of the method's performance, especially when both false positives and false negatives carry significant consequences.

Accuracy (Overall Classification Accuracy)

Although previously defined, for completeness, the overall accuracy was calculated as follows:

$$\text{Accuarcy} = \left(\frac{\text{Tp} + \text{TN}}{\text{TP} + \text{FP} + \text{FN} + \text{TN}}\right) * 100 \quad (7)$$

Where:

- **TN**: True Negatives correctly identified dry points.

While accuracy gives a general performance measure, it can be misleading when the data is imbalanced, hence the need for additional metrics like precision, recall, and F1 score, this equation shows that the ratio of correctly identified cases (both flooded and non-flooded) compared to the total number of cases evaluated. A higher accuracy percentage means the index is better at detecting flood areas.

Case Study 1: Wadi El-Allaqi (Sentinel-2 Data, August 2024 Event)

Interpretation of resulted indices' maps

A set of 5 panels was shown, each with four smaller images (Figures 5-9). In each panel, the top-left image showed the Normalized Difference Water Index (NDWI) map, the top-right image displayed the Modified Normalized Difference Water Index (MNDWI) map, and the bottom-left image showed the Surface Water Moisture Index (SMI) map. The top-middle image in each panel was the reference flood map, created directly from Sentinel-2 satellite images, which served as the accurate flood area for comparison. To check how well each index worked, 15 points were added to the same spots in all the images. Figures 10 shows the results for the 3 indexes (NDWI, MNDWI, SMI), including True Positives (TP), False Positives (FP), True Negatives (TN), and False Negatives (FN) for each index.

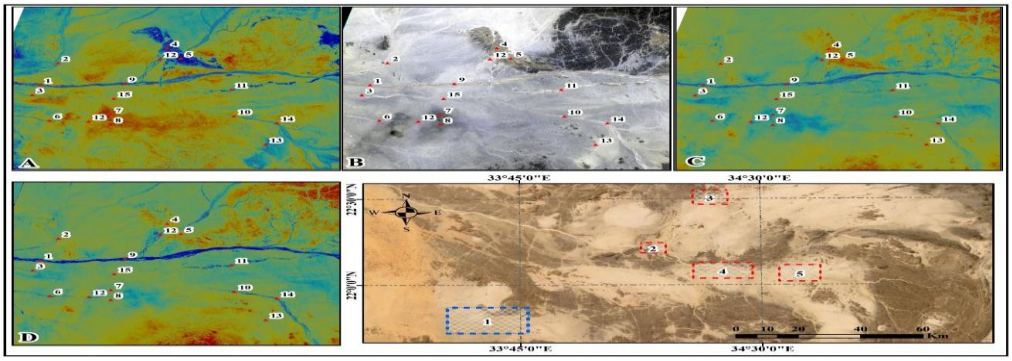

Figure 5: Flood Mapping Indices and Sentinel-2 Flood Inundation map for Panel 1.

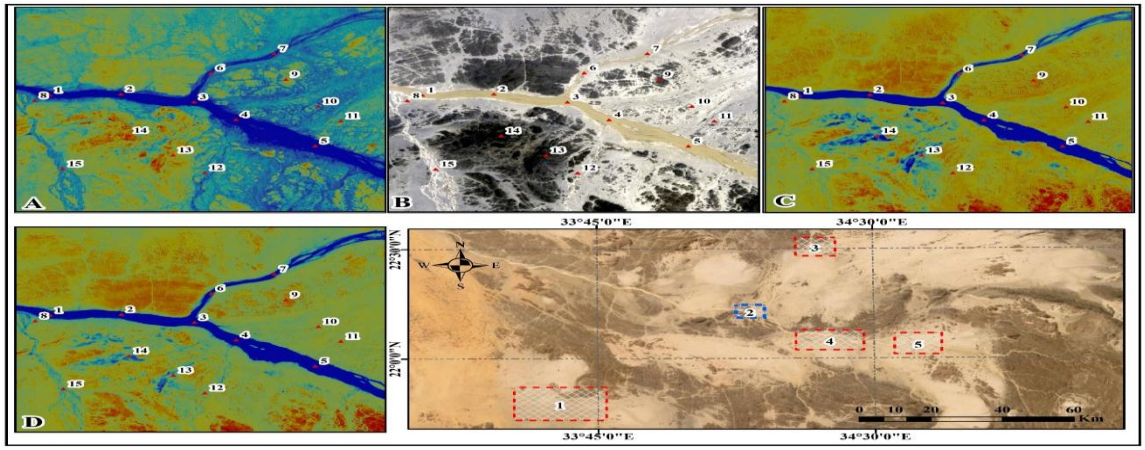

Figure 6: Flood Mapping Indices and Sentinel-2 Flood Inundation map for Panel 2.

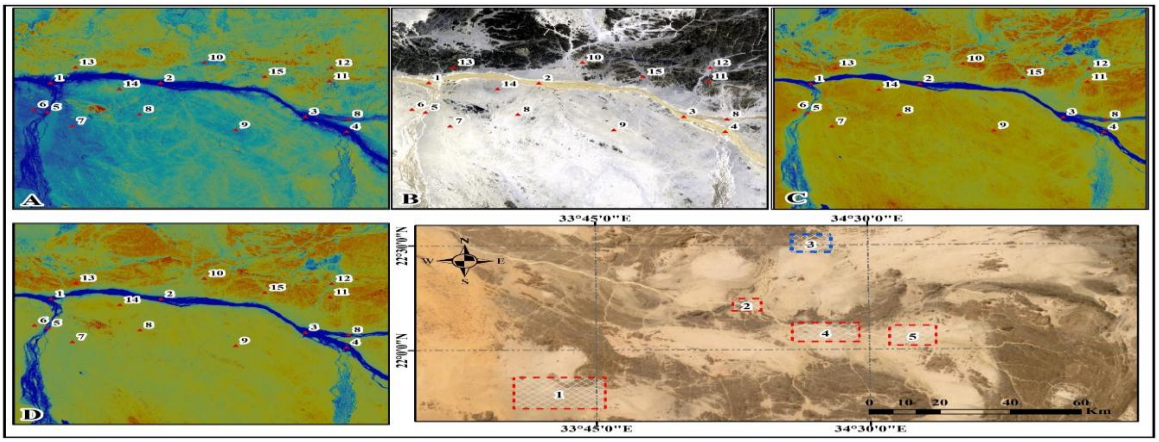

Figure 7: Flood Mapping Indices and Sentinel-2 Flood Inundation map for Panel 3.

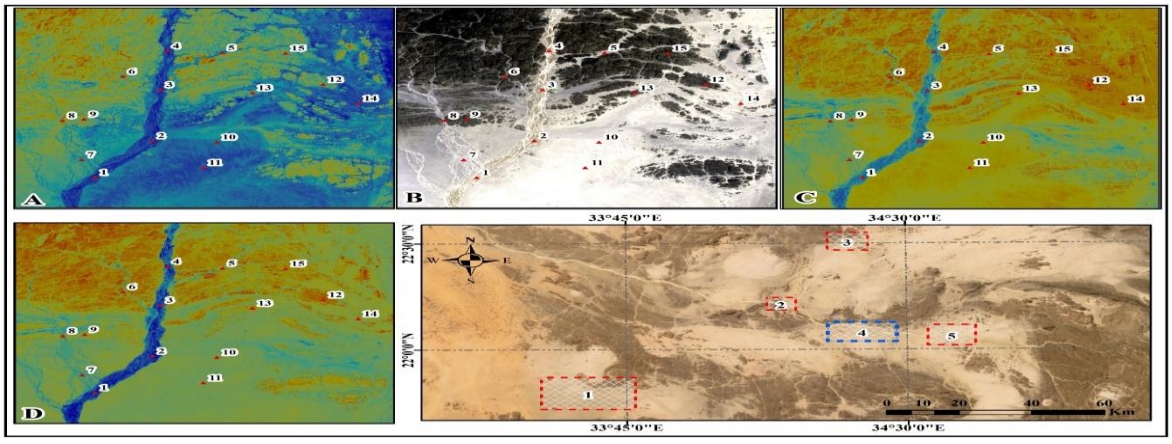

Figure 8: Flood Mapping Indices and Sentinel-2 Flood Inundation map for Panel 4.

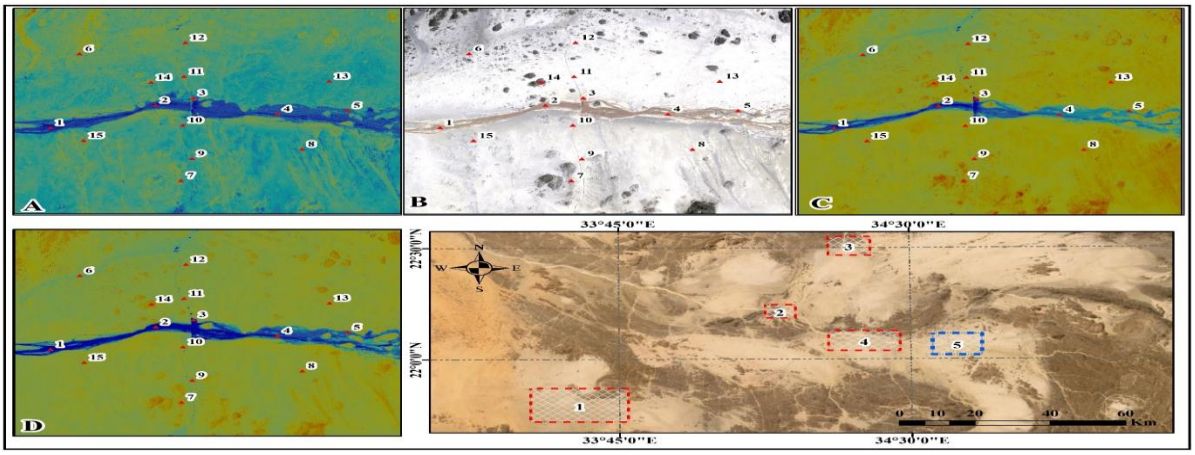

Figure 9: Flood Mapping Indices and Sentinel-2 Flood Inundation map for Panel 5.

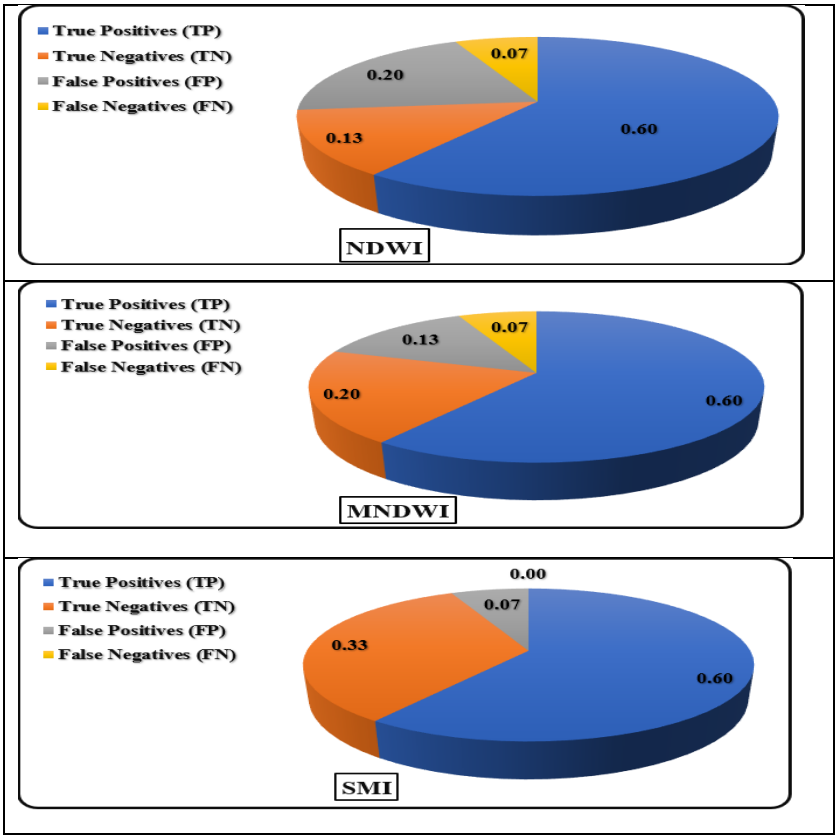

Figure 10: Percentage Distribution of Prediction Outcomes (TP, FP, TN, FN) for All indices.

The precision accuracy results showed that MNDWI and SMI performed consistently well (≥0.89) across all panels, even achieving perfect scores (1.0) in most cases, meaning they rarely mislabeled non-water as water as shown in (Figure 11). NDWI, however, had more variable results (0.54–0.83) and struggled in panel 3. This confirms that MNDWI and SMI are better choices when high precision is needed.

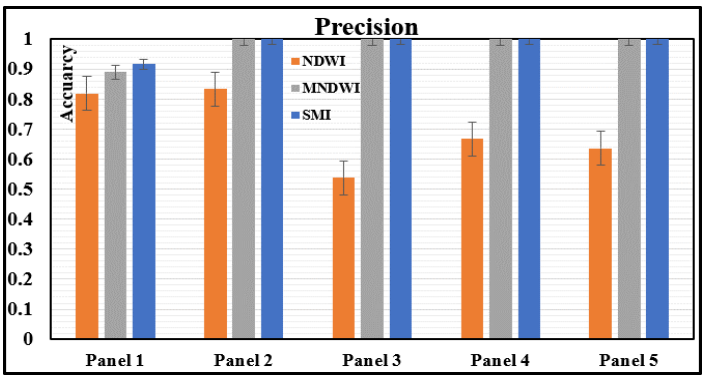

Figure 11: Precision Accuracy of NDWI, MNDWI, and SMI across Five Panels.

The recall accuracy results reveal that NDWI achieved the highest detection rate (0.875) in Panel 3 but showed inconsistency across other panels (0.636–0.769), indicating variability in identifying true water areas (Figure 12). MNDWI and SMI demonstrated more stable performance (0.667–0.838), with SMI excelling in Panel 1 (0.917) and both indices matching closely in Panels 4–5 (0.818–0.838). While NDWI occasionally outperformed in specific cases (e.g., Panel 3), its lower scores elsewhere suggest reliability issues, whereas MNDWI and SMI maintained balanced sensitivity to water detection.

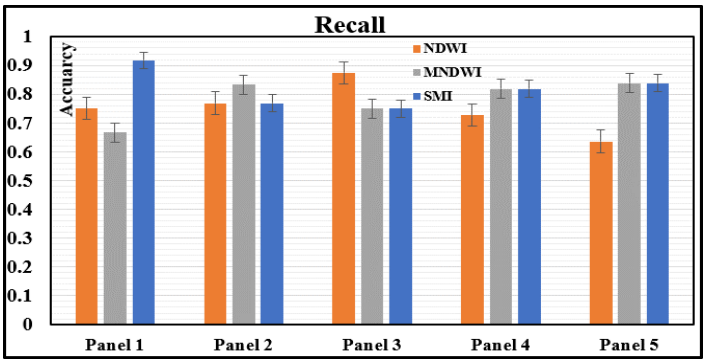

Figure 12: Recall Accuracy of NDWI, MNDWI, and SMI Across Five Panels.

The F1-score results (Figure 13) show that SMI performed best overall, with scores between 0.857 and 0.917 across all tests. It did especially well in panels 1, 4, and 5, where it scored 0.9 or higher. MNDWI also did well, with similar high scores in panels 2, 4, and 5. NDWI, however, had lower scores (0.636–0.8) and struggled in panels 3 and 5. This makes it clear that SMI and MNDWI are better choices than NDWI for accurately identifying water areas.

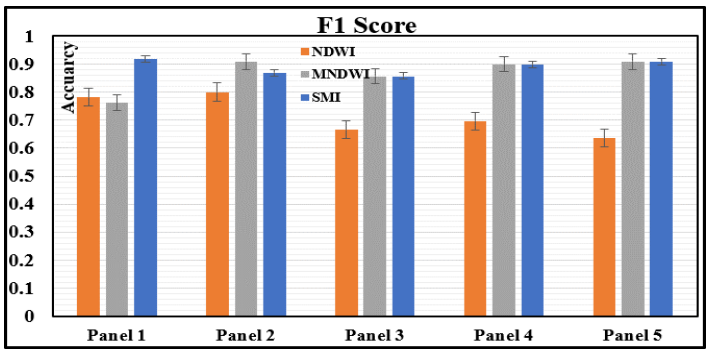

Figure 13: F1 Score Accuracy of NDWI, MNDWI, and SMI Across Five Panels.

The average results across all five panels show that SMI had the highest scores overall, with 0.98 precision, 0.82 recall, and 0.89 F1-score (Figure 14). MNDWI was close behind, scoring 0.98 precision, 0.78 recall, and 0.87 F1-score. NDWI performed noticeably worse, with lower averages in all three measures (0.70 precision, 0.75 recall, 0.72 F1-score). These results proved that – SMI gives the most accurate and consistent results for identifying water areas.

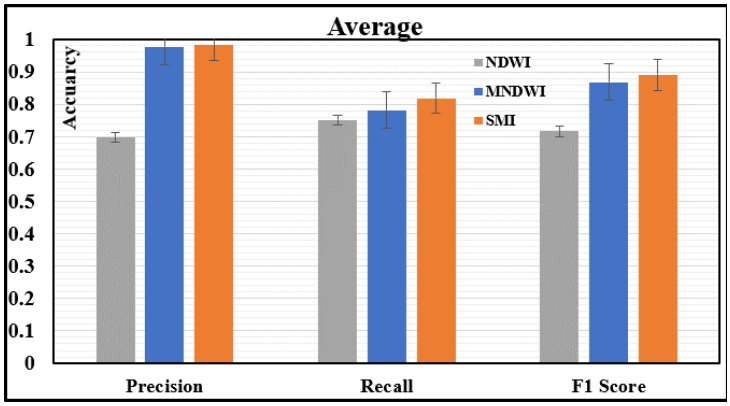

Figure 14: Average Precision, Recall, and F1-Scores of NDWI, MNDWI, and SMI Across All Panels.

This detailed study aimed to check how well the NDWI, MNDWI, and SMI indices work in accurately mapping the flood areas in the Wadi El-Allaqi basin. It also aimed to find out which index is the best for the dry conditions of the study area. Each index's average overall accuracy was determined for each of the five panels (1, 2, 3, 4, and 5), which together included 75 reference points. The results showed that the average accuracies for the (NDWI), (MNDWI), and (SMI) were 69.33%, 88%, and 94.66%, respectively., These results demonstrate how differently each index performs across the chosen panels in terms of precisely identifying flood inundation.

when we compare between the accuracies of three indices with each other, With nearly 1.07 and 1.36 better performance than the MNDWI and NDWI, respectively, the SMI was clearly identified as the most appropriate index for flood mapping in this arid environment by this thorough analysis, which included comparing the NDWI, MNDWI, and SMI indices across multiple locations within the Wadi El-Allaqi basin. The SMI is the suggested method for upcoming flood monitoring and assessment initiatives in the area since it was able to take advantage of the complementary advantages of both NDWI and MNDWI, producing a more accurate and thorough depiction of the flood inundation.

Case Study 2: Oman (Sentinel-2 Data, April 2024 Event)

Interpretation of resulted indices' maps

A similar approach was applied for Oman, where Sentinel-2_image corresponding to the 14 April 2024 rainfall event was used. As in the first case study, a set of panels was generated, each containing the NDWI, MNDWI, and SMI index maps, alongside the reference flood map derived from Sentinel-2_imagery. To ensure comparability, validation points were also distributed across all images for accuracy evaluation.

Figures 15–18 present the flood mapping results for the four panels in Oman, with each panel showing the three indices (NDWI, MNDWI, SMI) and the reference flood map. Figure 19 shows the results for the 3 indexes (NDWI, MNDWI, SMI), including True Positives (TP), False Positives (FP), True Negatives (TN), and False Negatives (FN) for each index.

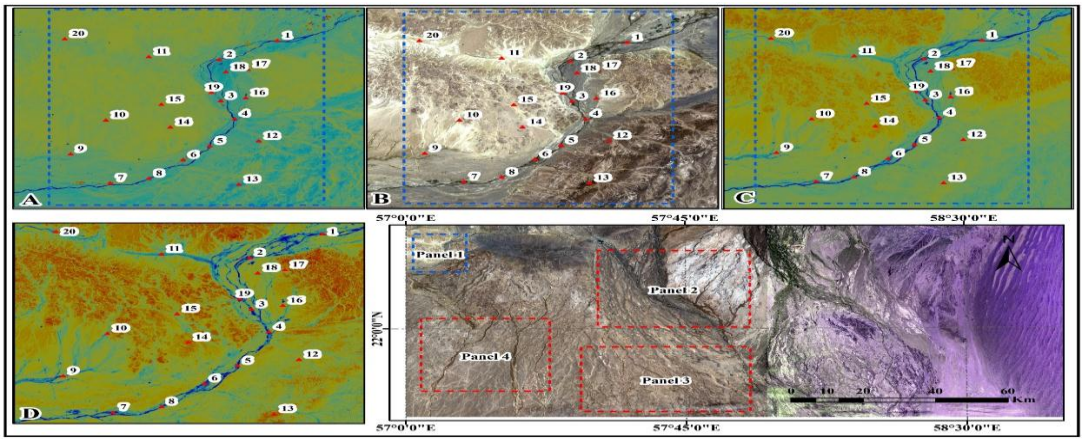

Figure 15: Flood Mapping Indices and Sentinel-2 Flood Inundation map for Panel 1.

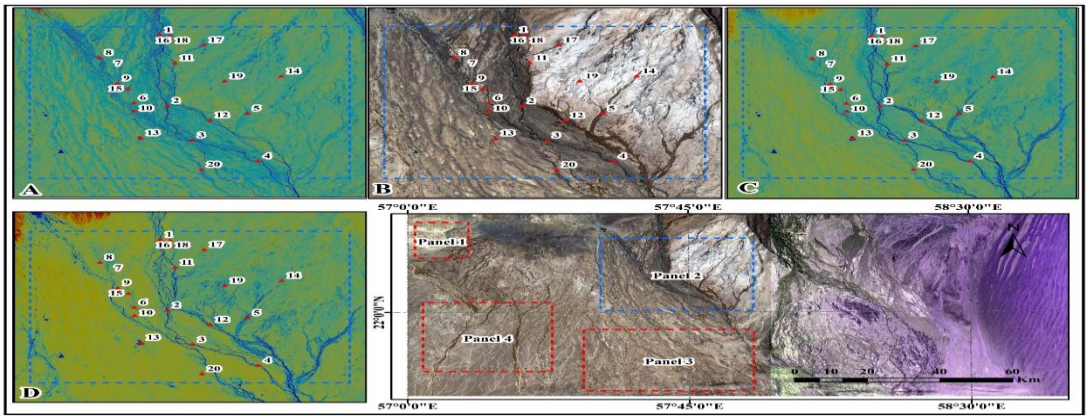

Figure 16: Flood Mapping Indices and Sentinel-2 Flood Inundation map for Panel 2.

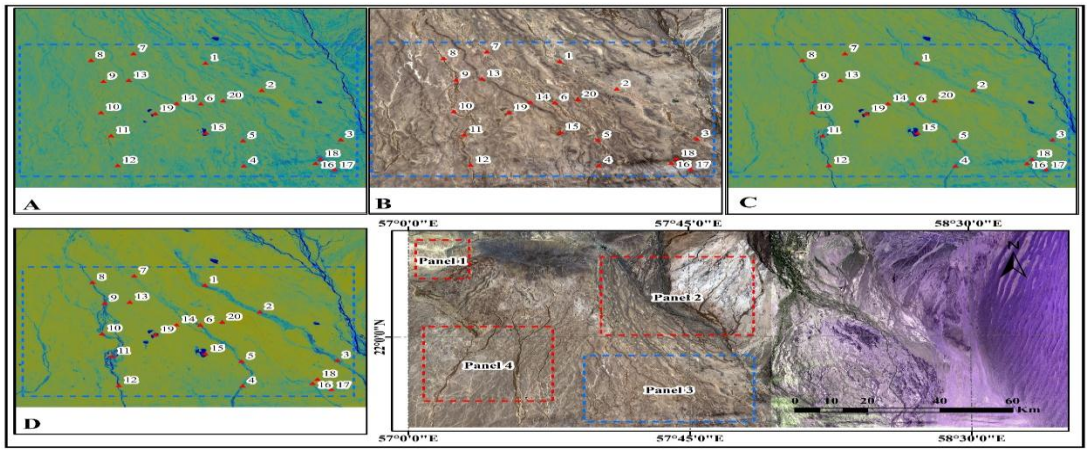

Figure 17: Flood Mapping Indices and Sentinel-2 Flood Inundation map for Panel 3.

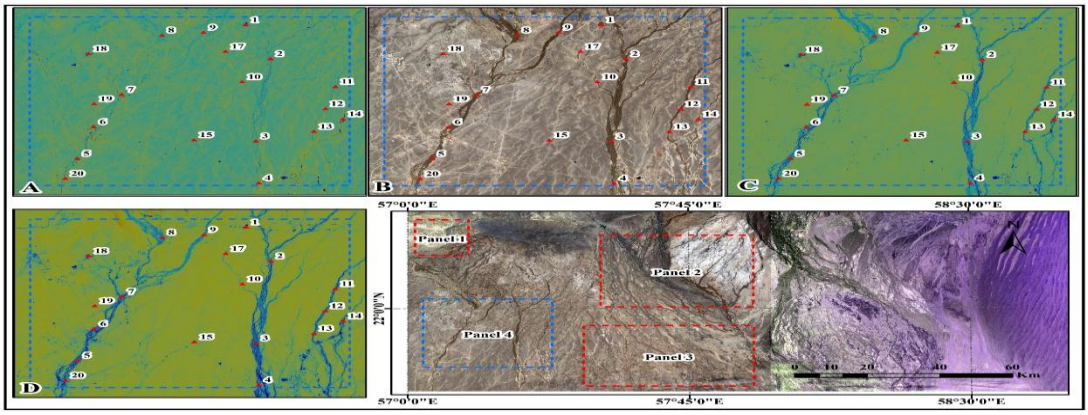

Figure 18: Flood Mapping Indices and Sentinel-2 8 Flood Inundation map for Panel 4.

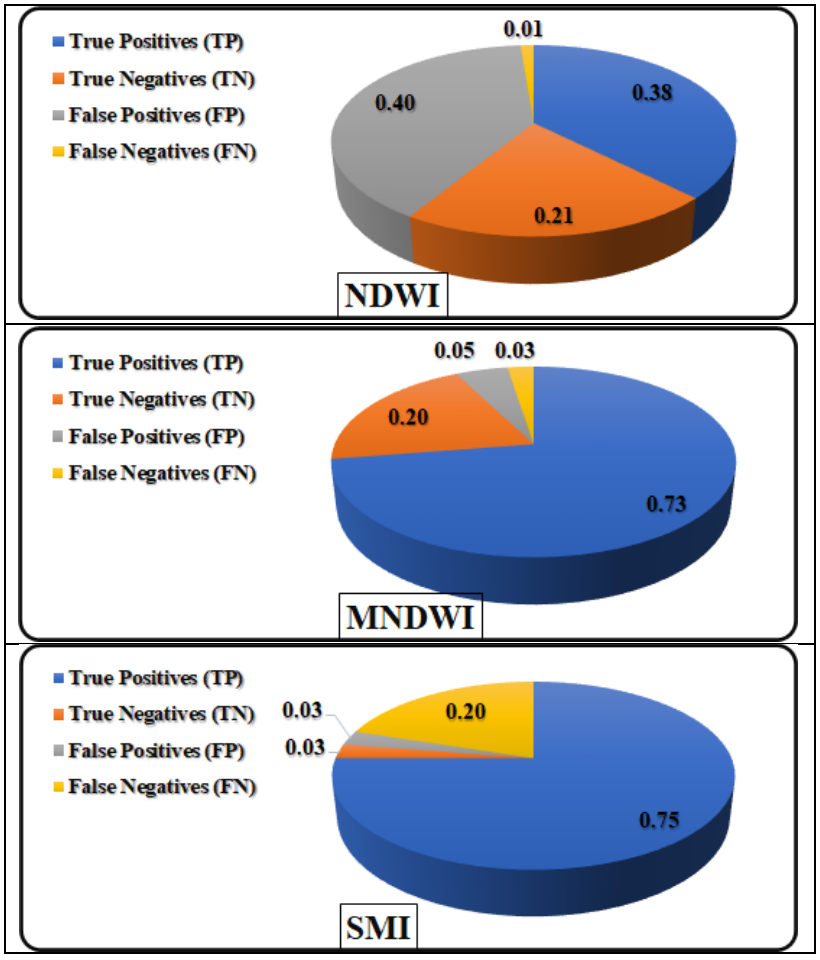

Figure 19: Percentage Distribution of Prediction Outcomes (TP, FP, TN, FN) for All indices.

The precision results across the four panels show clear performance differences among the three indices. **SMI** achieved the strongest and most consistent results, with perfect precision (1.0) in Panels 1 and 4, and very high values in Panels 2 (0.92) and 3 (0.94). MNDWI also performed reliably, with high precision in Panels 1 (0.82), 3 (0.78), and 4 (0.95), though it showed a

moderate drop in Panel 2 (0.60). In comparison, NDWI exhibited the greatest variability, with lower precision in Panels 1 and 2 (both 0.56), a moderate improvement in Panel 3 (0.60), and strong performance in Panel 4 (0.90). Overall, SMI and MNDWI demonstrated more stable and accurate flood identification than NDWI across the four panels.

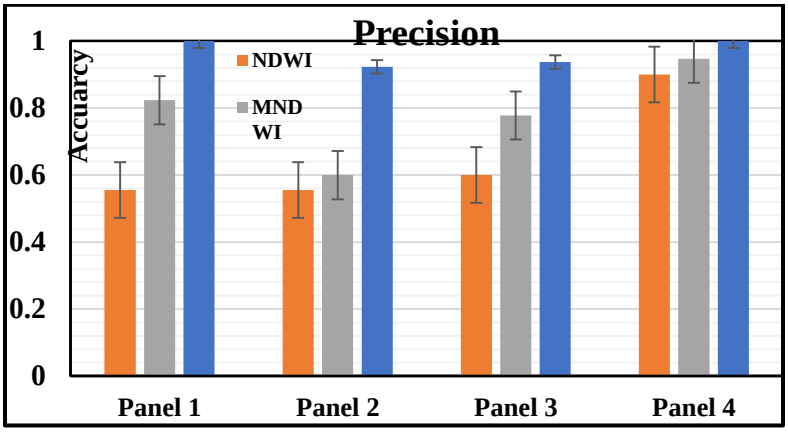

Figure 20: Precision Accuracy of NDWI, MNDWI, and SMI Across four Panels.

The recall results (Figure 21) reveal notable differences among the three indices across the four panels. MNDWI demonstrated the strongest and most consistent performance, achieving very high recall in all panels, with values of 0.93 in Panel 1, perfect recall (1.0) in Panels 2 and 3, and 0.95 in Panel 4. SMI also maintained high recall, matching MNDWI in Panel 1 (0.93) and Panel 2 (1.0), and performing well in Panel 3 (0.94) and Panel 4 (1.0). In contrast, NDWI showed substantial variability, with low recall in Panels 1 (0.33), 3 (0.38), and 4 (0.47), and only a moderate value in Panel 2 (0.83), indicating a high rate of missed flooded pixels. Overall, MNDWI and SMI clearly outperformed NDWI, offering more reliable and consistent detection of inundated areas across all panels.

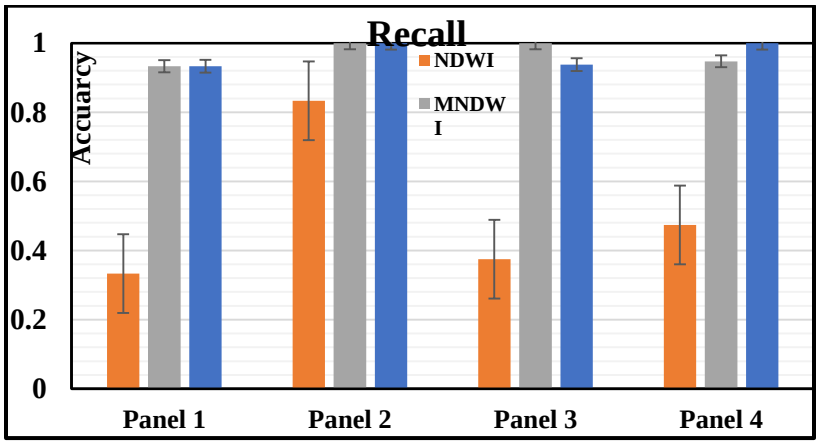

Figure 21: Recall Accuracy of NDWI, MNDWI, and SMI Across four Panels.

The F1-score analysis reinforces the overall superiority of SMI across the four panels. SMI achieved consistently high performance, with F1-scores of 0.97 in Panel 1, 0.96 in Panel 2, 0.94 in Panel 3, and a perfect score (1.0) in Panel 4, reflecting excellent balance between precision and recall. MNDWI also performed strongly, with high F1-scores in Panels 1 (0.88), 3 (0.88), and 4 (0.95), though it showed a moderate decline in Panel 2 (0.75). In contrast, NDWI displayed weaker and more variable performance, ranging from 0.42 in Panel 1 to 0.67 in Panel 2, with moderate values in Panels 3 (0.46) and 4 (0.62). Overall, SMI demonstrated the most stable and accurate flood detection, clearly outperforming both NDWI and MNDWI across all panels.

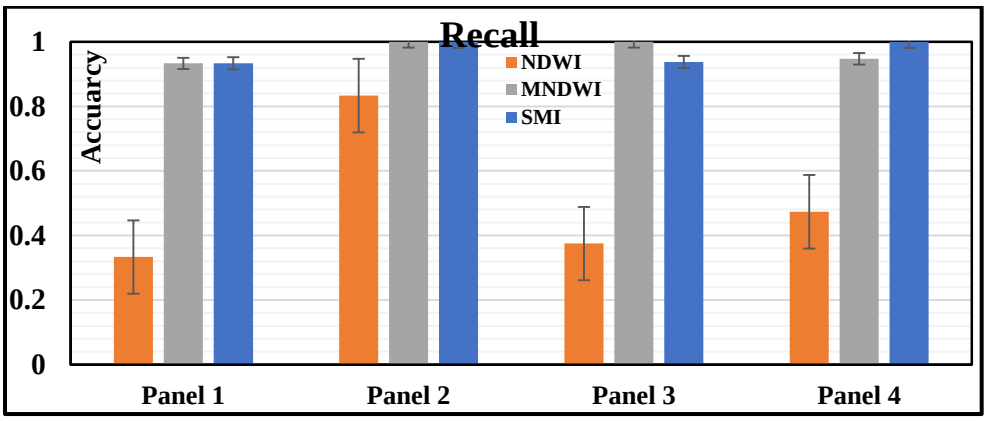

Figure 22: F1 Score Accuracy of NDWI, MNDWI, and SMI Across four Panels.

The overall average results across all panels highlight clear differences in the performance of the three indices. SMI achieved the strongest and most balanced performance, with high averages for precision (0.97), recall (0.97), and F1-score (0.97), confirming its stability and reliability for flood detection. MNDWI also performed well, particularly in recall, achieving an exceptionally high average value of 0.97, although its precision (0.79) and F1-score (0.86) were comparatively lower, indicating a tendency to detect most flooded pixels but with a higher rate of false positives. In contrast, NDWI showed weaker and more variable performance, with average precision of 0.65, recall of 0.50, and an overall F1-score of 0.54, reflecting limited consistency across panels. These results confirm that SMI provides the most accurate and stable flood mapping, while MNDWI may be preferable in applications where maximizing flood detection (high recall) is the primary objective.

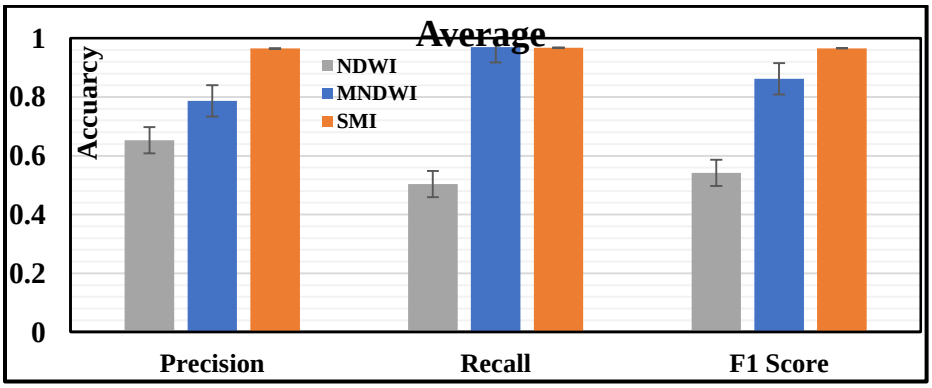

Figure 23: Average Precision, Recall, and F1-Scores of NDWI, MNDWI, and SMI Across All Panels.

CONCLUSIONS

This research highlights the increasing importance of remote sensing technologies in hydrological studies, especially in cases or regions where conventional ground-based measurement methods are constrained. The capability of satellite data to provide accurate and timely flood information significantly enhances disaster management efforts, supporting more efficient resource allocation and response planning.

By using advantages of Sentinel-2's spatial resolution and multi-spectral capabilities, the methodology strengthens the validation process, ensuring strong and reliable flood extent maps. These findings provide a critical foundation for improving flood risk management strategies, supporting disaster response efforts, and enhancing damage assessment frameworks in arid regions. The results offer practical insights for stakeholders tasked with mitigating flood impacts in similar environments, while the methodology itself serves as a replicable framework for future studies in data-scarce regions.

The study highlights the importance of selecting appropriate and suitable indices for flood inundation mapping using remote sensing data. Selecting the appropriate indices significantly enhances the accuracy and effectiveness of flood monitoring efforts, offering valuable insights to improve disaster response, resource allocation, and flood risk management strategies in regions with similar challenges. The effectiveness of three spectral indices Normalized Difference Water Index (NDWI), Modified Normalized Difference Water Index (MNDWI), and Surface Water Moisture Index (SMI) was examined in predicting flood inundation zones using available Sentinel-2 satellite imagery, particularly within the context of a significant flash flood event in Wadi El-Allaqi basin, Southeast of Egypt and in Oman.

The findings revealed that while all indices provided valuable insights into flood dynamics, SMI outperformed NDWI and MNDWI with an accuracy of 94.66% in delineating flood-affected areas in Wadi El-Allaqi, and achieved 95% accuracy in the Oman case study, confirming it as the most effective and consistent index across both regions. This superior performance was validated through ground truth comparisons and error matrix analysis, confirming SMI's precision in distinguishing water bodies from surrounding terrain. The advantage of SMI lies in its ability to integrate surface water and soil moisture information, a critical feature in arid environments where mixed spectral signals from vegetation and moist soils often hinder accurate water detection.

The integration of SMI into flood mapping practices can significantly enhance the reliability of flood assessments, providing a more comprehensive understanding of flood extents compared to NDWI and MNDWI. Future research should explore: (1) hybrid approaches combining SMI with machine learning techniques like Random Forest (RF) or Support Vector Machines (SVM) to classify water bodies more accurately by integrating spectral indices with terrain features, (2) calibration with ground truth data to refine index thresholds, and (3) application to other arid/semi-arid basins to validate methodology transferability. This could develop more robust models for flood detection and monitoring. In addition, integrating ground truth data, if available, could enhance the accuracy of flood assessments, where field measurements can help calibrate index thresholds and improve the overall validation process. The demonstrated effectiveness of SMI in this study, combined with its utilization of freely available Sentinel-2 imagery and standardized processing techniques, suggests strong potential for scalability to other arid and semi-arid regions globally. The methodology's robust performance (94.66% accuracy in Wadi El-Allaqi and 95% in Oman) in challenging desert flood conditions further confirms SMI as the best-performing index and indicates its likely transferability to similar hydrological basins. This scalability could significantly enhance flood monitoring capabilities in data-scarce regions facing similar flash flood risks.

The findings of this study contribute to the existing body of knowledge in remote sensing applications, highlighting the need for continuous refinement of methodologies used in flood monitoring.

DECLARATIONS

Conflict of Interest:

The author declares that there is no conflict of interest regarding the publication of this paper.

REFERENCES

1) Acharya, T. D., Subedi, A., & Lee, D. H. (2018). Evaluation of water indices for surface water extraction in a landsat 8 scene of Nepal. Sensors (Switzerland), 18(8). https://doi.org/10.3390/s18082580

2) Danso-Amoako, E., Scholz, M., Kalimeris, N., Yang, Q., & Shao, J. (2012). Predicting dam failure risk for sustainable flood retention basins: A generic case study for the wider Greater Manchester area. Computers, Environment and Urban Systems, 36(5), 423–433. https://doi.org/10.1016/j.compenvurbsys.2012.02.003

3) Dinh, D. A., Elmahrad, B., Leinenkugel, P., & Newton, A. (2019a). Time series of flood mapping in the Mekong Delta using high resolution satellite images. IOP Conference Series: Earth and Environmental Science, 266(1). https://doi.org/10.1088/1755-1315/266/1/012011

4) Dinh, D. A., Elmahrad, B., Leinenkugel, P., & Newton, A. (2019b). Time series of flood mapping in the Mekong Delta using high resolution satellite images. IOP Conference Series: Earth and Environmental Science, 266(1). https://doi.org/10.1088/1755-1315/266/1/012011

5) Feyisa, G. L., Meilby, H., Fensholt, R., & Proud, S. R. (2014). Automated Water Extraction Index: A new technique for surfacewatermappingusingLandsatimagery.RemoteSensingofEnvironment,140,2335.https://doi.org/10.1016/j.rse.2013.08.029

6) Ji, L., Zhang, L., & Wylie, B. (2009). Analysis of dynamic thresholds for the normalized difference water index. PhotogrammetricEngineeringandRemoteSensing, 75(11), 1307–1317. https://doi.org/10.14358/PERS.75.11.1307

7) Klemas, V. (2015). Remote sensing of floods and flood-prone areas: An overview. In Journal of Coastal Research (Vol. 31, Issue 4, pp. 1005–1013).CoastalEducation ResearchFoundation Inc. https://doi.org/10.2112/JCOASTRES-D-14-00160.1

8) Kobayashi, N., Tani, H., Wang, X., & Sonobe, R. (2020). Crop classification using spectral indices derived from sentinel-2aimageryJournalofInformationandTelecommunication,4(1),67–90.https://doi.org/10.1080/24751839.2019.1694765

9) Majidi Nezhad, M., Heydari, A., Fusilli, L., & Laneve, G. (2019). Land cover classification by using sentinel-2 images: A case study in the city of Rome. World Congress on Civil, Structural, and Environmental Engineering, 0. https://doi.org/10.11159/iceptp19.158

10) Martinis, S., Twele, A., & Voigt, S. (2009). Natural Hazards and Earth System Sciences Towards operational near real-time flood detection using a split-based automatic thresholding procedure on high resolution TerraSAR-X data. In Hazards Earth Syst. Sci (Vol. 9). www.nat-hazards-earth-syst-sci.net/9/303/2009/

11) McFeeters, S. K. (1996). The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features.InternationalJournalofRemoteSensing,17(7), 1425–1432. https://doi.org/10.1080/01431169608948714

12) McFEETERS, S. K. (1996). The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. International Journal of Remote Sensing, 17(7), 1425–1432.

13) Miranda, E., Mutiara, A. B., Ernastuti, & Wibowo, W. C. (2018). Classification of Land Cover from Sentinel-2 Imagery Using Supervised Classification Technique (Preliminary Study). Proceedings of 2018 International Conference on InformationManagementandTechnology,ICIMTech2018,69–74. https://doi.org/10.1109/ICIMTech.2018.8528122

14) Mosavi, A., Ozturk, P., & Chau, K. W. (2018). Flood prediction using machine learning models: Literature review. In Water (Switzerland) (Vol. 10, Issue 11). MDPI AG. https://doi.org/10.3390/w10111536

15) Rättich, M., Martinis, S., & Wieland, M. (2020). Automatic flood duration estimation based on multi-sensor satellite data. Remote Sensing, 12(4). https://doi.org/10.3390/rs12040643

16) Revilla-Romero, B., Thielen, J., Salamon, P., De Groeve, T., & Brakenridge, G. R. (2014). Evaluation of the satellite-based global flood Detection System for measuring river discharge: Influence of local factors. Hydrology and Earth System Sciences, 18(11), 4467–4484. https://doi.org/10.5194/hess-18-4467-2014

17) Rokni, K., Ahmad, A., Selamat, A., & Hazini, S. (2014). Water feature extraction and change detection using multitemporal landsat imagery. Remote Sensing, 6(5), 4173–4189. https://doi.org/10.3390/rs6054173

18) Schumann, G., & Di Baldassarre, G. (2010). The direct use of radar satellites for event-specific flood risk mapping. Remote Sensing Letters, 1(2), 75–84. https://doi.org/10.1080/01431160903486685

19) Smith, L. C. (1997). SATELLITE REMOTE SENSING OF RIVER INUNDATION AREA, STAGE, AND DISCHARGE: A REVIEW. In Ltd. Hydrol. Process (Vol. 11). John Wiley & Sons.

20) Solovey, T. (2020). Flooded wetlands mapping from sentinel-2 imagery with spectral water index: A case study of kampinos national park in central Poland. Geological Quarterly, 64(2), 492–505. https://doi.org/10.7306/gq.1509

21) Tsyganskaya, V., Martinis, S., Marzahn, P., & Ludwig, R. (2018). SAR-based detection of flooded vegetation–a review of characteristics and approaches. In International Journal of Remote Sensing (Vol. 39, Issue 8, pp. 2255–2293). Taylor and Francis Ltd. https://doi.org/10.1080/01431161.2017.1420938

22) Xu, H. (2006). Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery. International Journal of Remote Sensing, 27(14), 3025–3033. https://doi.org/10.1080/01431160600589179

23) Zhang, T. X., Su, J. Y., Liu, C. J., & Chen, W. H. (2019). Potential Bands of Sentinel-2A Satellite for Classification Problems in Precision Agriculture. International Journal of Automation and Computing, 16(1), 16–26. https://doi.org/10.1007/s11633-018-1143-x

24) EOS. (2020). Retrieved from : https://eos.com/sentinel-2/

25) ESA. (2015). 'Sentinel-2 User Handbook', ESA Standard Doc. (Vol. Issue 1 Rev 2.).

26) Twele et al., 2. A. (2016). Sentinel-1-based flood mapping: a fully automated processing chain. Remote Sens., 2990-3004. Retrieved from https://www.tandfonline.com/doi/full/10.1080/01431161.2016.1192304