SatVITS-Flood: Satellite Vegetation Index Time Series Flood detection model for hyperarid regions

Omer Burstein[1], Tamir Grodek[2], Yehouda Enzel[3], and David Helman[1]

[1]Institute of Environmental Sciences, Department of Soil and Water Sciences
[2]Institute of Earth Sciences, The Hebrew University of Jerusalem
[3]Hebrew University of Jerusalem

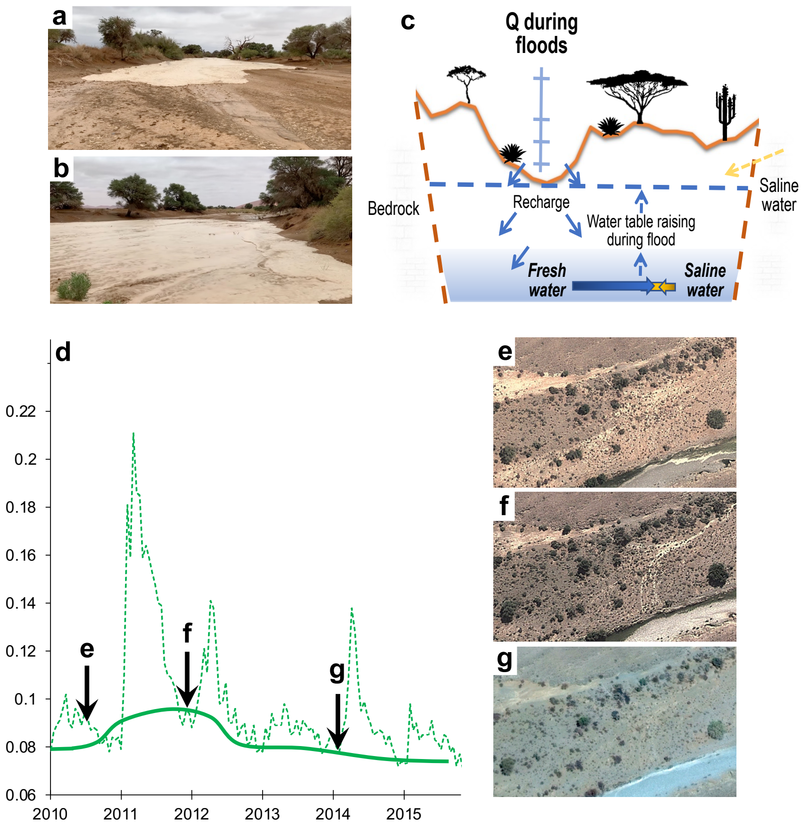
May 2, 2023

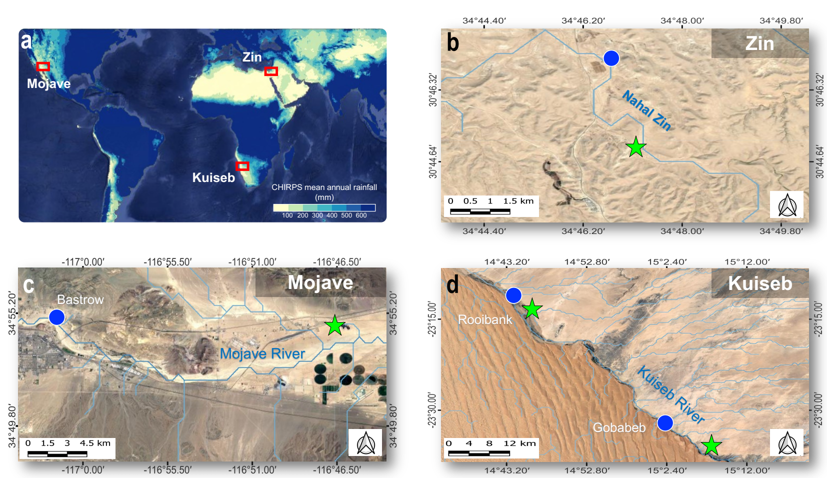
Abstract

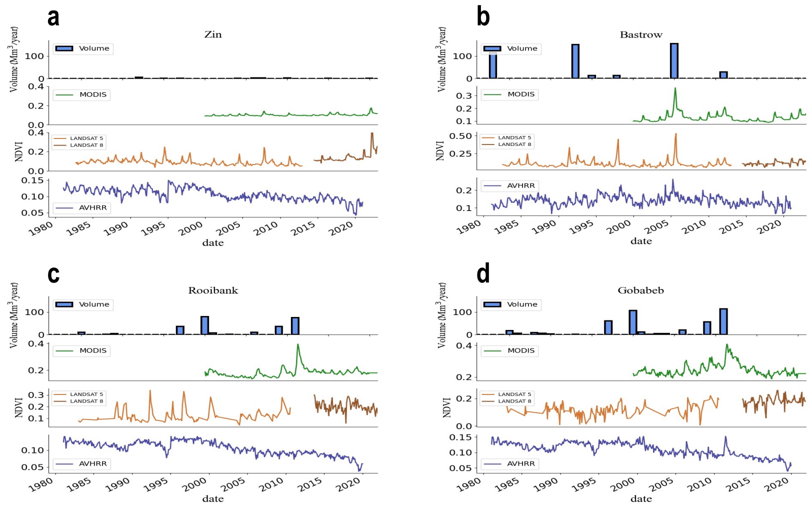
We present the Satellite Vegetation Index Time Series model for detecting historical floods in ungauged hyperarid regions (SatVITS-Flood). SatVITS-Flood is based on observations that floods are the primary cause of local vegetation expansion in hyperarid regions. To detect such expansion, we used two time series metrics: (1) trend change detection from the Breaks For Additive Season and Trend (BFAST-trend) and (2) a newly developed seasonal change metric based on Temporal Fourier Analysis (TFA) and the growing-season integral anomaly (TFA-GSIanom). The two metrics complement each other by capturing changes in perennial species following extreme, rare floods and ephemeral vegetation changes following more frequent floods. Metrics were derived from the time series of the normalized difference vegetation index (NDVI), the modified soil-adjusted vegetation index (MSAVI), and the normalized difference water index (NDWI), acquired from MODIS, Landsat, and AVHRR. The timing of the change was compared with the date of the flood and the magnitude of change with its volume and duration. We tested SatVITS-Flood in three regions on different continents with 40 years long, systematic, reliable gauge data. Our results indicate that SatVITS-Flood can predict flood occurrence with an accuracy of 78% and precision of 67% (Recall=0.69 and F1=0.68; p<0.01), and the flood volume and duration with NSE of 0.79 (RMSE=15.4 Mm3 event–1), and R2 of 0.69 (RMSE=5.7 days), respectively. SatVITS-Flood proved useful for detecting historical floods and may provide valuable long-term hydrological information in poorly-documented areas, which can help understand the impacts of climate change on the hydrology of hyperarid regions.

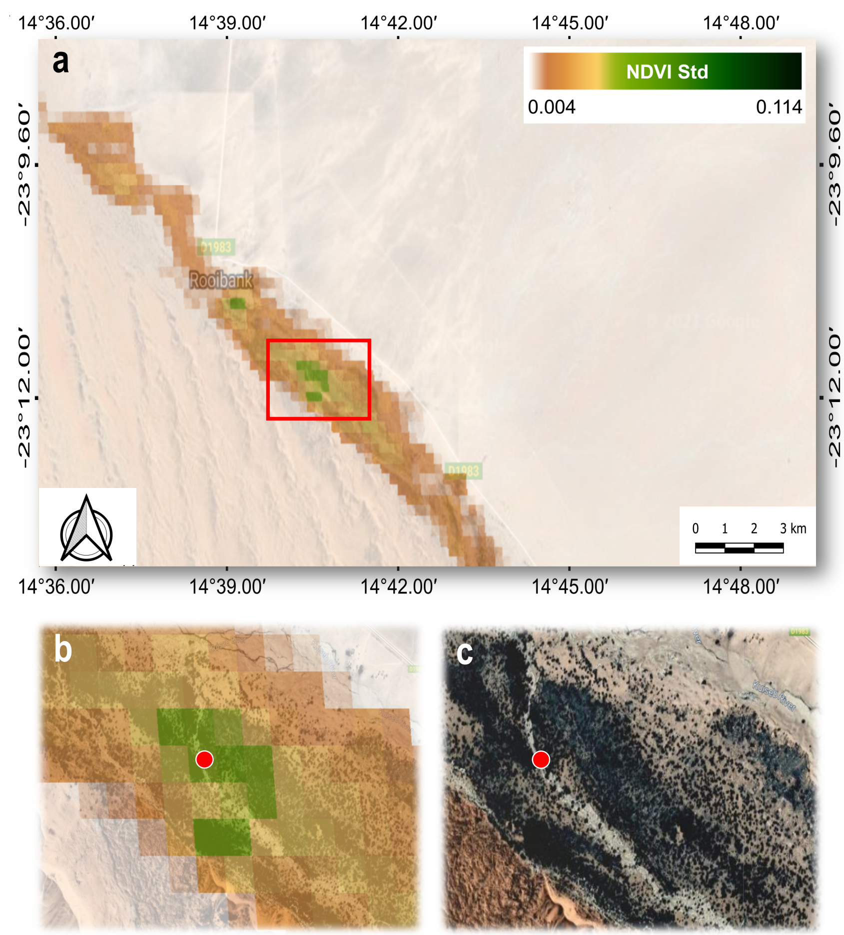
Hosted file

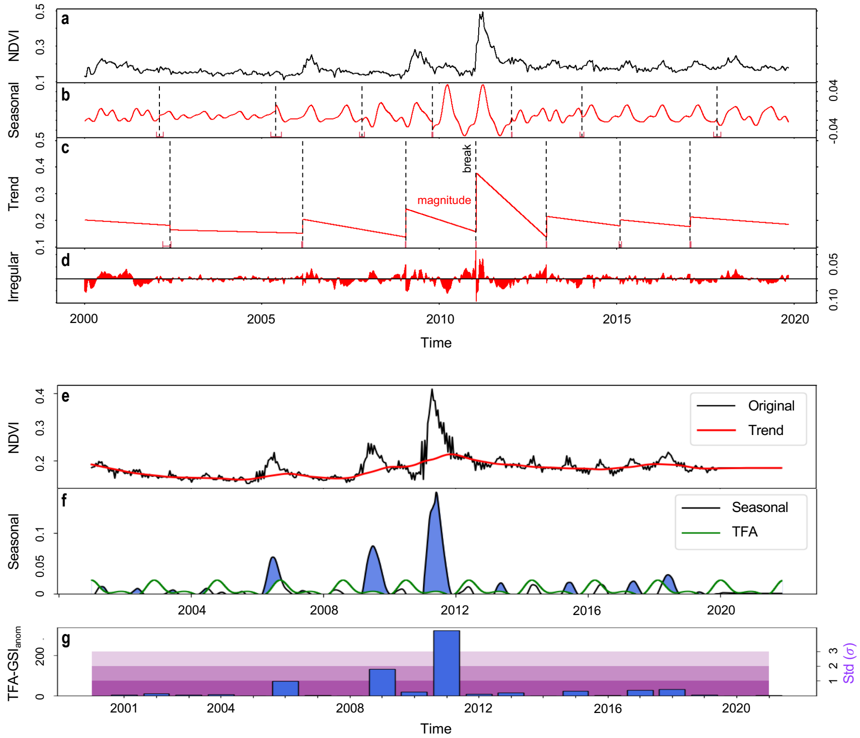
`961532_0_art_file_10911975_rtkpq3.docx` available at https://authorea.com/users/612348/articles/640299-satvits-flood-satellite-vegetation-index-time-series-flood-detection-model-for-hyperarid-regions


a
b
c
Q during floods
Bedrock
Recharge
Water table raising during flood
Saline water
Fresh water
Saline water
d
0.22
0.2
0.18
0.16
0.14
0.12
0.1
0.08
0.06
2010
2011
2012
2013
2014
2015
e
f
g

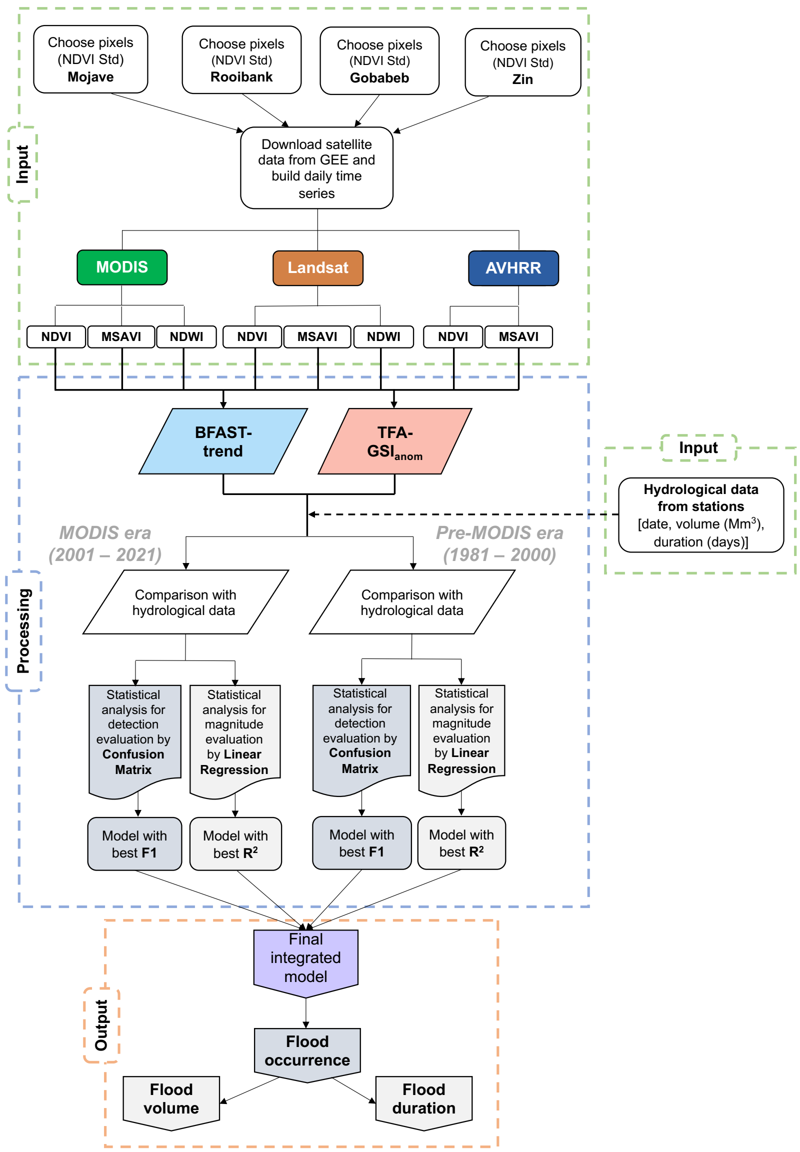

a
Mojave
Zin
Kuiseb
CHIRPS mean annual rainfall (mm)
b
Zin
Nahal Zin
c
Mojave
Bastrow
Mojave River
d
Kuiseb
Rooibank
Kuiseb River
Gobabeb

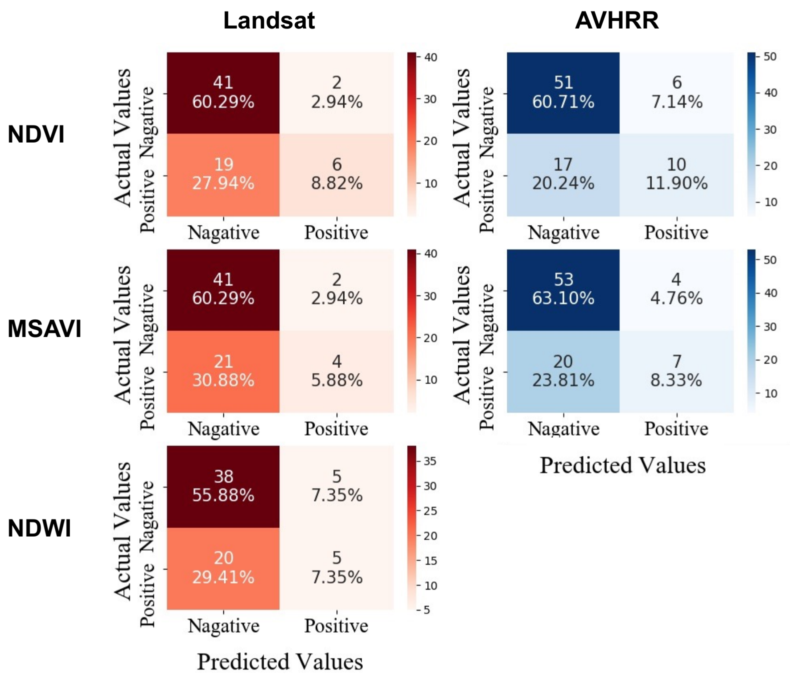

a
Zin
b
Bastrow
c
Rooibank
d
Gobabeb
Volume (Mm³/year)
Volume
MODIS
NDVI
LANDSAT 5
LANDSAT 8
AVHRR
date

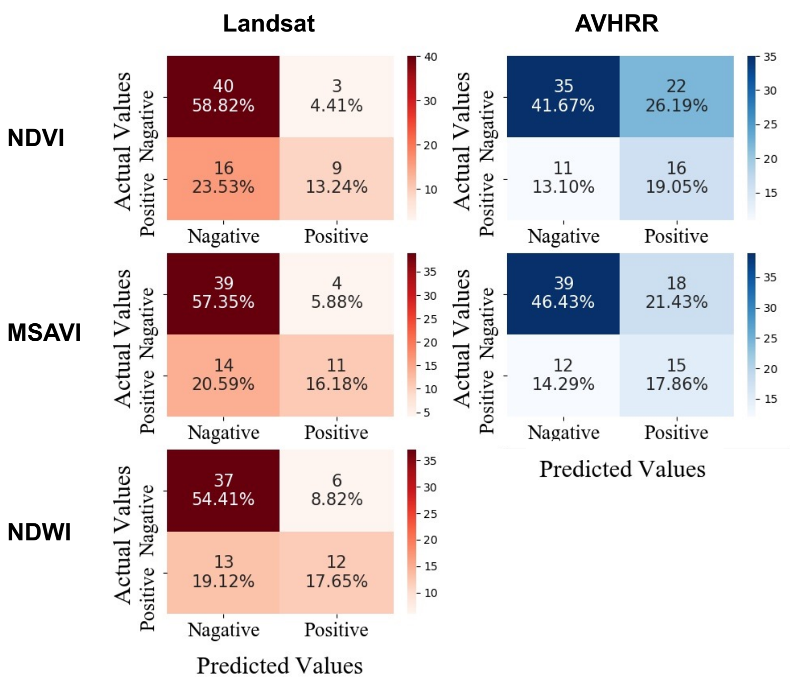

a
NDVI Std
0.004
0.114
14°36.00′
14°39.00′
14°42.00′
14°45.00′
14°48.00′
-23°9.60′
-23°12.00′
Rooibank
0 1 2 3 km
b
c

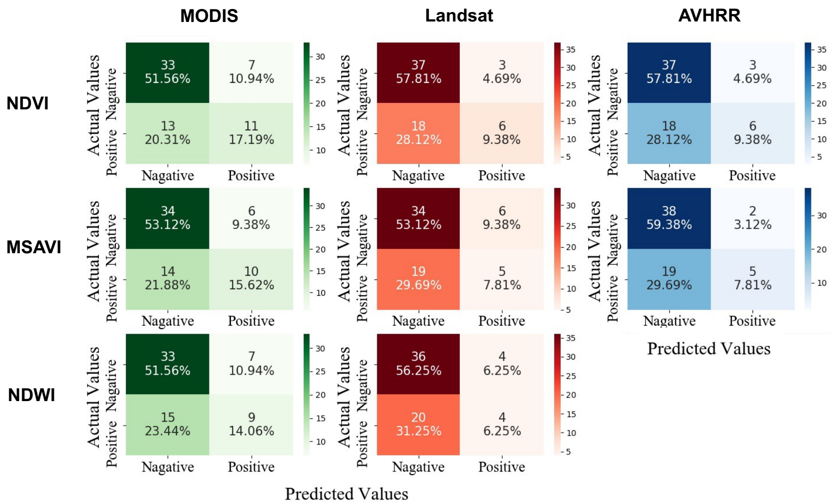


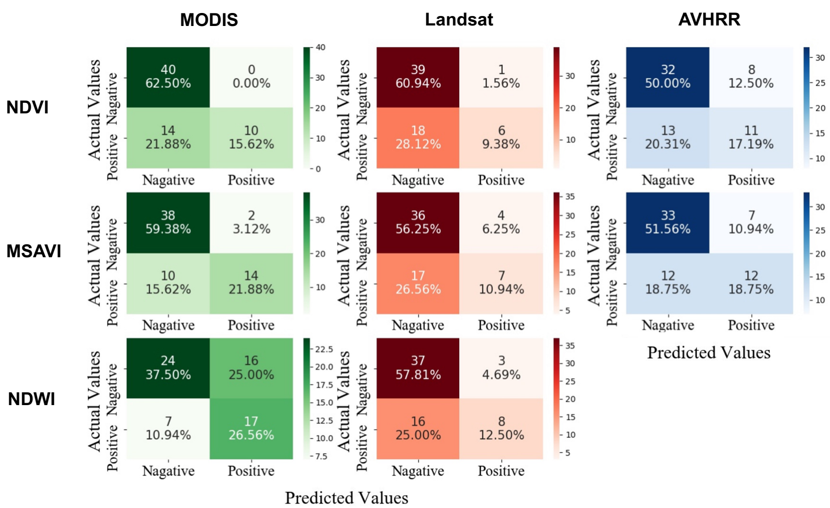
a
NDVI
b
Seasonal
c
Trend
break
magnitude
d
Irregular
Time
e
Original
Trend
f
Seasonal
TFA
g
TFA-GSI$_{anom}$
Std (σ)
Time


Input
Choose pixels (NDVI Std) Mojave
Choose pixels (NDVI Std) Rooibank
Choose pixels (NDVI Std) Gobabeb
Choose pixels (NDVI Std) Zin
Download satellite data from GEE and build daily time series
MODIS
Landsat
AVHRR
NDVI
MSAVI
NDWI
NDVI
MSAVI
NDWI
NDVI
MSAVI
Processing
BFAST-trend
TFA-GSI$_{anom}$
Input
Hydrological data from stations [date, volume (Mm^3), duration (days)]
MODIS era (2001 – 2021)
Pre-MODIS era (1981 – 2000)
Comparison with hydrological data
Comparison with hydrological data
Statistical analysis for detection evaluation by Confusion Matrix
Statistical analysis for magnitude evaluation by Linear Regression
Statistical analysis for detection evaluation by Confusion Matrix
Statistical analysis for magnitude evaluation by Linear Regression
Model with best F1
Model with best R^2
Model with best F1
Model with best R^2
Output
Final integrated model
Flood occurrence
Flood volume
Flood duration


Landsat
AVHRR
NDVI
MSAVI
NDWI
Actual Values
Nagative
Positive
41 60.29%
2 2.94%
19 27.94%
6 8.82%
51 60.71%
6 7.14%
17 20.24%
10 11.90%
41 60.29%
2 2.94%
21 30.88%
4 5.88%
53 63.10%
4 4.76%
20 23.81%
7 8.33%
38 55.88%
5 7.35%
20 29.41%
5 7.35%
Nagative
Positive
Predicted Values


Landsat
AVHRR
NDVI
MSAVI
NDWI
Actual Values
Nagative
Positive
40 58.82%
3 4.41%
16 23.53%
9 13.24%
35 41.67%
22 26.19%
11 13.10%
16 19.05%
39 57.35%
4 5.88%
14 20.59%
11 16.18%
39 46.43%
18 21.43%
12 14.29%
15 17.86%
37 54.41%
6 8.82%
13 19.12%
12 17.65%
Nagative
Positive
Predicted Values


MODIS
Landsat
AVHRR
NDVI
MSAVI
NDWI
Actual Values
Nagative
Positive
Predicted Values
33
51.56%
7
10.94%
13
20.31%
11
17.19%
37
57.81%
3
4.69%
18
28.12%
6
9.38%
37
57.81%
3
4.69%
18
28.12%
6
9.38%
34
53.12%
6
9.38%
14
21.88%
10
15.62%
34
53.12%
6
9.38%
19
29.69%
5
7.81%
38
59.38%
2
3.12%
19
29.69%
5
7.81%
33
51.56%
7
10.94%
15
23.44%
9
14.06%
36
56.25%
4
6.25%
20
31.25%
4
6.25%


MODIS
Landsat
AVHRR
NDVI
MSAVI
NDWI
Actual Values
Nagative
Positive
Predicted Values
40
62.50%
0
0.00%
14
21.88%
10
15.62%
39
60.94%
1
1.56%
18
28.12%
6
9.38%
32
50.00%
8
12.50%
13
20.31%
11
17.19%
38
59.38%
2
3.12%
10
15.62%
14
21.88%
36
56.25%
4
6.25%
17
26.56%
7
10.94%
33
51.56%
7
10.94%
12
18.75%
12
18.75%
24
37.50%
16
25.00%
7
10.94%
17
26.56%
37
57.81%
3
4.69%
16
25.00%
8
12.50%

a **Pre-MODIS**

		Landsat		AVHRR	
		BFAST	**GSI_{anom}**	**BFAST**	**GSI_{anom}**
NDVI	Accuracy	0.69	0.72	0.73	0.6
	Precision	**0.75**	**0.75**	0.62	0.42
	Recall	0.24	0.36	0.37	**0.59**
	F1	0.36	0.48	0.47	0.49
MSAVI	Accuracy	0.66	**0.73**	0.71	0.64
	Precision	0.67	0.73	0.64	0.45
	Recall	0.16	0.44	0.26	0.55
	F1	0.26	**0.55**	0.37	0.5
NDWI	Accuracy	0.63	0.72		
	Precision	0.5	0.66		
	Recall	0.2	0.48		
	F1	0.29	**0.55**		

b **MODIS**

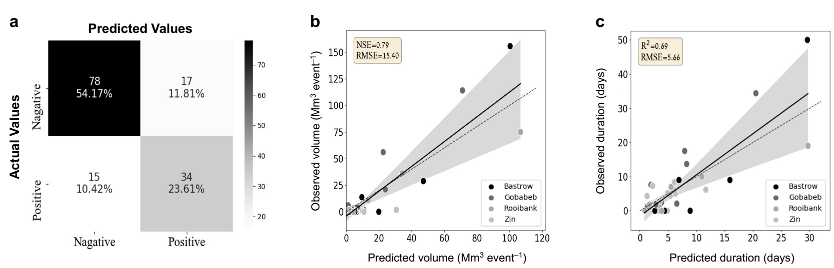
		MODIS		Landsat		AVHRR	
		BFAST	GSI_{anom}	BFAST	GSI_{anom}	BFAST	GSI_{anom}
NDVI	Accuracy	0.69	0.78	0.67	0.7	0.67	0.67
	Precision	0.61	**0.98**	0.67	0.85	0.67	0.57
	Recall	0.46	0.41	0.25	0.25	0.25	0.45
	F1	0.52	0.58	0.36	0.38	0.36	0.5
MSAVI	Accuracy	0.69	**0.81**	0.61	0.67	0.67	0.7
	Precision	0.62	0.87	0.45	0.63	0.71	0.63
	Recall	0.42	0.58	0.21	0.29	0.21	0.5
	F1	0.5	**0.7**	0.29	0.4	0.32	0.55
NDWI	Accuracy	0.66	0.64	0.62	0.7		
	Precision	0.56	0.51	0.5	0.72		
	Recall	0.38	**0.7**	0.17	0.33		
	F1	0.45	0.59	0.25	0.45		

a **Pre-MODIS**

			Landsat		AVHRR	
			BFAST	GSI_{anom}	BFAST	GSI_{anom}
NDVI	Volume (M m^3)	R^2	0.26	0.00	**0.43**	0.12
		RMSE	22.6	17.6	3.25	34.0
	Duration (days)	R^2	0.28	0.00	**0.28**	0.14
		RMSE	5.3	8.58	2.73	15.0
MSAVI	Volume (M m^3)	R^2	0.04	0.02	0.2	0.08
		RMSE	28.0	40.0	26.67	36.7
	Duration (days)	R^2	0.14	0.03	0.15	0.10
		RMSE	6.14	14.9	6.89	16.3
NDWI	Volume (M m^3)	R^2	0.42	0.04		
		RMSE	19.0	36.8		
	Duration (days)	R^2	0.21	0.04		
		RMSE	5.3	13.9		

b **MODIS**

			MODIS		Landsat		AVHRR	
			BFAST	GSI_{anom}	BFAST	GSI_{anom}	BFAST	GSI_{anom}
NDVI	Volume (M m^3)	R^2	0.64	0.26	0.22	0.53	0.46	0.36
		RMSE	26.5	41.8	9.55	35.0	40.2	34.6
	Duration (days)	R^2	0.53	0.18	0.25	0.57	0.51	0.39
		RMSE	8.9	12.4	2.33	10.3	11.5	10.2
MSAVI	Volume (M m^3)	R^2	0.52	0.51	0.28	0.38	0.07	0.10
		RMSE	31.9	31.43	37.5	35.2	56.6	41.0
	Duration (days)	R^2	0.41	0.41	0.37	0.42	0.05	0.10
		RMSE	10.2	9.85	10.8	10.5	16.2	12.1
NDWI	Volume (M m^3)	R^2	**0.75**	0.38	0.48	0.01		
		RMSE	16.5	27.6	14.3	44.2		
	Duration (days)	R^2	**0.62**	0.25	0.03	0.01		
		RMSE	5.7	9.8	10.5	14.5		



Hosted file

961532_0_supp_10918321_rtgm85.docx available at https://authorea.com/users/612348/articles/640299-satvits-flood-satellite-vegetation-index-time-series-flood-detection-model-for-hyperarid-regions

SatVITS-Flood: Satellite Vegetation Index Time Series Flood detection model for hyperarid regions

Omer Burstein[1,2], Tamir Grodek[3], Yehouda Enzel[3], David Helman[1,2,*]

1. Department of Soil and Water Sciences, Institute of Environmental Sciences, The Robert H. Smith Faculty of Agriculture, Food and Environment, The Hebrew University of Jerusalem, Rehovot 7610001, Israel
2. The Advanced School for Environmental Studies, The Hebrew University of Jerusalem, The Edmond J. Safra Campus, Givat Ram, Jerusalem 9190401, Israel
3. The Fredy and Nadine Herrmann Institute of Earth Sciences, The Hebrew University of Jerusalem, The Edmond J. Safra Campus, Givat Ram, Jerusalem 9190401, Israel

[*]Corresponding author: david.helman@mail.huji.ac.il

Abstract

We present the Satellite Vegetation Index Time Series model for detecting historical floods in ungauged hyperarid regions (SatVITS-Flood). SatVITS-Flood is based on observations that floods are the primary cause of local vegetation expansion in hyperarid regions. To detect such expansion, we used two time series metrics: (1) trend change detection from the Breaks For Additive Season and Trend (BFAST-trend) and (2) a newly developed seasonal change metric based on Temporal Fourier Analysis (TFA) and the growing-season integral anomaly (TFA-GSI_{anom}). The two metrics complement each other by capturing changes in perennial species following extreme, rare floods and ephemeral vegetation changes following more frequent floods. Metrics were derived from the time series of the normalized difference vegetation index (NDVI), the modified soil-adjusted vegetation index (MSAVI), and the normalized difference water index (NDWI), acquired from MODIS, Landsat, and AVHRR. The timing of the change was compared with the date of the flood and the magnitude of change with its volume and duration. We tested SatVITS-Flood in three regions on different continents with 40 years long, systematic, reliable gauge data. Our results indicate that SatVITS-Flood can predict flood occurrence with an accuracy of 78% and precision of 67% (Recall=0.69 and F1=0.68; p<0.01), and the flood volume and duration with NSE of 0.79 (RMSE=15.4 Mm^3 $event^{-1}$), and R^2 of 0.69 (RMSE=5.7 days), respectively. SatVITS-Flood proved useful for detecting historical floods and may provide valuable long-term hydrological information in poorly-documented areas, which can help understand the impacts of climate change on the hydrology of hyperarid regions.

Keywords: BFAST; flood; hyperarid; NDVI; satellite; vegetation index

1. Introduction

Occasional, rare floods in hyperarid regions are a vital water source for sustaining human, flora, and fauna life. Commonly generated due to their respective wetter, higher-elevation headwaters during the characteristic short rainy seasons, these flash floods flow along channels that cross deserts, transmit their discharge into local shallow aquifers (e.g., Enzel et al., 1993; Enzel and Wells, 1997), and form linear oases that provide year-round available water for flora, fauna, and human use. The vegetation along these elongated, shallow-aquifer-fed oases (Grodek et al., 2020) includes mainly drought-tolerant species and trees characteristic of wetter origin (Verdugo-Vásquez et al., 2021). The occasional flash floods replenish the groundwater and lead to a thriving vegetation growth (Grodek et al., 2020). The flood duration and magnitude are essential parameters for the recharge process (Dahan et al., 2008; Enzel et al., 1989; Enzel & Wells, 1997; Morin et al., 2009).

Managing water resources in hyperarid regions, however, is difficult due to the lack of gauge stations and reliable flood data (Benito et al., 2010). Obtaining such data is essential to support planning water resources, identifying areas that support plant and animal life, and predicting the impact of climate change on water availability (Zaman et al., 2012). Understanding the hydrology of such streams can improve water management and conservation decisions, ensuring the sustainability of ecosystems and human well-being in these sensitive watersheds (Chehbouni et al., 2008).

Different methods have been deployed to acquire hydrological information from ungauged areas. These include the use of numerical physical models (e.g., Puricelli et al., 2009), statistical techniques (e.g., Alfieri et al., 2013; Bonakdari et al., 2019), Machine Learning (ML) and Artificial Intelligence (AI) methods (e.g., Gizaw and Gan, 2016; Meresa, 2019; Pyayt et al., 2011; Shahabi et al., 2020; Xie et al., 2021), and remote sensing information acquired through sensors onboard drones and satellites (e.g., Cian et al., 2018; Gao et al., 2017; Li et al., 2015; Martinez and Le Toan, 2007; Wang et al., 2011; Xinyi et al., 2017; Yang et al., 2020). However, while physical (numerical) and empirical (statistical and ML/AI-based) models can provide valuable information on past floods and be applied in forecasting future floods, they are often heavily based on detailed weather data. This limits their use in climate-hydrology research because of the interdependency of the predicted (hydrological variables) and the climate/weather variables.

In contrast, remote sensing can provide direct, independent information on floods. Satellite imagery and aerial photography from planes and drones have been used to monitor past and real-time environmental and hydrological changes (Iqbal et al., 2023; Manfreda et al., 2018; Yang et al., 2020). The main advantage of satellites, however, is the large areal coverage, including remote regions. These, with their high temporal resolution, are, therefore, an ideal tool for monitoring floods in hyperarid regions. With roughly 50 years of Earth Observing Systems (EOS), satellites have provided indispensable information about the Earth's surface (Boyle et al., 2014).

Indeed, satellite information has been extensively used to study hydrological systems. Image analysis and advanced statistical tools have been deployed to acquire information and map flood inundation areas. For example, Dao et al. (2019) used a MODIS-Landsat image fusion with object-based image analysis to detect flood inundation in a heterogeneous vegetated area (Dao et al., 2019). Vekaria et al. (2022) used multi-temporal Sentinel-1 SAR images to detect floods in the Brahmaputra River, Assam (India) (Vekaria et al., 2022), and Tripathi et al. (2020) mapped flood inundation using multi-temporal optical and SAR satellite data in Darbhanga District, Bihar, India (Tripathi et al., 2020). DeVries et al. (2020) used Sentinel-1 and Landsat to monitor flood events on Google Earth Engine platform (DeVries et al., 2020). McCormack et al. (2022) used Sentinel-1 SAR imagery to map annual floods in Ireland from 2016 to 2021 (McCormack et al., 2022). Others combined satellite data with machine-learning methods to detect floods. For example, Rahman et al. (2021) used stacking hybrid machine-learning algorithms with radar and optical satellite data to detect floods in Bangladesh (Rahman et al., 2021), while Shahabi et al. (2020) used data from Sentinel-1 in ensemble models based on bagging as a meta-classifier and K-Nearest Neighbor (KNN) coarse, cosine, cubic, and weighted base classifiers to forecast flooding in a watershed in northern Iran (Shahabi et al., 2020).

These are only a few examples of the use of satellite data and methods in detecting floods, either for real-time or post-flood detection and mapping. However, some of these methods are based on hydrological data such as precipitation or soil humidity. Others are more oriented toward perennial rivers and are thus less effective for mapping floods in hyperarid ephemeral river channels. Recently launched satellites, such as Sentinel, TerraSAR-X, and COSMO-SkyMed, may be more suitable for these channels (Dasgupta et al., 2018; Grimaldi et al., 2020; Shen et al., 2019). Yet, their use is limited to detecting recent years' floods, which limits our ability to use this in climate-hydrology research that requires historical information (flood information for at least >30 years).

Here we propose a flood detection model based on time series analysis metrics of satellite-derived vegetation index (VI), explicitly focusing on hyperarid regions, combining four decades of information acquired via different satellite programs (Landsat, AVHRR, and MODIS). The proposed method has the advantage of providing hydrological information on flood occurrences (or frequency) and estimating their respective flood discharge, independent of weather or climate data or hydrological model. The analyses are conducted, first, for hyperarid streams with available gauged data to test the success of the method. Specifically, our study aims to fill the gap in the existing long-term historical data on floods in hyperarid regions and to contribute to understanding the hydrological cycle in such environments to improve water management and conservation decisions

2. The logic underlying the use of VI time series for flood detection in hyperarid regions

The general idea that floods in hyperarid regions can be detected using satellite-derived spectral-based indices leans on two main observations: (1) floods have an indirect impact on the riverbank vegetation by recharging the shallow aquifers and improving water quality (Dahan et al., 2008; Grodek et al., 2020) and (2) the changes in the vegetation cover, vigor, and growth, of both ephemeral (mainly annual herbaceous) and perennial (mainly evergreen, woody) plant species, can be detected via spectral-based vegetation indices from satellites (Grodek et al., 2020; Moses et al., 2021; Normandin et al., 2022).

Figure 1 illustrates an example of the indirect effect of a flood on the riverbank vegetation in a hyperarid region and how such an effect is detected in a time series of a vegetation index (VI).

[Figure 1]

First, the floodwater reaches the stream (Fig. 1a,b), replenishing the shallow aquifers along the route and raising its water table (Fig. 1c) (Enzel & Wells, 1997; Greenbaum et al., 2001; Morin et al., 2009). The new freshwater makes the local vegetation flourish and expand (Fig. 1e,f). At the same time (though sometimes with some lag of a few weeks or months), new annual species appear in the understory or around the woody vegetation, and germination of the woody species occurs (Fig. 1f). The annual vegetation, however, does not last for long due to the harsh dry, saline conditions and the shallower root system of these species.

The flourishing of the annual vegetation is well noticed in the VI time series through a gradual increase followed by a slow decrease to the baseline values at the end of the season (Fig. 1d). In contrast, the perennial (usually evergreen woody) vegetation has only moderate effect on the VI signal, with a slight increase in the baseline values, which can be noticed usually only after a few years or by comparing the signal with the signal from past years. Such a change is relatively more stable and lasts longer than the seasonal change of the ephemeral species, sometimes even for several years (Fig. 1d). The slow return to the original baseline values, or even lower values (e.g., point g in Fig. 1d), is noticeable when there no new water enters the system (i.e., during a few years without flood events). The deeper root system of the woody, drought-tolerant species allows them to use water even when the water table drops to lower levels following numerous years of drought.

The VI increase following the flood may indicate more than the occurrence of a flood and can provide additional hydrological information, such as the flood volume and duration (Grodek et al., 2020). This is because the volume and duration of floods are assumed to be correlated with vegetation expansion through more extensive growth and for an extended period of time when the flood volume is higher and its duration longer than when it is low and lasts for a shorter period of time.

Based on the abovementioned assumptions and previous observations (Grodek et al., 2020; Moses et al., 2021; Normandin et al., 2022), we used here satellite-derived VI time series metrics to develop the Satellite VI Time Series Flood detection model for hyperarid regions (SatVITS-Flood). Section 3 presents the data used to develop and evaluate the model. Section 4 presents the SatVITS-Flood model scheme and evaluation method.

3. Data sources

3.1. Study area and hydrology

To develop and test SatVITS-Flood, we selected four sites from three different regions on three continents (Fig. 2a). The four sites are Zin River (Israel) (Fig. 2b), Barstow reach of the Mojave River, California (USA) (Fig. 2c), and Rooibank and Gobabeb sites along the Kuiseb River (Namibia) (Fig. 2d). The primary criteria for choosing these sites were the availability and reliability of long-term hydrological data, for at least four decades. For example, the hydrological data from the Kuiseb River were collected through field expeditions as part of a long-term study conducted in Namibia by some of the co-authors. Thus, much information exists regarding these two sites. Data for Zin River (hereafter, Nahal Zin; Nahal is the Hebrew word for stream, either perennial or ephemeral) was collected through collaboration with local researchers from Israel and the Water Authority of Israel gauge stations, which is deemed to be very reliable. Mojave River was thoroughly studied by Y. Enzel (Enzel, 1992; Enzel et al., 1989) for the last decades (see also Enzel et al., 1989; Enzel and Wells, 1997).

[Figure 2]

At each site, we used systematic gauging station data after careful screening for errors (see the location of the stations in Fig. 2b-d). The annual flood volume (Mm^3; Fig. 3a-d) and the flood duration (days) of each event were derived, and the onset date of the flood was recorded. Following is a more detailed description of the investigated sites.

[Figure 3]

3.1.1. Zin River (Nahal Zin)

Nahal Zin is an ephemeral river in the central Negev desert, southern Israel. It heads at the northern flank of the Negev highland, at ca. 1,000 meters above sea level (m.a.s.l.), and ends in the Dead Sea, at ~-400 meters below sea level. It flows along 125 km with a mean slope of 1.12 % and drains an area of 1,400 km^2 (Greenbaum et al., 2000). The mean annual rainfall at the Zin headwaters is 90 mm $year^{-1}$, and 50-60 mm $year^{-1}$ at its terminus at the Dead Sea (Greenbaum et al., 1998). Flash floods occur up to a few times a year as a result of individual heavy rainstorms or prolonged storms associated with Active Red Sea Troughs, eastern Mediterranean low-pressure systems, or tropical plumes (Armon et al., 2018, 2019; Kahana et al., 2002). The shallow alluvium covering the riverbed is recharged by floodwater, feeding several springs (e.g., Ein Zin, Ein Aqrabim, Ein Avdat, and others; e.g., Greenbaum et al., 2000).

Data for Nahal Zin (Zin Elyon gauge station) were provided by the Israel Water Authority (Israel Water Authority, 2020). The station drains 135 km^2 at the headwaters and is located at 550 m.a.s.l. (Fig. 2b).

3.1.2. Mojave River

The Mojave River (Mojave Desert, southern California) heads on the northeastern San Bernardino Mountains at ca. 900 m.a.s.l., where the mean annual precipitation is 1000 mm year^{-1}; yet, around 90% of the watershed receives only <150 mm year^{-1} (∼93 mm year^{-1} in Barstow; Enzel, 1992). The floods flow for ∼180 km through Victorville and Barstow to Soda and Silver Lake playas (area of 9,500 km^2; 285 m.a.s.l.). The Mojave River is an intermittent river in which most floodwater is transmitted to the channel recharging the shallow alluvial aquifer. Therefore, only extreme floods reach the river mouth at the Silver Lake playa (Enzel et al., 1989; Enzel & Wells, 1997). The recharge of the alluvial aquifers along the river raises the water table, causing subsurface flows and, in several reaches, also surface base flow. Water extraction from the aquifers along the river since the early 1900s caused pronounced drops in water levels after floods and more pronouncedly in the long-term (Izbicki et al., 2004; Scott et al., 2000).

Data for gauge station number 10262500 (Fig. 2c) at Barstow, with a 637 m.a.s.l. and a watershed area of ∼3,340 km^2, were provided by the United States Geological Survey (USGS, 2023).

3.1.3. Kuiseb River (Rooibank and Gobabeb)

The Kuiseb River in western Namibia heads on the wetter Khomas Highland, a high plateau west of Windhoek, at ca. 2,000 m.a.s.l. The Kuiseb River flows across the Great Escarpment and the Namib sand sea into the Atlantic Ocean along 480 km, with a mean slope of 0.42% m km^{-1} (Morin et al., 2009) and a catchment area of 16,800 km^2. In Windhoek (located just outside of the Kuiseb River headwater), the mean annual precipitation over the last 130 years is 355 mm year^{-1}, whereas in Walvis Bay, on the coast of the Atlantic Ocean, the annual mean is only 9 mm year^{-1} (Grodek et al., 2020).

The Kuiseb River is an ephemeral river responding to heavy rainstorms at the wetter headwaters. The floods recharge the shallow alluvial aquifers along the hyperarid river reaches and supply essential water to sustain the green belts of lush oases downstream along the Namib desert. Only extreme floods reach the Atlantic Ocean, whilst most of them are blocked by the dunes of the Namib Sand Sea or infiltrate their route to the ocean (Grodek et al., 2020).

The hydrologic data were obtained from Morin et al. (2009) that corrected and verified the original hydrometric systematic data record. Specifically, we focused on the Gobabeb site (405 m.a.s.l., 14,300 km^2) and the Rooibank site (124 m.a.s.l., 16,400 km^2; Lehner and Grill, 2013). The distance between Gobabeb and Rooibank is ∼65 km (Fig. 2d).

3.2. Satellite data

We used data from different instruments onboard three satellite programs – Terra/Aqua MODerate resolution Imaging Spectroradiometer (MODIS), Landsat satellites program (Landsat 5 and 8), and the Advanced Very High-Resolution Radiometer (AVHRR) instrument onboard the National Oceanic and Atmospheric Administration (NOAA) family of polar-orbiting platforms (POES) and European MetOp satellites.

From each, we downloaded reflectance data or specific VI products (see Table 1 and more details in the following *Section 3.3*) using JavaScript codes in the Google Earth Engine (GEE) platform (Gorelick et al., 2017). Time series of vegetation indices indicative of various aspects of vegetation dynamics (e.g., Helman, 2018; Helman and Mussery, 2020) were constructed. Linear interpolation was used to resample the data to a daily timescale when daily data was unavailable (e.g., Helman et al., 2017). Because of the different periods of data availability of these satellite programs (AVHRR is available from 1981 and MODIS only from 2000. First Landsat images were acquired in the mid-70s, but reliable data, from the Landsat-5 mission, are available only from 1984; Table 1), we divided the datasets into two periods: for the pre-MODIS era, 1981 – 2000, and the MODIS era, 2001 – 2021. The models were independently trained on each of these 20-year-long intervals to overcome biases due to discrepancies in data size. Models were then integrated into a single predictive model for the entire 1981 – 2021 period (see Section 4 for more details).

Table 1. Main characteristics of the satellite programs and sensors and the vegetation indices derived from each sensor for this study.

Short name	**Satellites**	**Sensor/s**	**Period**	**Temporal resolution**	**Spatial resolution**	**Index**	**Product**
MODIS	Terra, Aqua	The MODerate resolution Imaging Spectroradiometer (MODIS)	2000 – 2021	16-day for NDVI (available twice a day) Daily for NDWI / MSAVI	250 m for NDVI / MSAVI 500 m for NDWI	NDVI, MSAVI, NDWI,	MOD13Q1 for NDVI MOD09GQ for MSAVI MCD43A4 for NDWI
Landsat	Landsat 5, Landsat 8	Multispectral Scanner (MSS) and Thematic Mapper (TM) for Landsat 5, Operational Land Imager (OLI) for Landsat 8	1984 – 2012 for Landsat 5 2013 – 2021 for Landsat 8	16-day	30 m	NDVI, MSAVI, NDWI	LT05/C01/T1_TOA for Landsat 5, LC08/C01/T1_TOA for Landsat 8
AVHRR	POES and MetOp satellites	The Advanced Very-High-Resolution Radiometer (AVHRR)	1981 – 2021	Daily (available 4 times a day)	5566 m	NDVI, MSAVI	NOAA/CDR/AVHRR/NDVI/V5 for NDVI NOAA/CDR/AVHRR/SR/V5 for MSAVI

3.3. Vegetation indices (VIs)

Three VIs were derived from the satellite data: (i) the normalized difference vegetation index (NDVI), (ii) the modified soil-adjusted vegetation index (MSAVI), and (iii) the normalized difference water index (NDWI). Since AVHRR does not have information at the shortwave infrared wavelength (SWIR), we derived NDWI only for MODIS and Landsat. The idea underlying the use of these three VIs is to exploit the strengths and advantages of each index, which may complement each other by providing unique information related to their specific characteristics. Following is a concise description of each VI and its main characteristics.

3.3.1. NDVI

NDVI is the most commonly used VI (see Box 1 in Helman, 2018). It is based on the reflectance at the red wavelength range (ρ_{Red}), usually from 620 nm – 700 nm, and the near-infrared (NIR) range (ρ_{NIR}), around 780 nm – 900 nm (Rouse et al., 1974):

$$\mathrm{NDVI} = \frac{\rho_{\mathrm{NIR}} - \rho_{\mathrm{Red}}}{\rho_{\mathrm{NIR}} + \rho_{\mathrm{Red}}} \qquad (1)$$

This index ranges between −1 and +1 for non-vegetated and fully vegetated surfaces, respectively, with negative values usually corresponding to watered surfaces and low positive values for bare ground or sparsely vegetated areas. It has been shown to correlate well with plant chlorophyll content, leaf expansion, canopy structure, and other plant-related characteristics (e.g., Carlson and Ripley, 1997; Gamon et al., 1995; Glenn et al., 2008).

We used the 250-m 16-day NDVI product of MODIS (MOD13Q1), which uses reflectance at the wavelength range of 645 nm (sur_refl_b01) and 858 nm (sur_refl_b02) for the red and NIR bands, respectively. For Landsat, we used the ranges of 630–690 nm for Landsat 5 (B3) and 640–670 nm for Landsat 8 (B4) for the red band, and 760–900 nm and 850–880 nm for Landsat 5 (B4) and Landsat 8 (B5) for the NIR to calculate the NDVI. For AVHRR, we used 640 nm for the red band and 860 nm for the NIR (SREFL_CH1 and SREFL_CH2, respectively) from the NOAA/CDR/AVHRR/SR/V5 product to derive NDVI.

Figure 3a-d present examples of NDVI time series of selected pixels in the four study sites derived from the three satellite programs.

3.3.2. MSAVI

The modified version of the Soil Adjusted Vegetation Index (MSAVI) was designed to better reduce the influence of bare soil reflectance on the vegetation signal by replacing SAVI's soil adjustment factor *L* with a soil brightness correction factor (Qi et al., 1994). MSAVI was proved more reliable than SAVI in adjusting for soil influences (Abderrazak et al., 1996; Qi et al., 1994). This is because, unlike SAVI, which requires prior knowledge about the vegetation cover and density to select the optimal *L*, MSAVI uses the slope of the soil line from a plot of red versus NIR brightness values. The formula for generating MSAVI can be then simply reduced to (Qi et al., 1994):

$$\mathrm{MSAVI} = \frac{2\,\rho_{\mathrm{NIR}} + 1 - \sqrt{(2\,\rho_{\mathrm{Red}} + 1)^2 - 8\,(\rho_{\mathrm{NIR}} - \rho_{\mathrm{Red}})}}{2} \qquad (2)$$

Although many variations of soil-adjusted indices exist (e.g., OSAVI, TSAVI, etc.), MSAVI is considered the most suitable for monitoring vegetation changes in hyperarid regions where the vegetation cover is usually less than 25% (Rondeaux et al., 1996).

We used the red and NIR ranges from MODIS, 620–670 nm (sur_refl_b01) and 841–876 nm (sur_refl_b02), respectively, both from the daily 250 m MOD09GQv006 product, in equation (2). The same bands from the same products as in NDVI were used to derive MSAVI from Landsat and AVHRR.

3.3.3. NDWI

NDWI is a spectral index based on the reflection at the NIR range and the short wave infrared (SWIR) range, 1400–3000 nm (B. Gao, 1996):

$$\mathrm{NDWI} = \frac{\rho_{\mathrm{NIR}} - \rho_{\mathrm{SWIR}}}{\rho_{\mathrm{NIR}} + \rho_{\mathrm{SWIR}}} \qquad (3)$$

NDWI values vary between –1 and +1 depending on the leaf water content, vegetation type, and cover. Although some other variations of this index exist (e.g., Azar et al., 2023; Helman and Mussery, 2020), this version is most robust for detecting vegetation water content because of its use of the reflectance at the SWIR bands in Eq. 3, which is known to be sensitive to changes in mesophyll water content (Ceccato et al., 2001; Tucker, 1980). Since the water use of the vegetation in hyperarid regions is typically low, we expect NDWI to be highly sensitive to new water supply by floods.

To derive NDWI, we used the NIR and SWIR ranges, 841–876 nm (Nadir_Reflectance_Band2) and 1628–1652 nm (Nadir_Reflectance_Band6), from the 16-day 500 m composite MODIS product MCD43A4v006. For Landsat 5 and 8, we used the SWIR bands of 1550–1750 nm (B5) and 1570–1650 nm (B6), respectively. The same NIR bands as in NDVI and MSAVI were used for Landsat. We did not derive NDWI for AVHRR since it does not have reflectance at the SWIR band.

4. The SatVITS-Flood model scheme and evaluation method

SatVITS-Flood is based on specific change detection (and magnitude of change) in the VI time series. However, to build the model, our first challenge was to assess specific areas where we expect the maximal effect of the flood on the VIs. To do so, we generated a map of the seasonal standard deviation of NDVI (NDVI Std; Fig 4) on GEE, as NDVI was shown to be high sensitivity to small changes in vegetation cover and growth in dry environments (Grodek et al., 2020; Helman et al., 2014; Helman et al., 2014a; Helman and Mussery, 2020).

[Figure 4]

Pronounced irregular changes (large NDVI Std) are assumed to be caused by the impact of floods unless the area is irrigated. Thus, pixels with high standard deviations are interpreted as strong vegetation responses to water supply (whether by transmission loss, groundwater recharge, or manmade irrigation). Next, we inspected the map and selected the closest pixels to the riverbank with the highest NDVI Std, identified as non-agricultural pixels (visually inspecting the high spatial resolution RGB images on GEE; Fig 4). The time series of the three VIs were then generated for these pixels for each site (Fig 3a-d). Each time series was analyzed for changes using two methods (see explained methods in *Section 4.1*) to detect the timing and magnitude of these changes. The date of change and its magnitude were compared with the known year of occurrence and volume and duration of the flood for each site.

4.1. Two approaches of abrupt change detection in time series

To detect abrupt changes in the time series of the VIs, two complementary approaches were used: (1) the Breaks For Additive Season and Trend (BFAST; Fig. 5a-d) and (2) a newly developed seasonal change metric, based on Temporal Fourier Analysis (TFA) and the calculation of the growing-season integral anomaly of the VI (TFA-GSI_{anom}; Fig. 5e-g). The use of the BFAST method is aimed to point at abrupt changes in the baseline of the time series (BFAST-trend), which would correspond mainly to changes in the perennial vegetation (mostly evergreen woody species) (Helman, 2018; Helman et al., 2015). In contrast, TFA-GSI_{anom} provides the change in the seasonal signal, which corresponds with changes in the ephemeral species (David Helman et al., 2015). Below is a description of the two methods.

4.1.1. BFAST-trend

BFAST is an iterative algorithm developed to detect abrupt changes in the time series (Verbesselt, Hyndman, Newnham, et al., 2010; Verbesselt, Hyndman, Zeileis, et al., 2010). It first decomposes the time series into seasonal, trend, and residual components (Fig. 5b,c,d) and then applies time series analysis techniques to detect abrupt changes in the trend and seasonal signals, delivering the timing of the change and its magnitude. Here we used BFAST only to detect changes in the trend (BFAST-trend; Fig. 5c). Abrupt changes in the VI seasonal component are mostly negative (abrupt drop in values), meaning some damage to the ephemeral vegetation (usually herbaceous), which can be a result of, for example, fire or other disturbances (Verbesselt et al., 2012).

Our goal is to search for a positive change due to water supply by floods. Therefore, the change detection method for the seasonal component should be focused on a positive anomaly relative to past ‘regular’ years (i.e., no flood years). The TFA-GSI_{anom} developed in this study, explained in the following subsection, is more suitable for that purpose.

BFAST is considered robust against noise, not influenced by changes in the amplitude of the seasonal component, which confirms its applicability to time series with varying noise and seasonal amplitudes. It can be applied to any time series data without the need to select a specific reference period or define a change trajectory and has been widely used with remote sensing data to identify vegetation cover changes, near real-time ecosystem disturbances, and even streamflow abrupt changes (Fang et al., 2018; Mardian et al., 2021; Verbesselt et al., 2012; Xu et al., 2022; Zhao et al., 2015).

[Figure 5]

Here we used different parameters for BFAST. First, we defined the h parameter determining the minimum time between breakpoints. Optimizing this h parameter is particularly important to enable the detection of as many possible extreme changes in the VI, potentially caused by flood events. At the same time, it should also be optimized to prevent change detection due to noise (Watts & Laffan, 2014), as those could increase the model's false positive detection rate. Accordingly, we adopted an h-parameter that yields a minimum laps time of two years between breakpoints, as recommended in previous studies (Bai & Perron, 2003; Fang et al., 2018; Verbesselt, Hyndman, Zeileis, et al., 2010). Because each satellite's data set covers a slightly different period, we determined a different h value for each data set: 0.1 for MODIS, 0.08 for Landsat, and 0.06 for AVHRR. This was done by a trial-and-error process in which we inspected each time series of hydrological and VI data. We chose the "harmonic" seasonal model because it is considered to be more adapted to natural vegetation phenological changes rather than the "dummy" seasonal model, which is more suitable for changes in crops (Verbesselt, Hyndman, Newnham, et al., 2010).

The magnitude of the break in the BFAST-trend was used to derive the flood volume and duration by simply correlating it with the gauge-derived hydrological data. Both the timing and magnitude of change in the BFAST-trend were derived using the bfast R package (Verbesselt, Hyndman, Newnham, et al., 2010; Verbesselt, Hyndman, Zeileis, et al., 2010), freely available at https://search.r-project.org/CRAN/refmans/bfast/html/bfast.html.

4.1.2. TFA-GSI_{anom}

The time series metric of TFA-GSI_{anom} allows for identifying positive seasonal anomalies related to ephemeral vegetation based on the correlation between the integral over the seasonal VI signal and the biomass of such vegetation (Helman et al., 2014; Helman et al., 2014a; Helman and Mussery, 2020). We expect a significant positive deviation from ephemeral plants' regular mean annual biomass following flood events. We also expect such deviation to correlate with the flood event intensity (Grodek et al., 2020). Thus, to develop a robust metric, we follow the subsequent steps:

(1) We first extracted the long-term trend of the VI time series (red line in Fig. 5e) by using the LOcal WEighted Scatterplot Smoothing (LOWESS) technique (Cleveland, 1979) with a window of w = 0.1, which corresponds to 10% of the data (e.g., Bianchi et al., 1999).

(2) To eliminate noise in the original VI time series related to radiometric issues and/or atmospheric conditions, we used LOWESS again but with a narrower window of w = 0.02, which corresponds to 2% of the data (e.g., Helman et al., 2019, 2017).

(3) We subtracted the trend from the smoothed time series, remaining only with the seasonal signal (black line in Fig. 5f).

(4) We applied a Temporal Fourier Analysis (TFA), which derives the mean annual cycle of the seasonal signal in the time series. The TFA is based on the sum of cosine and sine series with different amplitudes and phases corresponding to well-known environmental cycles (Lensky & Dayan, 2011). In this case, the TFA-based seasonal VI signal (the green line in Fig. 5f) is assumed to describe the average expected growth and senescence cycle of ephemeral plants, considering conditions of most years. Here we used the harmonics that correspond to the annual, biennial, and triennial cycles of seasonal changes, which are the most significant to the biological periodicity (Blum et al., 2013; Lensky & Dayan, 2011; Scharlemann et al., 2008).

(5) The next step was subtracting the seasonal VI from the TFA seasonal signal for each year, considering only positive values (the blue area underneath the curve in Fig. 5f).

(6) We then calculated the integral over the area, computing the growing season integral (GSI) anomaly per year (blue bars in Fig. 5g).

(7) Finally, we considered anomalous TFA-GSI_{anom} values to be those that exceed at least one standard deviation ($+1\sigma$) (horizontal purple bars in Fig. 5g).

Years with an anomalous TFA-GSI_{anom} value of $>1\sigma$ were classified as flood years. As in BFAST-trend, the TFA-GSI_{anom} value was used to derive the flood volume and duration by simply correlating it with the gauge-derived hydrological data.

The Python package statsmodels 0.13.5 (Seabold & Perktold, 2010) was used to smooth the time series with LOWESS (statsmodels.nonparametric.smoothers_lowess.lowess) and the package scipy.integrate was used to calculate the integral over the seasonal signal.

4.2. Independent models and model integration

Since we have two time-series metrics (BFAST-trend and TFA-GSI_{anom}), two VIs (NDVI and MSAVI), and two satellites (Landsat and AVHRR), plus two metrics for one VI (NDWI) and one satellite (Landsat) for the pre-MODIS era, this gives us a total of 10 models for the pre-MODIS period. For the MODIS period, we have two metrics, two VIs (NDVI and MSAVI), and three satellites (including MODIS), plus two metrics, one VI (NDWI), and two satellites (MODIS and Landsat), which gives us a total of 16 additional models.

In total, there are 26 models for flood detection, 26 for flood volume, and 26 for flood duration models for the entire period of 1981-2021. From these models, we had to choose the two best-performing models for flood detection: one for the pre-MODIS period and one for the MODIS period. Then, based on the detected floods from these two "best" models, we developed linear regression models by correlating the magnitude of the floods from the 26 BFAST-trend and TFA-GSI_{anom} models with the flood volume and duration. The idea was to again choose the two best models for flood volume and duration per period. Finally, we integrated the 2 + 2 + 2 best models into a single flood detection, flood volume, and flood duration model based on their evaluation scores (see *Section 4.3* below).

Figure 6 presents the full SatVITS-Flood model scheme. The codes for SatVITS-Flood can be found in our lab's GitHub account: https://github.com/M-M-VS-Lab/-SatVITS-Flood-Model (Burstein et al., 2023).

[Figure 6]

4.3. Statistical analyses and model evaluation

The evaluation of each model result was conducted using two methods, (i) a confusion matrix, which summarizes the classification performance of the flood detection model, and (ii) a linear regression for predicting the volume and duration of the detected floods. We used four statistical metrics based on the confusion matrix results to evaluate each model quantitatively – namely, the accuracy, precision, recall, and F1 scores (Table 2).

The confusion matrix compares the predicted and actual values of the model, providing the true positive (TP; how many events were correctly detected), true negative (TN; how many non-evens were correctly detected), false positive (FP; how many floods events were detected but did not occur), and false negative (how many floods were undetected by the model) cases detected by applying each model. Based on these metrics, the accuracy of the model is computed. However, the use of the accuracy metric alone can be misleading (Kulkarni et al., 2020); this is mainly because imbalanced data may result in a highly accurate model, while this might be an overestimation since most years are no flood years (typical for hyperarid regions). The precision metric is then suggested because it accounts for the model's accuracy while considering also FP cases. The recall metric complements this by indicating the model's strength to identify TP with as few missed cases as possible. The two metrics influence each other, and there is a built-in tradeoff between a model's correctness (precision) and coverage (recall).

To summarize this tradeoff, the F1 score is often used (Sasaki, 2007). The F1 score integrates the precision and recall of the model as a single weighted harmonic mean, with a score of 1 for the perfect model and 0 for the worst. We use the F1 score to select the best-performing model for flood detection for the pre-MODIS and MODIS intervals. Still, we also present and discuss the other statistical metrics (accuracy, precision, and recall).

Table 2. Statistical scores used to evaluate the different models in this study. TP, TN, FP, and FN, are the number of true positive, true negative, false positive, and false negative cases. Q_o and Q_p are for observed and predicted values.

Confusion Matrix				NSE
Accuracy	Precision	Recall	F1	
$\frac{TP + TN}{TP + TN + FP + FN}$	$\frac{TP}{TP + FP}$	$\frac{TP}{TP + FN}$	$2 \times \frac{Precision \times Recall}{Precision + Recall}$	$1 - \frac{\sum_{t=1}^{T} (Q_o^t - Q_p^t)^2}{\sum_{t=1}^{T} (Q_o^t - \overline{Q}_o)^2}$

For the volume and duration prediction models, we selected the best models based on the highest coefficient of determination (R^2), and the minimum root mean square error (RMSE) of their predictions. The SatVITS-Flood model was finally evaluated for flood volume with the Nash–Sutcliffe model efficiency coefficient (NSE), widely used to assess the predictive skill of hydrological models (Table 2). NSE scores approaching 1 and generally greater than 0 are considered acceptable for model performance.

All statistical analyses were performed using Python code. For the confusion matrix statistical scores (accuracy, precision, recall, and F1 scores), the packages sklearn.metrics.accuracy_score, sklearn.metrics.precision_score, sklearn.metrics.recall_score, and score1sklearn.metrics.f1_score were used.

5. Results

5.1. Confusion matrix of the flood detection models

For the pre-MODIS interval, true positive cases (TP) from both BFAST-trend and TFA-GSI_{anom} metrics were higher when using AVHRR VIs than when using Landsat VIs by up to 44% (Fig. 7 and 8). Overall, TFA-GSI_{anom} (Fig. 7) was able to detect almost double the number of floods than BFAST-trend (Fig. 8) but also derived many more false positive cases (FP).

[Figure 7]

[Figure 8]

FP cases from TFA-GSI_{anom} were ~2 to 6 times higher than from BFAST-trend, meaning that although it was more successful in detecting floods, it also caught much more false events.

For the MODIS period, VIs from MODIS performed better at detecting floods, with the only exception of TFA-GSI_{anom} AVHRR NDVI, which was able to detect the maximum number of floods (11 compared to the 10 floods detected by TFA-GSI_{anom} MODIS NDVI; Fig. 9). MODIS VIs also performed better at avoiding FP predictions when using TFA-GSI_{anom} (Fig. 10). When using BFAST-trend, however, MODIS NDVI had twice the number of FP cases than Landsat and

AVHRR, and the same number as MSAVI from Landsat and twice the number of MSAVI from AVHRR (Fig. 9).

[Figure 9]

[Figure 10]

5.2. Overall performance and comparison of models

Figure 11 summarizes the flood detection capability of each time series metric (BFAST-trend and TFA-GSI_{anom}) using the different VIs (NDVI, MSAVI, and NDWI) from the three satellite programs (MODIS, Landsat, and AVHRR) for both the pre-MODIS and MODIS intervals.

Results for the accuracy and precision are mixed with better scores achieved by TFA-GSI_{anom} when Landsat and MODIS VIs are used (except for MODIS NDWI) and lower when AVHRR VIs are used. Overall, the most precise models (highest ratio of TP to the sum of TP and FP) were those of BFAST-trend and TFA-GSI_{anom} using Landsat NDVI (Fig. 11a) and TFA-GSI_{anom} using MODIS NDVI (Fig. 11b), for pre-MODIS and MODIS periods, respectively.

[Figure 11]

TFA-GSI_{anom} had better recall and F1 scores than BFAST-trend (for both time intervals, with all satellites and VIs except MODIS NDVI). The F1-score, the most indicative of the overall precision and recall of the model, was the highest for MSAVI for the two periods (Landsat for the pre-MODIS period and MODIS for the MODIS period).

The relatively moderate scores are due to the small number of flood occurrences during the study period (only 38% and 37% of the years for the pre-MODIS and MODIS periods, respectively), which is typical for hyperarid regions.

5.3. Assessing flood volume and duration

Next, we used the magnitude of BFAST-trend breaks and TFA-GSI_{anom} values to build linear regression models for assessing flood volume and duration. Since the detection model is independent of the flood volume and duration models, we used the volume and duration data of the maximum number of detected floods from the best detection models with all possible combinations of time series metrics, satellites, and VIs.

The statistics for predicting the flood volume and duration are summarized in Figure 12. The individual correlations are presented in Figs S1-S8 in *Supplementary Material.*

[Figure 12]

The highest correlations (R^2) for flood volume and duration during the pre-MODIS interval were of BFAST-trend with AVHRR NDVI (0.43 and 0.28, respectively), which also gave the lowest RMSE (3.25 Mm^3 $event^{-1}$ and 2.7 days for volume and duration, respectively). For the MODIS period, the highest scores were, again, of BFAST-trend, but with MODIS NDWI ($R^2 = 0.75$ and 0.62 for volume and duration, with RMSE of 16.5 Mm^3 $event^{-1}$ and 5.7 days).

5.4. The integrated SatVITS-Flood model

Using the above statistical scores, we integrated the six best-performing models (one for detection, one for flood volume, and one for flood duration for each of the two intervals) into a single SatVITS-Flood model.

Table 3 summarizes the flood detection, volume, and duration models for each period. SatVITS-Flood was actually composed of only four models since the best models for volume were also the best for flood duration.

Table 3. The six best models, indicating the time series metric (*BFAST-trend* or *TFA-GSI$_{anom}$*), the satellite platform (Landsat, AVHRR, or MODIS), and the vegetation index (NDVI, MSAVI, and NDWI) for flood detection, volume, and duration for the pre-MODIS (1981-2000) and MODIS (2001-2021) periods. Models were selected based on their performance using statistical scores (see explanation in the text).

	Model	
Prediction	Pre-MODIS (1981-2000) era	MODIS (2001-2021) era
Detection	*TFA-GSI$_{anom}$* Landsat MSAVI	*TFA-GSI$_{anom}$* MODIS MSAVI
Volume	*BFAST-trend* AVHRR NDVI	*BFAST-trend* MODIS NDWI
Duration	*BFAST-trend* AVHRR NDVI	*BFAST-trend* MODIS NDWI

We applied SatVITS-Flood using all data from both periods, showing that it can detect 70% of the floods that occurred between 1981 and 2021 in the four sites (Fig. 13a). Only 18% of the cases were falsely detected as floods. The overall accuracy of SatVITS-Flood was 0.78, with a precision of 0.67, a recall of 0.69, and an F1-score of 0.68.

SatVITS-Flood was also good at predicting the volume and duration of the detected floods with NSE of 0.79 and RMSE of 15.4 Mm^3 $flood^{-1}$ for flood volume (Fig. 13b) and R^2 of 0.69 and RMSE of 5.7 days for flood duration (Fig. 13c).

[Figure 13]

6. Discussion

The SatVITS-Flood model scheme effectively detects and correlates floods with changes derived from satellite vegetation indices. These results, aligned with previous research, demonstrate the potential of time series analysis of satellite vegetation indices in the flood monitoring and prediction (Fu & Burgher, 2015; Grodek et al., 2020; Manning et al., 2020; Moses et al., 2021). The use of NDVI as a measure of vegetation sensitivity to changes in water supply, and the selection of pixels with a high NDVI annual standard deviation closest to the riverbank, provided an effective way to identify areas likely to be affected by floods, consistent with previous studies that have used NDVI for similar related purposes (Helman et al., 2014; Helman et al., 2014; Helman and Mussery, 2020). Using both BFAST-trend and TFA-GSI_{anom} methods for detecting changes in the time series allowed for a more comprehensive analysis of both perennial and ephemeral vegetation response to floods. Such responses could be leveraged to detect rare floods and estimate the hydrological parameters of the floods, as shown initially in previous studies (Grodek et al., 2020; Moses et al., 2021; Normandin et al., 2022), and more extensively in this study.

The confusion matrix of the flood detection models showed that AVHRR true positive cases were generally higher than those from Landsat, whether when using BFAST-trend or TFA-GSI_{anom} and with NDVI or MSAVI for the period of 1981-2001. The more continuous (daily) temporal resolution of the AVHRR data set could be the reason for that, especially considering that Landsat's revisit time is once every 16 days (Table 1). Although we used linear interpolation to generate a daily data set from Landsat, in cases when there is more than one date of missing data, the result would largely affect the time series and, subsequently, the ability of the metrics to identify significant changes (see, e.g., Helman and Mussery, 2020). This is a known limitation of Landsat satellites, whose data availability is highly affected by weather conditions due to their low revisit frequency (Wulder et al., 2008, 2019).

Both metrics, TFA-GSI_{anom} and BFAST-trend, successfully captured the different aspects of the floods in the hyperarid regions. While TFA-GSI_{anom} outperformed BFAST-trend in detecting flood occurrence, BFAST-trend was better in predicting the hydrological characteristics of the floods, i.e., their volume and duration. Thus, the final SatVITS-Flood model, which is an integrated scheme combining both metrics, was an improved version of each separated model with a much better performance at detecting floods and assessing flood characteristics (Table 3). The likely reason for the better flood detectability of TFA-GSI_{anom} might be its robustness in tracking short-term annual changes. BFAST-trend is good at detecting long-term changes in the baseline trend, which is more representative of the perennial, typically woody vegetation. As such, BFAST-trend could be particularly useful for detecting floods in hyperarid regions since such events are rare. Thus, their effect on the perennial vegetation (and, consequently, on the VI trend) is well pronounced (e.g., Fig. 1d-g). Yet, when floods occur in consecutive years, the ability of the BFAST-trend to detect such events is limited. It can usually detect the first event after several dry years. Still, miss following consecutive events as their impact on the perennial vegetation would be rather gradual. TFA-GSI_{anom}, on the other hand, is more sensitive to short-term changes in the ephemeral plant species and, therefore, can capture such consecutive events since their impact on the ephemeral vegetation would still be anomalous with respect to the long-term seasonal pattern (D. Helman & Mussery, 2020; David Helman et al., 2014). However, since TFA-GSI_{anom} is also sensitive to small changes not necessarily occurring due only to floods (e.g., a relatively anomalous local precipitation amount in specific years), its magnitude of change may be less related to the hydrological characteristics of the flood.

Among the satellite sensors, MODIS, even with its moderate spatial resolution relative to the channel widths (250 meters), proved the most successful in providing reliable VI time series for detecting floods and assessing their volume and duration (Table 3). Having highly reliable VI products based on robust criteria by selecting the best daily observations to produce the 16-day product (Huete et al., 2002), MODIS has provided indispensable information on the Earth's surface for the last 23 years (since 2000 for Terra and 2002 for Aqua satellites). Moreover, BFAST has been widely applied to MODIS time series with great success due to the continuousness and robustness of its products (e.g., Fang et al., 2018; Lambert et al., 2015). Previous studies have compared different products from MODIS, Landsat, and AVHRR, showing, at times, some discrepancies among their products (e.g., Jiang et al., 2017; Tong and He, 2013). For example, Jiang et al. (2017) showed that different leaf area index (LAI) products derived from MODIS and AVHRR had distinctive trends, interannual variabilities, and uncertainty variations. Here we observed differences in magnitude and trend when comparing the same VI from the various satellite platforms in the same place (e.g., Fig. 3a-d). The differences likely resulted from each sensor's unique mechanical and optical characteristics. Also, the same VI was derived from slightly different wavelength ranges and bands due to such sensor characteristics (see details in *Section 3.3*). This might at least affect the VI magnitude, providing different values for the same VI when retrieved from the various sensors (satellites). The distinctive spatial and even temporal resolution (Table 1) is obviously another reason for the observed discrepancy. The coarser spatial resolution of AVHRR (5.7 km) compared to MODIS (250 m and 500 m) and Landsat (30 m) may affect not only the VI magnitude (which would usually result in lower values for AVHRR compared to MODIS and Landsat) but also the trend, as such a large area would probably include various land use covers, each with its own history and dynamics.

Finally, the best combination of the different satellite sensors, VIs, and time series metrics provided a robust flood detection observational-based model. SatVITS-Flood was capable of detecting floods at high accuracy and precision, with an overall F1 of 0.68. It also provided a good assessment of flood volume (NSE = 0.79) and duration (R^2 = 0.69), with acceptable RMSE of 15.4 Mm^3 $event^{-1}$ and 5.7 days. Unlike many remote sensing flood detection tools that intend to provide real-time or near real-time flood detection warnings, our SatVITS-Flood model mainly aims to provide historical long-term (40-year) flood information in mainly hyperarid, ungauged areas. Such hydrological information is indispensable for climate-hydrology research in these fragile regions where water is the most precious resource for humans, animals, and vegetation. The fact that SatVITS-Flood is independent of weather or climate data makes it suitable for studying climatic impacts on the hydrological systems of the global hyperarid regions. Moreover, the outcomes of such research may improve our understanding of how hydrological systems work in those regions. It may serve to calibrate and improve numerical hydrological models, which show severe limitations in predicting floods in extremely dry areas (e.g., Lorenzo Alfieri et al., 2013).

Caveats and future directions

The study proves the reliability of using satellite products in generating long-term flood records. However, further research is needed to validate the model in other regions with different vegetation types and hydrological conditions. Additionally, the effect of factors, such as land use changes and urbanization, on the relationship between vegetation indices and floods needs further study. These factors can significantly affect the relationship between vegetation indices and floods and may lead to different results in other regions.

It is also important to note that the study used a linear model for predicting flood volume and duration, which may only be appropriate for some types of floods. Different characteristics of floods may have other relationships with changes in vegetation indices, and a non-linear model may be more appropriate in some cases.

Finally, the model was evaluated but still needs to be validated with additional, independent data. This was not done here due to the relatively small amount of reliable data on floods in hyperarid regions. Additional data from new sources, such as social media and citizen reports (de Bruijn et al., 2019), may enable a more robust evaluation of the model at a global scale in the future. Using machine learning techniques such as Random Forest and Neural Networks may provide additional information on the hydrological characteristics of the floods and boost the model's performance.

7. Conclusion

In conclusion, the SatVITS-Flood model scheme developed in this study is a promising tool for monitoring and predicting hyperarid floods based on changes in vegetation indices derived from satellites. The use of several VIs and both BFAST-trend and TFA-GSI_{anom} methods for detecting changes in the time series provided a comprehensive analysis of the vegetation response to floods. The results showed a strong correlation between the timing and magnitude of vegetation changes and floods, which can be used to predict or at least largely improve information on flood occurrences, their volumes, and durations.

The main purpose of SatVITS-Flood is to provide historical long-term flood information in remote, mainly ungauged hyperarid regions, where such information is rare and vital for the livelihood of humans and the interaction with fauna and flora. Such hydrological information is independent of weather data and thus can be used for climate-hydrology research to deepen our understanding of how climatic changes may impact hyperarid hydrological systems. However, it is essential to consider the limitations of the proposed model, and thus further research is needed to validate the results in other regions and under different conditions.

The codes for the SatVITS-Flood model are being updated and will be freely available soon via our lab's GitHub account: https://github.com/M-M-VS-Lab/-SatVITS-Flood-Model (Burstein et al., 2023).

Acknowledgments

The authors thank Israel Water Authority for sharing the hydrological data for Nahal Zin and USGS for making their data open and accessible from which we downloaded the hydrological data for Mojave River. This paper is part of an M.Sc. research conducted by O.B. and supervised by D.H. O.B. and D.H. thank the Ring Foundation and the Israel Women League Foundation for partly supporting O.B. through his studies.

References

Abderrazak, B., Morin, D., Bonn, F., & Huete, A. (1996). A review of vegetation indices. *Remote Sensing Reviews*, *13*, 95–120. https://doi.org/10.1080/02757259509532298

Alfieri, L, Burek, P., Dutra, E., Krzeminski, B., Muraro, D., Thielen, J., & Pappenberger, F. (2013). GloFAS – global ensemble streamflow forecasting and flood early warning. *Hydrol. Earth Syst. Sci.*, *17*(3), 1161–1175. https://doi.org/10.5194/hess-17-1161-2013

Alfieri, Lorenzo, Burek, P., Dutra, E., Krzeminski, B., Muraro, D., Thielen, J., & Pappenberger, F. (2013). GloFAS–Global Ensemble Streamflow Forecasting and Flood Early Warning. *Hydrology and Earth System Sciences*, *17*, 1161–1175. https://doi.org/10.5194/hess-17-1161-2013

Armon, M., Dente, E., Smith, J. A., Enzel, Y., & Morin, E. (2018). Synoptic-Scale Control over Modern Rainfall and Flood Patterns in the Levant Drylands with Implications for Past Climates. *Journal of Hydrometeorology*, *19*(6), 1077–1096. https://doi.org/10.1175/JHM-D-18-0013.1

Armon, M., Morin, E., & Enzel, Y. (2019). Overview of modern atmospheric patterns controlling rainfall and floods into the Dead Sea: Implications for the lake's sedimentology and paleohydrology. *Quaternary Science Reviews*, *216*, 58–73. https://doi.org/https://doi.org/10.1016/j.quascirev.2019.06.005

Azar, M., Mulero, G., Yaara, O.-S., Helman, D., & Klein, T. (2023). Aboveground responses to belowground root damage detected by non-destructive sensing metrics in three tree species. *Forestry*.

Bai, J., & Perron, P. (2003). Computation and Analysis of Multiple Structural-Change. *Journal of Applied Econometrics*, *18*.

Benito, G., Rohde, R., Seely, M., Külls, C., Dahan, O., Enzel, Y., et al. (2010). Management of Alluvial Aquifers in Two Southern African Ephemeral Rivers: Implications for IWRM. *Water Resources Management*, *24*(4), 641–667. https://doi.org/10.1007/s11269-009-9463-9

BIANCHI, M., BOYLE, M., & HOLLINGSWORTH, D. (1999). A comparison of methods for trend estimation. *Applied Economics Letters*, *6*(2), 103–109. https://doi.org/10.1080/135048599353726

Blum, M., Lensky, I. M., & Nestel, D. (2013). Estimation of olive grove canopy temperature from MODIS thermal imagery is more accurate than interpolation from meteorological stations. *Agricultural and Forest Meteorology*, *176*, 90–93. https://doi.org/https://doi.org/10.1016/j.agrformet.2013.03.007

Bonakdari, H., Zaji, A. H., Binns, A. D., & Gharabaghi, B. (2019). Integrated Markov chains and uncertainty analysis techniques to more accurately forecast floods using satellite signals. *Journal of Hydrology*, *572*, 75–95. https://doi.org/https://doi.org/10.1016/j.jhydrol.2019.02.027

Boyle, S. A., Kennedy, C. M., Torres, J., Colman, K., Pérez-Estigarribia, P. E., & de la Sancha, N. U. (2014). High-resolution satellite imagery is an important yet underutilized resource in conservation biology. *PloS One*, *9*(1), e86908. https://doi.org/10.1371/journal.pone.0086908

de Bruijn, J. A., de Moel, H., Jongman, B., de Ruiter, M. C., Wagemaker, J., & Aerts, J. C. J. H. (2019). A global database of historic and real-time flood events based on social media. *Scientific Data*, *6*(1), 311. https://doi.org/10.1038/s41597-019-0326-9

Burstein, O., Enzel, Y., Grodek, T., & Helman, D. (2023). SatVITS-Flood-Model [Software]. Retrieved from https://github.com/M-M-VS-Lab/-SatVITS-Flood-Model

Carlson, T. N., & Ripley, D. A. (1997). On the relation between NDVI, fractional vegetation cover, and leaf area index. *Remote Sensing of Environment*, *62*(3), 241–252. https://doi.org/https://doi.org/10.1016/S0034-4257(97)00104-1

Ceccato, P., Flasse, S., Tarantola, S., Jacquemoud, S., & Grégoire, J.-M. (2001). Detecting vegetation leaf water content using reflectance in the optical domain. *Remote Sensing of Environment*, *77*(1), 22–33. https://doi.org/https://doi.org/10.1016/S0034-4257(01)00191-2

Chehbouni, A., Escadafal, R., Duchemin, B., Boulet, G., Simonneaux, V., Dedieu, G., et al. (2008). An integrated modelling and remote sensing approach for hydrological study in arid and semi-arid regions: the SUDMED Programme. *International Journal of Remote Sensing*, *29*(17–18), 5161–5181. https://doi.org/10.1080/01431160802036417

Cian, F., Marconcini, M., & Ceccato, P. (2018). Normalized Difference Flood Index for rapid flood mapping: Taking advantage of EO big data. *Remote Sensing of Environment*, *209*, 712–730. https://doi.org/https://doi.org/10.1016/j.rse.2018.03.006

Cleveland, W. S. (1979). Robust Locally Weighted Regression and Smoothing Scatterplots. *Journal of the American Statistical Association*, *74*(368), 829–836. https://doi.org/10.1080/01621459.1979.10481038

Dahan, O., Tatarsky, B., Enzel, Y., Kulls, C., Seely, M. K., & Benito, G. (2008). Dynamics of Flood Water Infiltration and Ground Water Recharge in Hyperarid Desert. *Groundwater*, *46*.

Dao, P. D., Mong, N. T., & Chan, H.-P. (2019). Landsat-MODIS image fusion and object-based image analysis for observing flood inundation in a heterogeneous vegetated scene. *GIScience & Remote Sensing*, *56*(8), 1148–1169. https://doi.org/10.1080/15481603.2019.1627062

Dasgupta, A., Grimaldi, S., Ramsankaran, R. A. A. J., Pauwels, V. R. N., & Walker, J. P. (2018). Towards operational SAR-based flood mapping using neuro-fuzzy texture-based approaches. *Remote Sensing of Environment*, *215*, 313–329. https://doi.org/https://doi.org/10.1016/j.rse.2018.06.019

DeVries, B., Huang, C., Armston, J., Huang, W., Jones, J. W., & Lang, M. W. (2020). Rapid and robust monitoring of flood events using Sentinel-1 and Landsat data on the Google Earth Engine. *Remote Sensing of Environment*, *240*, 111664.

https://doi.org/https://doi.org/10.1016/j.rse.2020.111664

Enzel, Y. (1992). Flood Frequency of the Mojave River and the Formation of Late Holocene Playa Lakes, Southern California, USA. *The Holocene*, *2*(1), 11–18. https://doi.org/10.1177/095968369200200102

Enzel, Y., & Wells, S. G. (1997). Extracting Holocene paleohydrology and paleoclimatology information from modern extreme flood events: An example from southern California. *Geomorphology*, *19*(3), 203–226. https://doi.org/https://doi.org/10.1016/S0169-555X(97)00015-9

Enzel, Y., Cayan, D. R., Anderson, R. Y., & Wells, S. G. (1989). Atmospheric circulation during Holocene lake stands in the Mojave Desert: evidence of regional climate change. *Nature*, *341*(6237), 44–47. https://doi.org/10.1038/341044a0

Enzel, Y., Ely, L. L., House, P. K., Baker, V. R., & Webb, R. H. (1993). Paleoflood evidence for a natural upper bound to flood magnitudes in the Colorado River Basin. *Water Resources Research*, *29*(7), 2287–2297. https://doi.org/https://doi.org/10.1029/93WR00411

Fang, X., Zhu, Q., Ren, L., Chen, H., Wang, K., & Peng, C. (2018). Large-scale detection of vegetation dynamics and their potential drivers using MODIS images and BFAST: A case study in Quebec, Canada. *Remote Sensing of Environment*, *206*, 391–402. https://doi.org/https://doi.org/10.1016/j.rse.2017.11.017

Fu, B., & Burgher, I. (2015). Riparian vegetation NDVI dynamics and its relationship with climate, surface water and groundwater. *Journal of Arid Environments*, *113*, 59–68. https://doi.org/https://doi.org/10.1016/j.jaridenv.2014.09.010

Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., et al. (2015). The climate hazards infrared precipitation with stations - A new environmental record for monitoring extremes. *Scientific Data*, *2*, 1–21. https://doi.org/10.1038/sdata.2015.66

Gamon, J., Field, C., Goulden, M., Griffin, K., Hartley, A., Joel, G., et al. (1995). Relationships Between NDVI, Canopy Structure, and Photosynthesis in Three Californian Vegetation Types. *Ecological Applications*, *5*, 28–41. https://doi.org/10.2307/1942049

Gao, B. (1996). NDWI—A normalized difference water index for remote sensing of vegetation liquid water from space. *Remote Sensing of Environment*, *58*(3), 257–266. https://doi.org/https://doi.org/10.1016/S0034-4257(96)00067-3

Gao, Z., Long, D., Tang, G., Zeng, C., Huang, J., & Hong, Y. (2017). Assessing the potential of satellite-based precipitation estimates for flood frequency analysis in ungauged or poorly gauged tributaries of China's Yangtze River basin. *Journal of Hydrology*, *550*, 478–496. https://doi.org/https://doi.org/10.1016/j.jhydrol.2017.05.025

Gizaw, M. S., & Gan, T. Y. (2016). Regional Flood Frequency Analysis using Support Vector Regression under historical and future climate. *Journal of Hydrology*, *538*, 387–398. https://doi.org/https://doi.org/10.1016/j.jhydrol.2016.04.041

Glenn, E., Huete, A., Nagler, P., & Nelson, S. (2008). Relationship Between Remotely-sensed Vegetation Indices, Canopy Attributes and Plant Physiological Processes: What Vegetation Indices Can and Cannot Tell Us About the Landscape. *Sensors*, *8*. https://doi.org/10.3390/s8042136

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017). Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment*, *202*, 18–27. https://doi.org/https://doi.org/10.1016/j.rse.2017.06.031

Greenbaum, N., Margalit, A., Schick, A., Sharon, D., & Baker, V. (1998). A high magnitude storm and flood in a hyperarid catchment, Nahal Zin, Negev Desert, Israel. *Hydrological Processes*, *12*, 1–23. https://doi.org/10.1002/(SICI)1099-1085(199801)12:1<1::AID-HYP559>3.0.CO;2-6

Greenbaum, N., Schick, A., & Baker, V. (2000). The palaeoflood record of a hyperarid catchment, Nahal Zin, Negev Desert, Israel. *Earth Surface Processes and Landforms*, *25*, 951–971. https://doi.org/10.1002/1096-9837(200008)25:9<951::AID-ESP110>3.0.CO;2-8

Greenbaum, N., Schwartz, U., Schick, A., & Enzel, Y. (2001). Paleofloods and the Estimation of Long Term Transmission Losses and Recharge to the Lower Nahal Zin Alluvial Aquifer, Negev Desert, Israel. *American Geophysical Union, Water Resources and Application*, *5*, 311–328. https://doi.org/10.1029/WS005p0311

Grimaldi, S., Xu, J., Li, Y., Pauwels, V. R. N., & Walker, J. P. (2020). Flood mapping under vegetation using single SAR acquisitions. *Remote Sensing of Environment*, *237*, 111582. https://doi.org/https://doi.org/10.1016/j.rse.2019.111582

Grodek, T., Morin, E., Helman, D., Lensky, I., Dahan, O., Seely, M., et al. (2020). Eco-hydrology and geomorphology of the largest floods along the hyperarid Kuiseb River, Namibia. *Journal of Hydrology*, *582*, 124450. https://doi.org/https://doi.org/10.1016/j.jhydrol.2019.124450

Helman, D., & Mussery, A. (2020). Using Landsat satellites to assess the impact of check dams built across erosive gullies on vegetation rehabilitation. *Science of the Total Environment*, *730*. https://doi.org/10.1016/j.scitotenv.2020.138873

Helman, D., Mussery, A., Lensky, I. M., & Leu, S. (2014). Detecting changes in biomass productivity in a different land management regimes in drylands using satellite-derived vegetation index. *Soil Use and Management*, *30*(1), 32–39. https://doi.org/10.1111/sum.12099

Helman, D., Lensky, I. M., Osem, Y., Rohatyn, S., Rotenberg, E., & Yakir, D. (2017). A biophysical approach using water deficit factor for daily estimations of evapotranspiration and CO2 uptake in Mediterranean environments. *Biogeosciences*, *14*(17). https://doi.org/10.5194/bg-14-3909-2017

Helman, D., Bonfil, D. J., & Lensky, I. M. (2019). Crop RS-Met: A biophysical evapotranspiration and root-zone soil water content model for crops based on proximal sensing and meteorological data. *Agricultural Water Management*, *211*. https://doi.org/10.1016/j.agwat.2018.09.043

Helman, David. (2018). Land surface phenology: What do we really ‘see’ from space? *Science of The Total Environment*, *618*, 665–673. https://doi.org/https://doi.org/10.1016/j.scitotenv.2017.07.237

Helman, David, Lensky, I. M., Mussery, A., & Leu, S. (2014). Rehabilitating degraded drylands by creating woodland islets: Assessing long-term effects on aboveground productivity and soil fertility. *Agricultural and Forest Meteorology*, *195–196*, 52–60.

https://doi.org/10.1016/j.agrformet.2014.05.003

Helman, David, Lensky, I., Tessler, N., & Osem, Y. (2015). A Phenology-Based Method for Monitoring Woody and Herbaceous Vegetation in Mediterranean Forests from NDVI Time Series. *Remote Sensing*, *7*, 12314–12335. https://doi.org/10.3390/rs70912314

Huete, A., Didan, K., Miura, T., Rodriguez, E. P., Gao, X., & Ferreira, L. G. (2002). Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sensing of Environment*, *83*(1–2), 195–213. https://doi.org/10.1016/S0034-4257(02)00096-2

Iqbal, U., Riaz, M. Z., Zhao, J., Barthelemy, J., & Perez, P. (2023). Drones for Flood Monitoring, Mapping and Detection: A Bibliometric Review. *Drones*. https://doi.org/10.3390/drones7010032

Israel Water Authority. (2020). Tide report - historical data. Zin Elyon gauge station [Dataset]. Retrieved from https://www.gov.il/he/departments/topics/surface-water/govil-landing-page

Izbicki, J. A., Stamos, C. L., Nishikawa, T., & Martin, P. (2004). Comparison of ground-water flow model particle-tracking results and isotopic data in the Mojave River ground-water basin, southern California, USA. *Journal of Hydrology*, *292*(1), 30–47. https://doi.org/https://doi.org/10.1016/j.jhydrol.2003.12.034

Jiang, C., Ryu, Y., Fang, H., Myneni, R., Claverie, M., & Zhu, Z. (2017). Inconsistencies of interannual variability and trends in long-term satellite leaf area index products. *Global Change Biology*, *23*(10), 4133–4146. https://doi.org/https://doi.org/10.1111/gcb.13787

Kahana, R., Ziv, B., Enzel, Y., & Dayan, U. (2002). Synoptic climatology of major floods in the Negev Desert, Israel. *International Journal of Climatology*, *22*(7), 867–882. https://doi.org/https://doi.org/10.1002/joc.766

Kulkarni, A., Chong, D., & Batarseh, F. A. (2020). 5 - Foundations of data imbalance and solutions for a data democracy. In F. A. Batarseh & R. B. T.-D. D. Yang (Eds.) (pp. 83–106). Academic Press. https://doi.org/https://doi.org/10.1016/B978-0-12-818366-3.00005-8

Lambert, J., Denux, J.-P., Verbesselt, J., Balent, G., & Cheret, V. (2015). Detecting Clear-Cuts and Decreases in Forest Vitality Using MODIS NDVI Time Series. *Remote Sensing*. https://doi.org/10.3390/rs70403588

Lehner, B., & Grill, G. (2013). Global River Hydrography and Network Routing: Baseline Data and New Approaches to Study the World's Large River Systems. *Hydrological Processes*, *27*, 2171–2186. https://doi.org/10.1002/hyp.9740

Lensky, I., & Dayan, U. (2011). Detection of Finescale Climatic Features from Satellites and Implications for Agricultural Planning. *Bull. Amer. Meteor. Soc.*, *92*. https://doi.org/10.1175/2011BAMS3160.1

Li, Z., Yang, D., Gao, B., Jiao, Y., Hong, Y., & Xu, T. (2015). Multiscale hydrologic applications of the latest satellite precipitation products in the Yangtze river basin using a distributed hydrologic model. *Journal of Hydrometeorology*, *16*(1), 407–426. https://doi.org/10.1175/JHM-D-14-0105.1

Manfreda, S., McCabe, M. F., Miller, P. E., Lucas, R., Pajuelo Madrigal, V., Mallinis, G., et al. (2018). On the Use of Unmanned Aerial Systems for Environmental Monitoring. *Remote Sensing* . https://doi.org/10.3390/rs10040641

Manning, A., Julian, J. P., & Doyle, M. W. (2020). Riparian vegetation as an indicator of stream channel presence and connectivity in arid environments. *Journal of Arid Environments*, *178*(March), 104167. https://doi.org/10.1016/j.jaridenv.2020.104167

Mardian, J., Berg, A., & Daneshfar, B. (2021). Evaluating the temporal accuracy of grassland to cropland change detection using multitemporal image analysis. *Remote Sensing of Environment*, *255*, 112292. https://doi.org/https://doi.org/10.1016/j.rse.2021.112292

Martinez, J.-M., & Le Toan, T. (2007). Mapping of flood dynamics and spatial distribution of vegetation in the Amazon floodplain using multitemporal SAR data. *Remote Sensing of Environment*, *108*(3), 209–223. https://doi.org/https://doi.org/10.1016/j.rse.2006.11.012

McCormack, T., Campanyà, J., & Naughton, O. (2022). A methodology for mapping annual flood extent using multi-temporal Sentinel-1 imagery. *Remote Sensing of Environment*, *282*, 113273. https://doi.org/https://doi.org/10.1016/j.rse.2022.113273

Meresa, H. (2019). Modelling of river flow in ungauged catchment using remote sensing data: application of the empirical (SCS-CN), Artificial Neural Network (ANN) and Hydrological Model (HEC-HMS). *Modeling Earth Systems and Environment*, *5*. https://doi.org/10.1007/s40808-018-0532-z

Morin, E., Grodek, T., Dahan, O., Benito, G., Kulls, C., Jacoby, Y., et al. (2009). Flood routing and alluvial aquifer recharge along the ephemeral arid Kuiseb River, Namibia. *Journal of Hydrology*, *368*(1), 262–275. https://doi.org/https://doi.org/10.1016/j.jhydrol.2009.02.015

Moses, O., Blamey, R. C., & Reason, C. J. C. (2021). Relationships between NDVI, river discharge and climate in the Okavango River Basin region. *International Journal of Climatology*, (June), 1–23. https://doi.org/10.1002/joc.7267

Normandin, C., Paillou, P., Lopez, S., Marais, E., & Scipal, K. (2022). Monitoring the Dynamics of Ephemeral Rivers from Space: An Example of the Kuiseb River in Namibia. *Water*. https://doi.org/10.3390/w14193142

Puricelli, M., Francés, F., Velez, J., & Lopez Unzu, F. (2009). Parameter extrapolation to ungauged basins with a hydrological distributed model in a regional framework. *Hydrology and Earth System Sciences*, *13*, 229–246. https://doi.org/10.5194/hessd-4-909-2007

Pyayt, A. L., Mokhov, I. I., Lang, B., Krzhizhanovskaya, V. V., & Meijer, R. J. (2011). Machine learning methods for environmental monitoring and flood protection. *World Academy of Science, Engineering and Technology*, *78*, 118–123.

Qi, J., Chehbouni, A., Huete, A. R., Kerr, Y. H., & Sorooshian, S. (1994). A modified soil adjusted vegetation index. *Remote Sensing of Environment*, *48*(2), 119–126. https://doi.org/https://doi.org/10.1016/0034-4257(94)90134-1

Rahman, M., Chen, N., Elbeltagi, A., Islam, M. M., Alam, M., Pourghasemi, H. R., et al. (2021). Application of stacking hybrid machine learning algorithms in delineating multi-type flooding in Bangladesh. *Journal of Environmental Management*, *295*, 113086. https://doi.org/https://doi.org/10.1016/j.jenvman.2021.113086

Rondeaux, G., Steven, M., & Baret, F. (1996). Optimization of soil-adjusted vegetation indices. *Remote Sensing of Environment*, *55*(2), 95–107. https://doi.org/https://doi.org/10.1016/0034-4257(95)00186-7

Rouse, J. W., Hass, R. H., Schell, J. A., Deering, D. W., & Harlan, J. C. (1974). *Monitoring the vernal advancement and retrogradation (greenwave effect) of natural vegetation*. (F. A. O. of the UN, Ed.). Greenbelt, Maryland: Texas A & M University, Remote Sensing Center.

Sasaki, Y. (2007). The truth of the F-measure. *Teach Tutor Mater*, *1*(5), 1–5.

Scharlemann, J. P. W., Benz, D., Hay, S. I., Purse, B. V, Tatem, A. J., Wint, G. R. W., & Rogers, D. J. (2008). Global Data for Ecology and Epidemiology: A Novel Algorithm for Temporal Fourier Processing MODIS Data. *PLOS ONE*, *3*(1), e1408. Retrieved from https://doi.org/10.1371/journal.pone.0001408

Scott, M. L., Lines, G. C., & Auble, G. T. (2000). Channel incision and patterns of cottonwood stress and mortality along the Mojave River, California. *Journal of Arid Environments*, *44*(4), 399–414. https://doi.org/https://doi.org/10.1006/jare.1999.0614

Seabold, S., & Perktold, J. (2010). Statsmodels: Econometric and Statistical Modeling with Python [Software]. *Proceedings of the 9th Python in Science Conference*, *2010*, 92–96. https://doi.org/10.25080/majora-92bf1922-011

Shahabi, H., Shirzadi, A., Ghaderi, K., Omidvar, E., Al-Ansari, N., Clague, J. J., et al. (2020). Flood Detection and Susceptibility Mapping Using Sentinel-1 Remote Sensing Data and a Machine Learning Approach: Hybrid Intelligence of Bagging Ensemble Based on K-Nearest Neighbor Classifier. *Remote Sensing*. https://doi.org/10.3390/rs12020266

Shen, X., Anagnostou, E. N., Allen, G. H., Robert Brakenridge, G., & Kettner, A. J. (2019). Near-real-time non-obstructed flood inundation mapping using synthetic aperture radar. *Remote Sensing of Environment*, *221*, 302–315. https://doi.org/https://doi.org/10.1016/j.rse.2018.11.008

Tong, A., & He, Y. (2013). Comparative analysis of SPOT, Landsat, MODIS, and AVHRR normalized difference vegetation index data on the estimation of leaf area index in a mixed grassland ecosystem. *Journal of Applied Remote Sensing*, *7*(1), 73599. https://doi.org/10.1117/1.JRS.7.073599

Tripathi, G., Pandey, A. C., Parida, B. R., & Kumar, A. (2020). Flood Inundation Mapping and Impact Assessment Using Multi-Temporal Optical and SAR Satellite Data: a Case Study of 2017 Flood in Darbhanga District, Bihar, India. *Water Resources Management*, *34*(6), 1871–1892. https://doi.org/10.1007/s11269-020-02534-3

Tucker, C. J. (1980). Remote sensing of leaf water content in the near infrared. *Remote Sensing of Environment*, *10*(1), 23–32. https://doi.org/https://doi.org/10.1016/0034-4257(80)90096-6

USGS. (2023). National Water Information System. USGS 10262500 MOJAVE R A BARSTOW CA [Dataset]. Retrieved from https://waterdata.usgs.gov/nwis/inventory?agency_code=USGS&site_no=10262500

Vekaria, D., Chander, S., Singh, R. P., & Dixit, S. (2022). A change detection approach to flood inundation mapping using multi-temporal Sentinel-1 SAR images, the Brahmaputra River, Assam (India): 2015–2020. *Journal of Earth System Science*, *132*(1), 3. https://doi.org/10.1007/s12040-022-02020-x

Verbesselt, J., Hyndman, R., Newnham, G., & Culvenor, D. (2010). Detecting trend and seasonal changes in satellite image time series [Software]. *Remote Sensing of Environment*, *114*(1), 106–115. https://doi.org/https://doi.org/10.1016/j.rse.2009.08.014

Verbesselt, J., Hyndman, R., Zeileis, A., & Culvenor, D. (2010). Phenological change detection while accounting for abrupt and gradual trends in satellite image time series [Software]. *Remote Sensing of Environment*, *114*(12), 2970–2980. https://doi.org/https://doi.org/10.1016/j.rse.2010.08.003

Verbesselt, J., Zeileis, A., & Herold, M. (2012). Near real-time disturbance detection using satellite image time series. *Remote Sensing of Environment*, *123*, 98–108. https://doi.org/https://doi.org/10.1016/j.rse.2012.02.022

Verdugo-Vásquez, N., Gutiérrez-Gamboa, G., Díaz-Gálvez, I., Ibacache, A., & Zurita-Silva, A. (2021). Modifications Induced by Rootstocks on Yield, Vigor and Nutritional Status on Vitis vinifera Cv Syrah under Hyper-Arid Conditions in Northern Chile. *Agronomy* . https://doi.org/10.3390/agronomy11050979

Wang, J., Hong, Y., Li, L., Gourley, J. J., Khan, S. I., Yilmaz, K. K., et al. (2011). The coupled routing and excess storage (CREST) distributed hydrological model. *Hydrological Sciences Journal*, *56*(1), 84–98. https://doi.org/10.1080/02626667.2010.543087

Watts, L. M., & Laffan, S. W. (2014). Effectiveness of the BFAST algorithm for detecting vegetation response patterns in a semi-arid region. *Remote Sensing of Environment*, *154*, 234–245. https://doi.org/https://doi.org/10.1016/j.rse.2014.08.023

Wulder, M. A., White, J. C., Goward, S. N., Masek, J. G., Irons, J. R., Herold, M., et al. (2008). Landsat continuity: Issues and opportunities for land cover monitoring. *Remote Sensing of Environment*, *112*(3), 955–969. https://doi.org/https://doi.org/10.1016/j.rse.2007.07.004

Wulder, M. A., Loveland, T. R., Roy, D. P., Crawford, C. J., Masek, J. G., Woodcock, C. E., et al. (2019). Current status of Landsat program, science, and applications. *Remote Sensing of Environment*, *225*, 127–147. https://doi.org/https://doi.org/10.1016/j.rse.2019.02.015

Xie, S., Wu, W., Mooser, S., Wang, Q. J., Nathan, R., & Huang, Y. (2021). Artificial neural network based hybrid modeling approach for flood inundation modeling. *Journal of Hydrology*, *592*, 125605. https://doi.org/https://doi.org/10.1016/j.jhydrol.2020.125605

Xinyi, S., Yang, H., Ke, Z., & Zengchao, H. (2017). Refining a Distributed Linear Reservoir Routing Method to Improve Performance of the CREST Model. *Journal of Hydrologic Engineering*, *22*(3), 4016061. https://doi.org/10.1061/(ASCE)HE.1943-5584.0001442

Xu, L., Herold, M., Tsendbazar, N.-E., Masiliūnas, D., Li, L., Lesiv, M., et al. (2022). Time series analysis for global land cover change monitoring: A comparison across sensors. *Remote Sensing of Environment*, *271*, 112905. https://doi.org/https://doi.org/10.1016/j.rse.2022.112905

Yang, S., Li, C., Lou, H., Wang, P., Wang, J., & Ren, X. (2020). Performance of an Unmanned Aerial Vehicle (UAV) in Calculating the Flood Peak Discharge of Ephemeral Rivers Combined with the Incipient Motion of Moving Stones in Arid Ungauged Regions. *Remote Sensing* . https://doi.org/10.3390/rs12101610

Zaman, M. A., Rahman, A., & Haddad, K. (2012). Regional flood frequency analysis in arid regions: A case study for Australia. *Journal of Hydrology*, *475*, 74–83. https://doi.org/https://doi.org/10.1016/j.jhydrol.2012.08.054

Zhao, G., Li, E., Mu, X., Wen, Z., Rayburg, S., & Tian, P. (2015). Changing trends and regime shift of streamflow in the Yellow River basin. *Stochastic Environmental Research and Risk Assessment*, *29*(5), 1331–1343. https://doi.org/10.1007/s00477-015-1058-9

Tables

Table 1. Main characteristics of the satellite programs and sensors and the vegetation indices derived from each sensor for this study.

Short name	Satellites	Sensor/s	Period	Temporal resolution	Spatial resolution	Index	Product
MODIS	Terra, Aqua	The MODerate resolution Imaging Spectroradiometer (MODIS)	2000 – 2021	16-day for NDVI (available twice a day) Daily for NDWI / MSAVI	250 m for NDVI / MSAVI 500 m for NDWI	NDVI, MSAVI, NDWI,	MOD13Q1 for NDVI MOD09GQ for MSAVI MCD43A4 for NDWI
Landsat	Landsat 5, Landsat 8	Multispectral Scanner (MSS) and Thematic Mapper (TM) for Landsat 5, Operational Land Imager (OLI) for Landsat 8	1984 – 2012 for Landsat 5 2013 – 2021 for Landsat 8	16-day	30 m	NDVI, MSAVI, NDWI	LT05/C01/T1_TOA for Landsat 5, LC08/C01/T1_TOA for Landsat 8
AVHRR	POES and MetOp satellites	The Advanced Very-High-Resolution Radiometer (AVHRR)	1981 – 2021	Daily (available 4 times a day)	5566 m	NDVI, MSAVI	NOAA/CDR/AVHRR/NDVI/V5 for NDVI NOAA/CDR/AVHRR/SR/V5 for MSAVI

Table 2. Statistical scores used to evaluate the different models in this study. TP, TN, FP, and FN, are the number of true positive, true negative, false positive, and false negative cases. Q_o and Q_p are for observed and predicted values.

Confusion Matrix				**NSE**
Accuracy	Precision	Recall	F1	
$\frac{TP + TN}{TP + TN + FP + FN}$	$\frac{TP}{TP + FP}$	$\frac{TP}{TP + FN}$	$2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$	$1 - \frac{\sum_{t=1}^{T} (Q_o^t - Q_p^t)^2}{\sum_{t=1}^{T} (Q_o^t - \bar{Q}_o)^2}$

Table 3. The six best models, indicating the time series metric (*BFAST-trend* or *TFA-GSI*$_{anom}$), the satellite platform (Landsat, AVHRR, or MODIS), and the vegetation index (NDVI, MSAVI, and NDWI) for flood detection, volume, and duration for the pre-MODIS (1981-2000) and MODIS (2001-2021) periods. Models were selected based on their performance using statistical scores (see explanation in the text).

	Model	
Prediction	Pre-MODIS (1981-2000) era	MODIS (2001-2021) era
Detection	*TFA-GSI*$_{anom}$ Landsat MSAVI	*TFA-GSI*$_{anom}$ MODIS MSAVI
Volume	*BFAST-trend* AVHRR NDVI	*BFAST-trend* MODIS NDWI
Duration	*BFAST-trend* AVHRR NDVI	*BFAST-trend* MODIS NDWI

Figure captions

Figure 1. Illustration of the indirect effect of floods on the riverbank vegetation in a typical hyperarid region (Tsauchab River, Sesriem, Namibia; Lat/Lon: -24.654/15.650) and its detection by MODIS NDVI time series. View of (**a-b**) the January 2021 flood bore. Such floods (**c**) reach and raise the water table in the downstream alluvial aquifer. This has an immediate effect on the riverbank vegetation, which can be detected in the satellite vegetation index time series. An example is given in **d** for MODIS-derived NDVI (unitless) for the same location, showing an increase in the baseline of the NDVI from 2010 to 2012 following two large floods and a gradual decrease towards 2014 following two consecutive dry years. The pictures in **e-g** correspond to the dates marked by the arrows in **d**. The effect on the NDVI signal is evident in both the baseline (thick line) and original time series with its seasonal signal (dotted thin line), which correspond to changes in the perennial and ephemeral vegetation, respectively. Images **a** and **b** are from a YouTube video uploaded by Chris Kloppers (https://www.youtube.com/watch?v=dVrLIfnT-EE&t=51s). Images **e-g** are from Google Earth. The MODIS NDVI time series was downloaded through GEE.

Figure 2. Location of the four study sites in three hyperarid regions from three continents. (**a**) The general location of the study areas is shown on top of a mean annual precipitation map produced from the CHIRPS product (Funk et al., 2015). The specific locations of the hydrological gauge stations (blue circles) and the selected pixels (green stars) are shown for (**b**) Zin River in Israel, (**c**) Bastrow station in Mojave River (California, USA), and Rooibank and Gobabeb stations at the Kuiseb River (Namibia).

Figure 3. Flood volumes (Mm^3 $year^{-1}$) and NDVI time series for the four sites in Fig. 2. NDVI time series were derived from MODIS (green line), both Landsat 5 and 8 missions (brown lines), and AVHRR (purple line). Notice that there is a gap in the Landsat data due to failures in Landsat 7, which was supposed to connect Landsat 5 and 8 missions.

Figure 4. (**a**) Example of an NDVI Std map (long-term standard deviation NDVI values) around Rooibank, showing high values in a vegetated area near the riverbank. The VI time series were selected based on high NDVI Std pixels for the model development. Images **b** and **c** show a close-up view of the rectangular red area in **a**.

Figure 5. The steps in generating the time series analysis metrics, (**a-d**) BFAST-trend and (**e-g**) TFA-GSI_{anom}. (**a**) An example of MODIS NDVI time series from a selected pixel at Rooibank, and BFAST-derived (**b**) seasonal, (**c**) trend, and (**d**) reminder (irregular) components. BFAST breaks in the trend (BFAST-trend), with their corresponding magnitudes, are shown in **c**. (**e**) The same NDVI time series (back line) as in **a** and its LOWESS-derived trend (red line). (**f**) The remaining seasonal signal after subtracting the trend from the original time series. The smoothed seasonal signal is shown as a blue line, the Temporal Fourier Analysis (TFA) of the smoothed seasonal signal as a green line, and the difference between them as the blue area underneath the

curve in **f**. (**g**) The positive integral over the blue area in **f** per year is the TFA-GSI$_{anom}$. When TFA-GSI$_{anom}$ exceeds one standard deviation (+1σ) it is regarded as a potentially flood year. The magnitude of breaks in BFAST-trend in **c** and the TFA-GSI$_{anom}$ values in **g** were used to derive the flood volume and duration through a simple linear regression with hydrological data.

Figure 6. The SatVITS-Flood model scheme.

Figure 7. Confusion matrix for flood detection for the pre-MODIS interval of 1981-1999 using BFAST-trend with Landsat and AVHRR VIs. Numbers indicate the number of events and the percentage of each event category from the total. Notice that NDWI was not produced from AVHRR because it does not have information on reflectance at the SWIR band.

Figure 8. Confusion matrix for flood detection for the pre-MODIS interval of 1981-1999 using TFA-GSI$_{anom}$ with Landsat and AVHRR VIs. Numbers indicate the number of floods and the percentage of each detection category from the total. Notice that NDWI was not produced from AVHRR because it does not have information on reflectance at the SWIR band.

Figure 9. Confusion matrix for flood detection for the MODIS interval of 2000-2021 for BFAST-trend with MODIS, Landsat, and AVHRR VIs. Numbers indicate the number of floods and the percentage of detection categories from the total. Notice that NDWI was not produced from AVHRR because it does not have information on reflectance at the SWIR band.

Figure 10. Confusion matrix for flood detection for the MODIS interval of 2000-2021 for TFA-GSI$_{anom}$ with MODIS, Landsat, and AVHRR VIs. Numbers indicate the number of floods and the percentage of detection categories from the total. Notice that NDWI was not produced from AVHRR because it does not have information on reflectance at the SWIR band.

Figure 11. Flood detection statistics (accuracy, precision, recall, and F1-score; see Table 2 for the formulation of these statistical metrics) for BFAST-trend and TFA-GSI$_{anom}$ when using different VIs from the three satellites for the (**a**) pre-MODIS (1981-1999) and (**b**) MODIS (2000-2021) intervals. The best scores are highlighted in bold.

Figure 12. Flood volume and duration prediction scores (R^2 and RMSE) for the magnitude of change calculated from BFAST-trend and TFA-GSI$_{anom}$ using VIs from the three satellites for the (**a**) pre-MODIS (1981-1999) and (**b**) MODIS (2000-2021) intervals. The best scores are highlighted in bold.

Figure 13. (**a**) Confusion matrix of the flood detection capability of the integrated SatVITS-Flood model for the entire study interval of 1981–2021 and the four sites (Bastrow, Gobabeb, Rooibank, and Zin). Prediction of (**b**) flood volume and (**c**) duration of the detected floods using the integrated SatVITS-Flood model. The shaded bands in **b** and **c** indicate the range of the 95% confidence of the model, while the dashed line indicates the 1:1 line.

Figure 1.

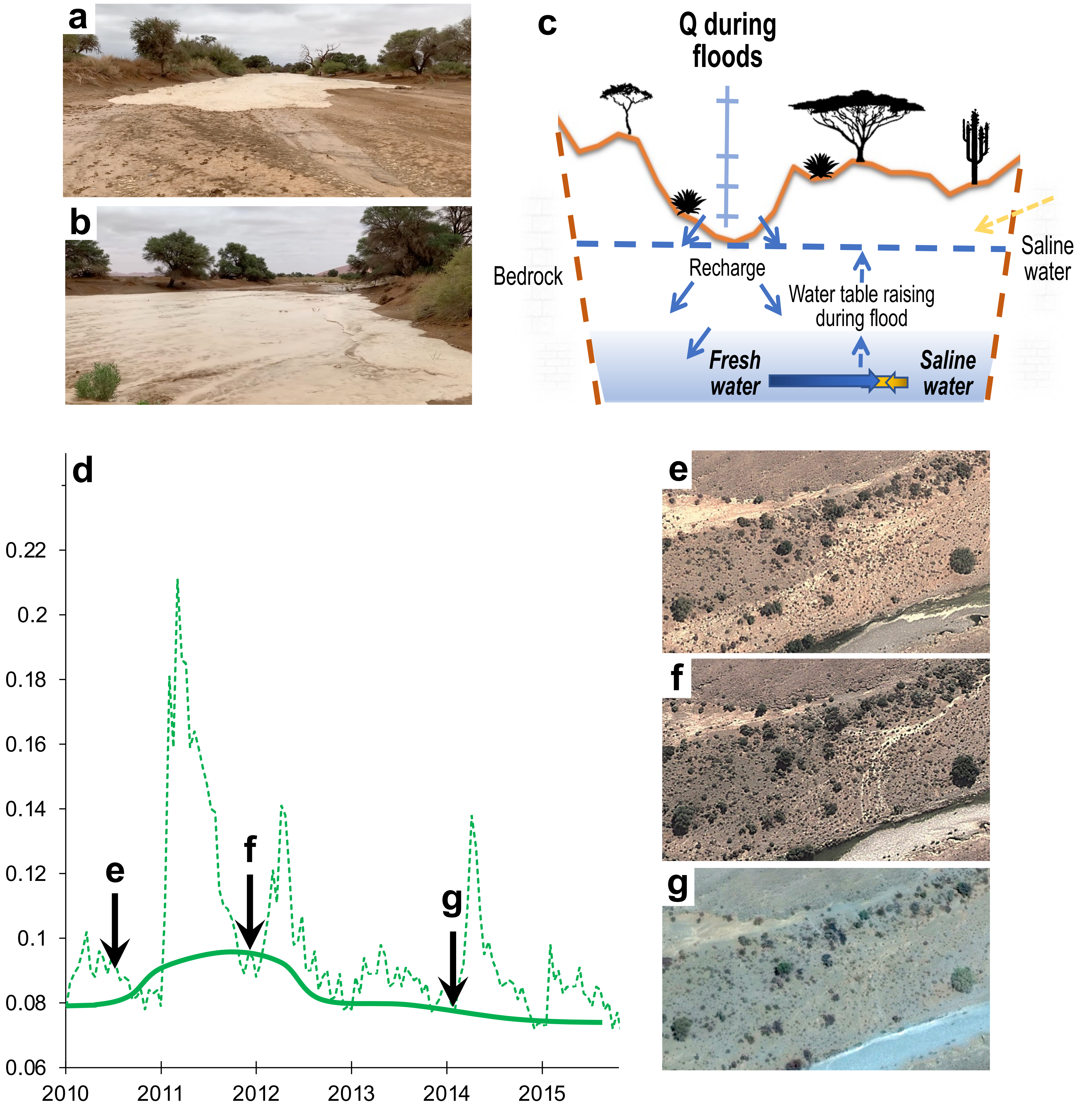

a
b
c
Q during floods
Bedrock
Recharge
Saline water
Water table raising during flood
Fresh water
Saline water
d
0.22
0.2
0.18
0.16
0.14
0.12
0.1
0.08
0.06
2010
2011
2012
2013
2014
2015
e
f
g

Figure 2.

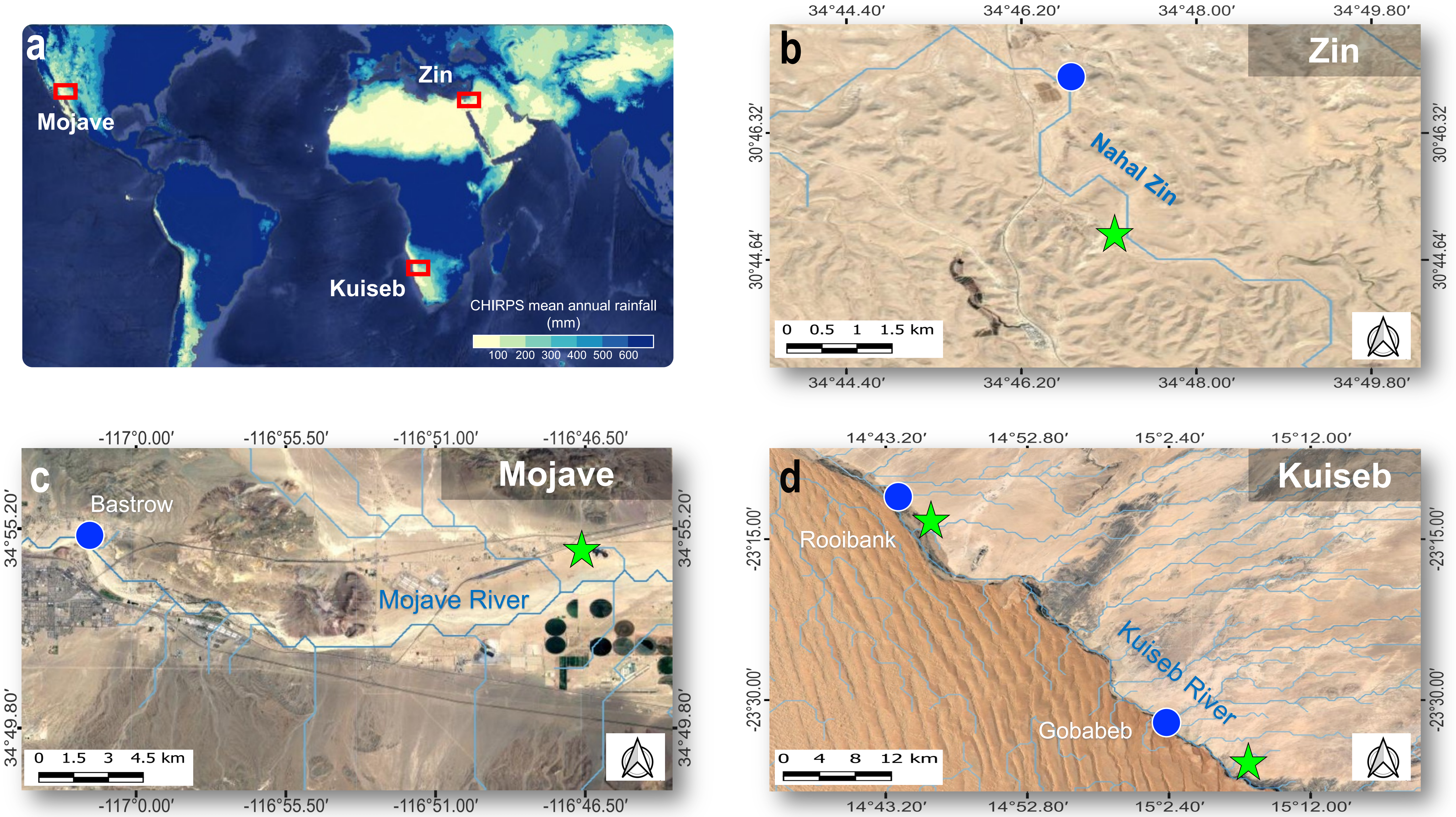
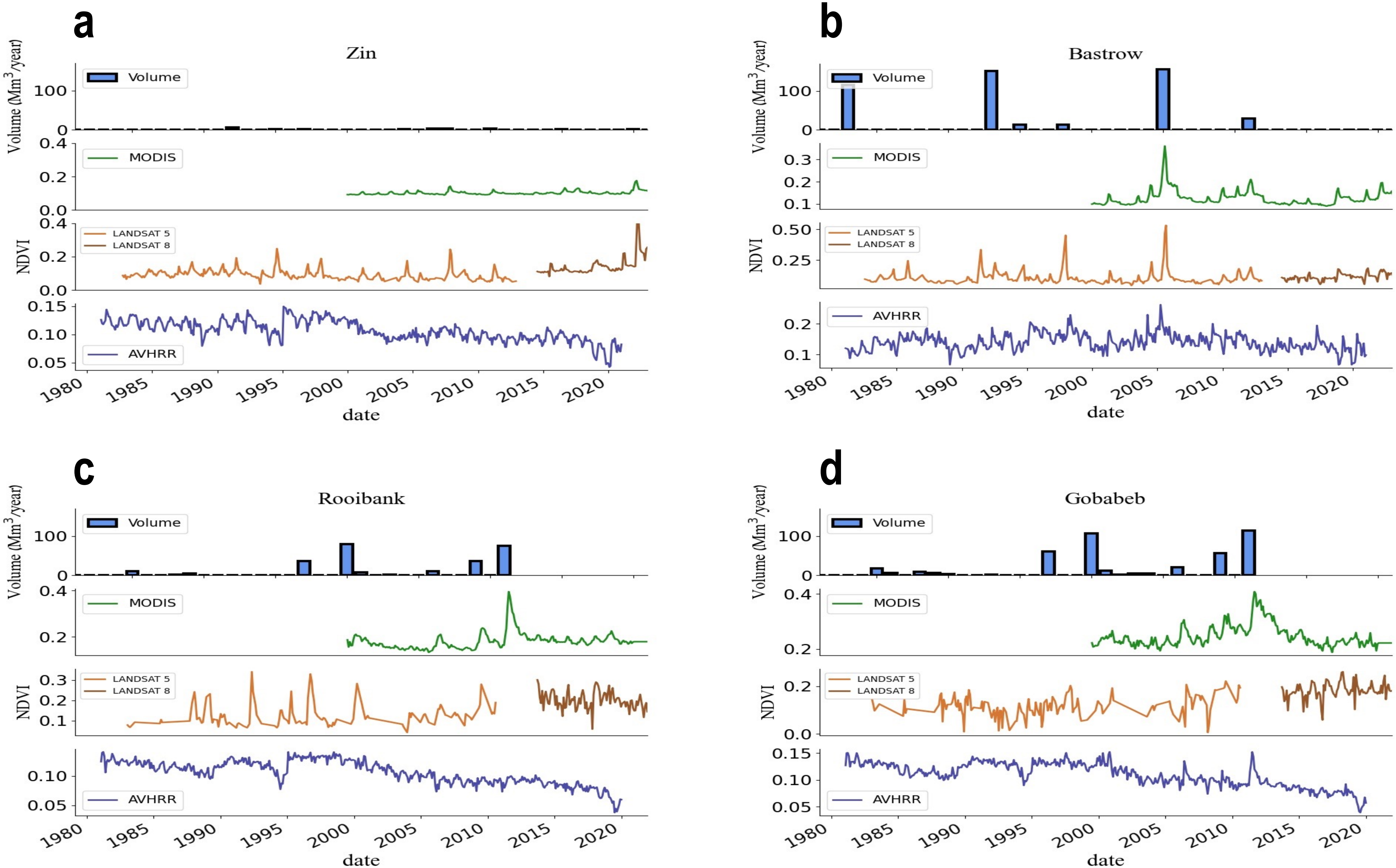

a
Zin
Mojave
Kuiseb
CHIRPS mean annual rainfall (mm)
100 200 300 400 500 600
b
Zin
34°44.40′
34°46.20′
34°48.00′
34°49.80′
30°46.32′
30°44.64′
Nahal Zin
0 0.5 1 1.5 km
c
Mojave
-117°0.00′
-116°55.50′
-116°51.00′
-116°46.50′
34°55.20′
34°49.80′
Bastrow
Mojave River
0 1.5 3 4.5 km
d
Kuiseb
14°43.20′
14°52.80′
15°2.40′
15°12.00′
-23°15.00′
-23°30.00′
Rooibank
Kuiseb River
Gobabeb
0 4 8 12 km

Figure 3.



a
Zin
b
Bastrow
c
Rooibank
d
Gobabeb
Volume (Mm^3/year)
Volume
MODIS
NDVI
LANDSAT 5
LANDSAT 8
AVHRR
date

Figure 4.

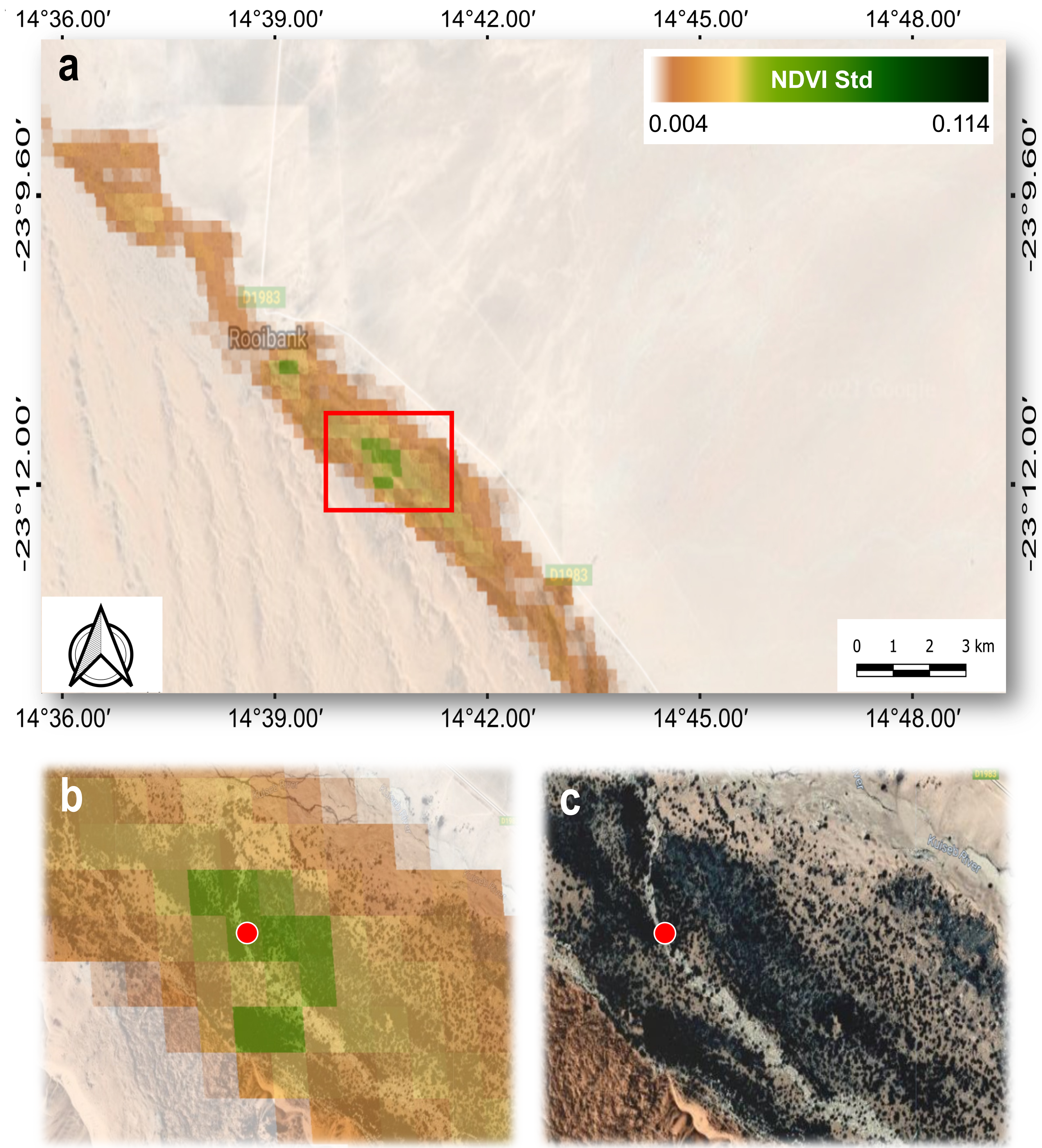

14°36.00′
14°39.00′
14°42.00′
14°45.00′
14°48.00′
a
NDVI Std
0.004
0.114
-23°9.60′
-23°12.00′
D1983
Rooibank
0 1 2 3 km
b
c
Kuiseb River

Figure 5.

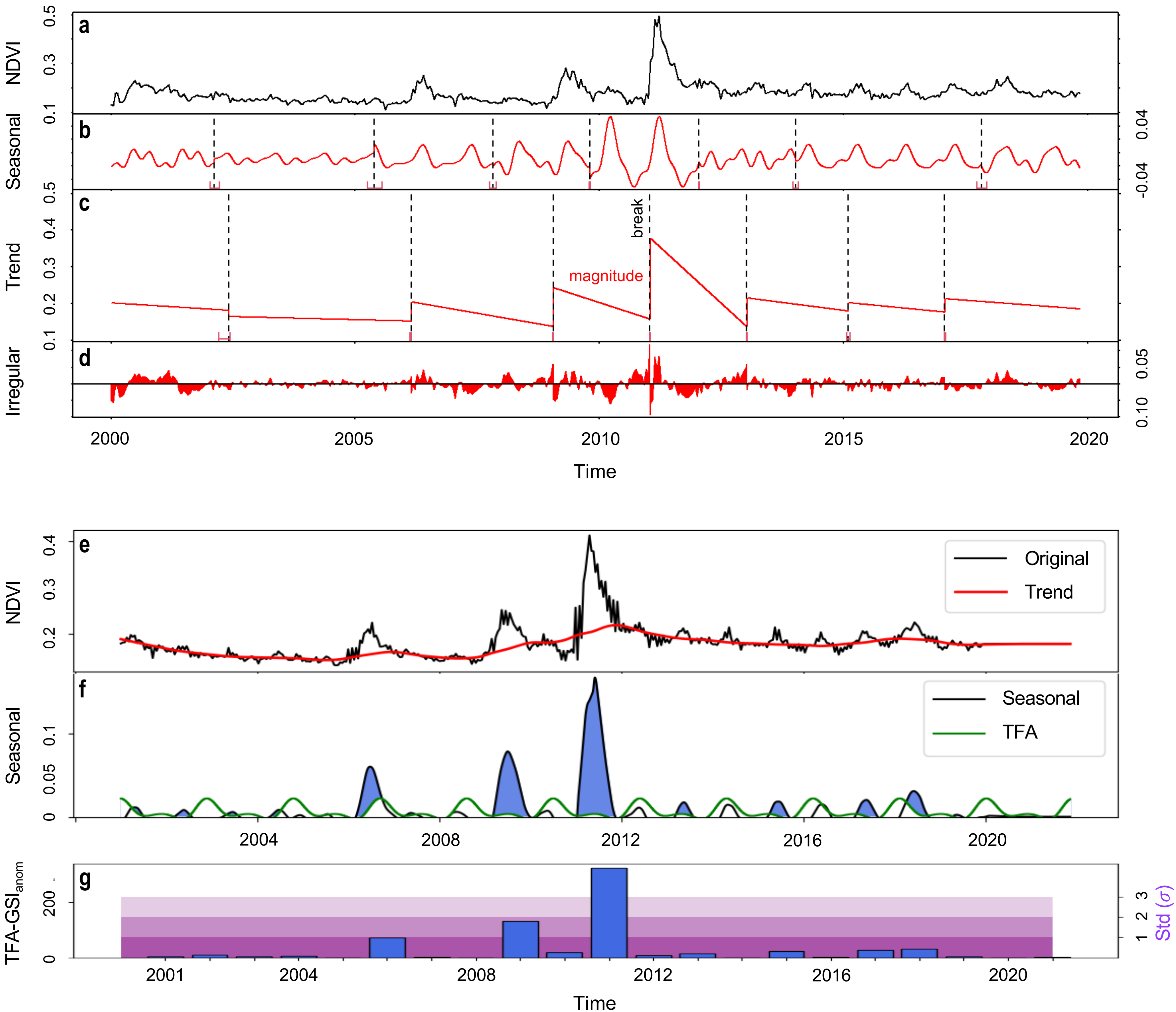

a
b
c
d
e
f
g
NDVI
Seasonal
Trend
Irregular
break
magnitude
Time
Original
Trend
Seasonal
TFA
TFA-GSI$_{anom}$
Std (σ)

Figure 6.

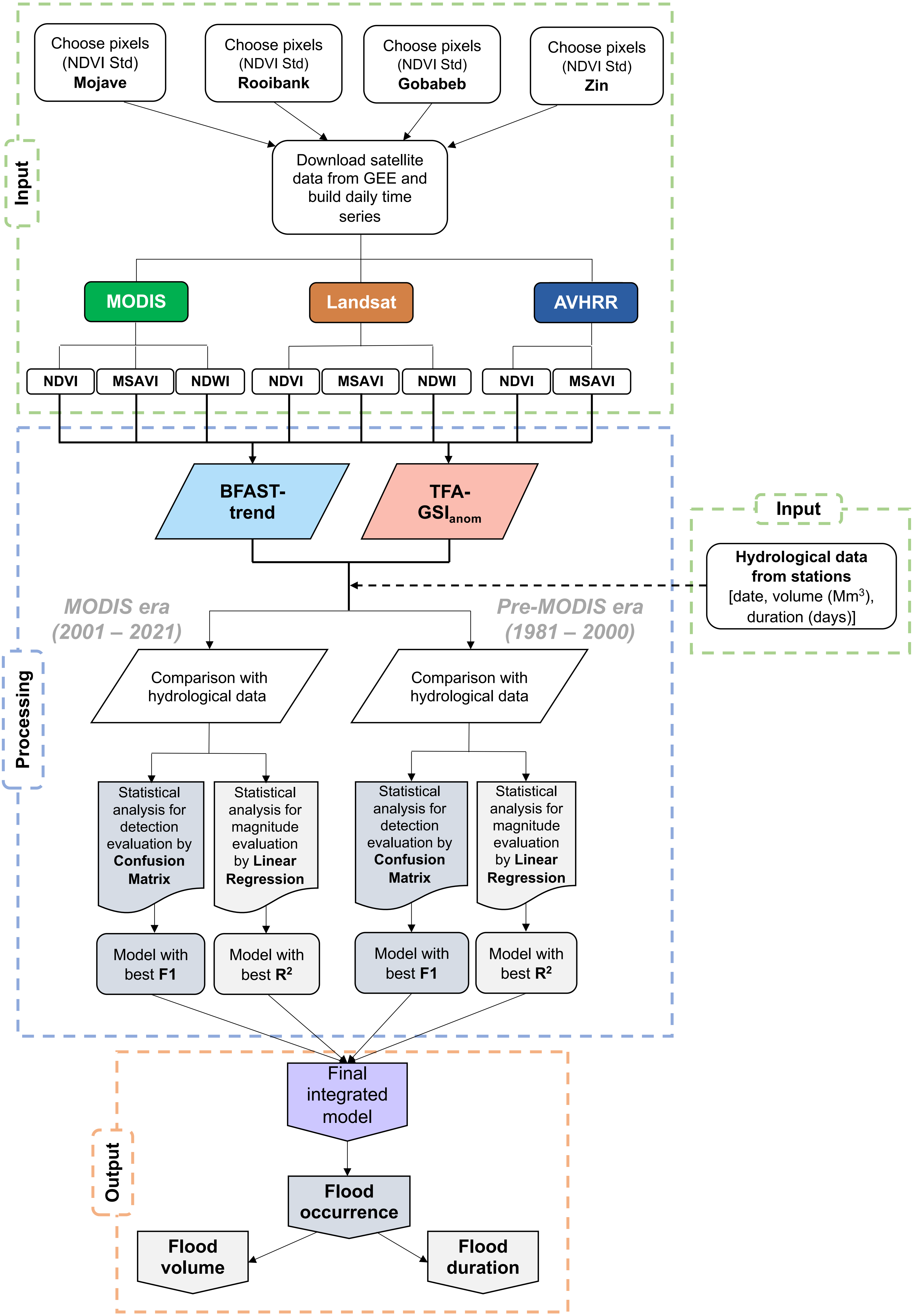

Input
Choose pixels (NDVI Std) Mojave
Choose pixels (NDVI Std) Rooibank
Choose pixels (NDVI Std) Gobabeb
Choose pixels (NDVI Std) Zin
Download satellite data from GEE and build daily time series
MODIS
Landsat
AVHRR
NDVI
MSAVI
NDWI
NDVI
MSAVI
NDWI
NDVI
MSAVI
Processing
BFAST-trend
TFA-GSI$_{anom}$
Input
Hydrological data from stations
[date, volume (Mm^3), duration (days)]
MODIS era (2001 – 2021)
Pre-MODIS era (1981 – 2000)
Comparison with hydrological data
Comparison with hydrological data
Statistical analysis for detection evaluation by Confusion Matrix
Statistical analysis for magnitude evaluation by Linear Regression
Statistical analysis for detection evaluation by Confusion Matrix
Statistical analysis for magnitude evaluation by Linear Regression
Model with best F1
Model with best R^2
Model with best F1
Model with best R^2
Output
Final integrated model
Flood occurrence
Flood volume
Flood duration

Figure 7.

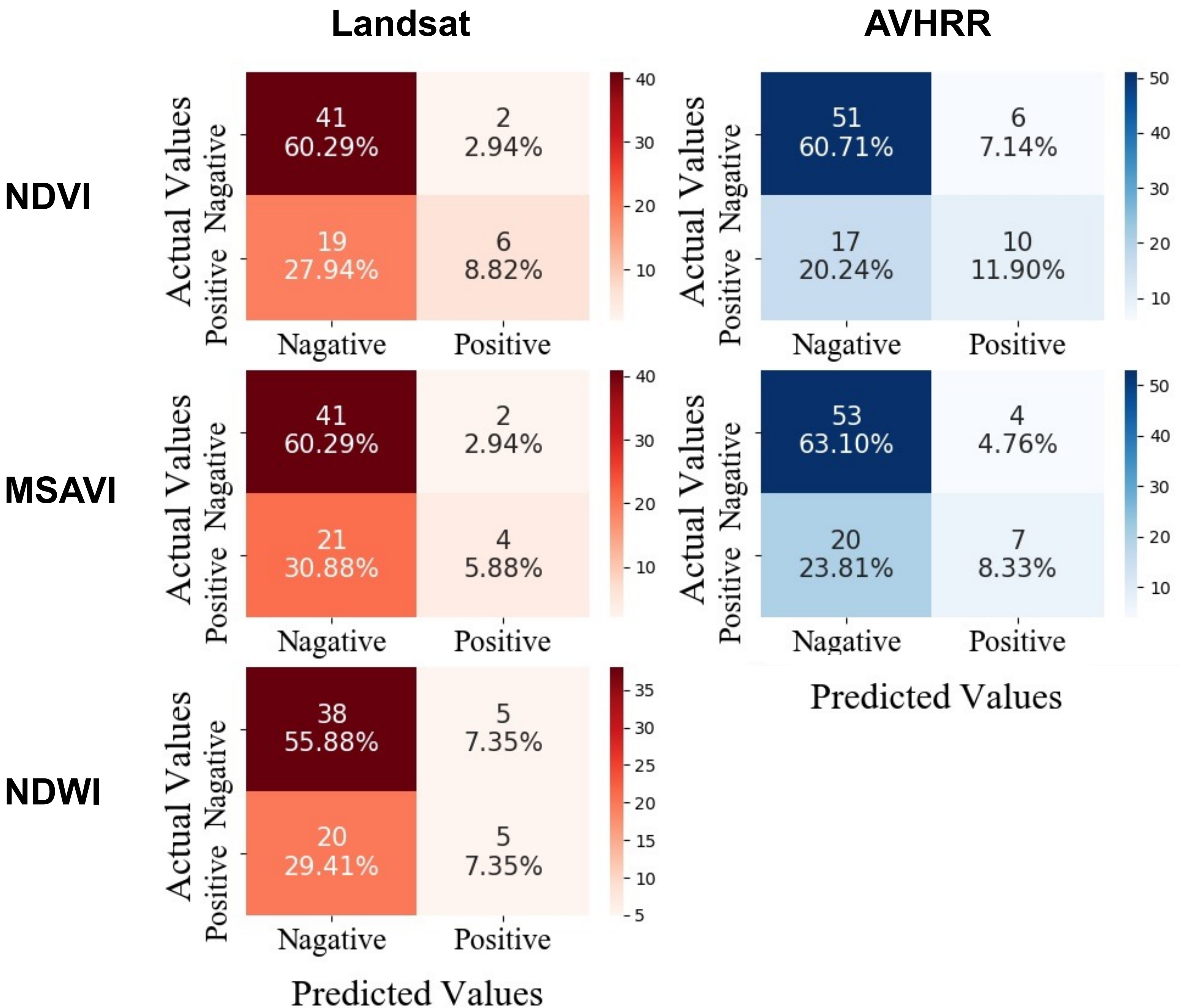

Landsat
AVHRR
NDVI
MSAVI
NDWI
Actual Values
Nagative
Positive
Predicted Values
41
60.29%
2
2.94%
19
27.94%
6
8.82%
51
60.71%
6
7.14%
17
20.24%
10
11.90%
41
60.29%
2
2.94%
21
30.88%
4
5.88%
53
63.10%
4
4.76%
20
23.81%
7
8.33%
38
55.88%
5
7.35%
20
29.41%
5
7.35%

Figure 8.

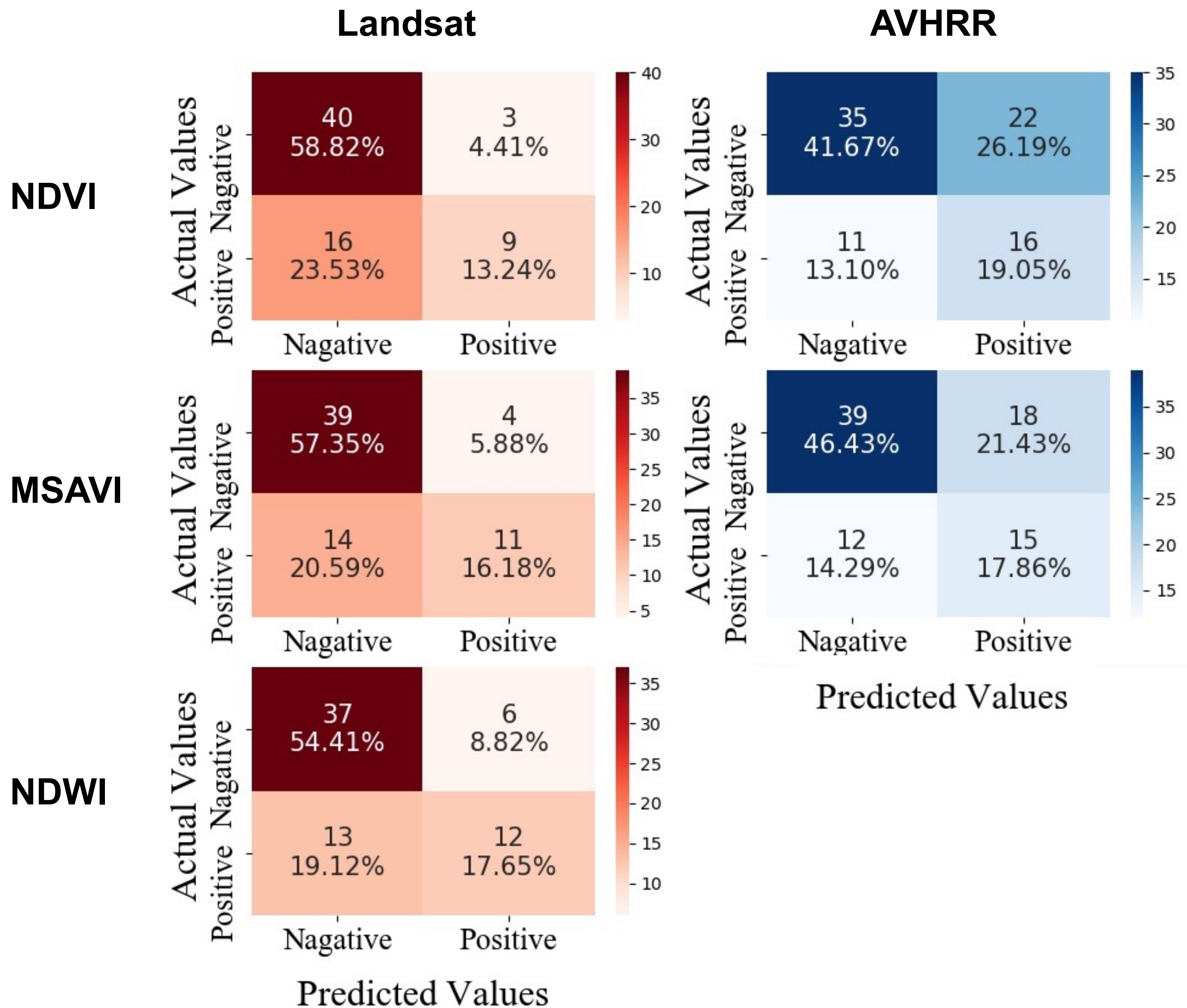
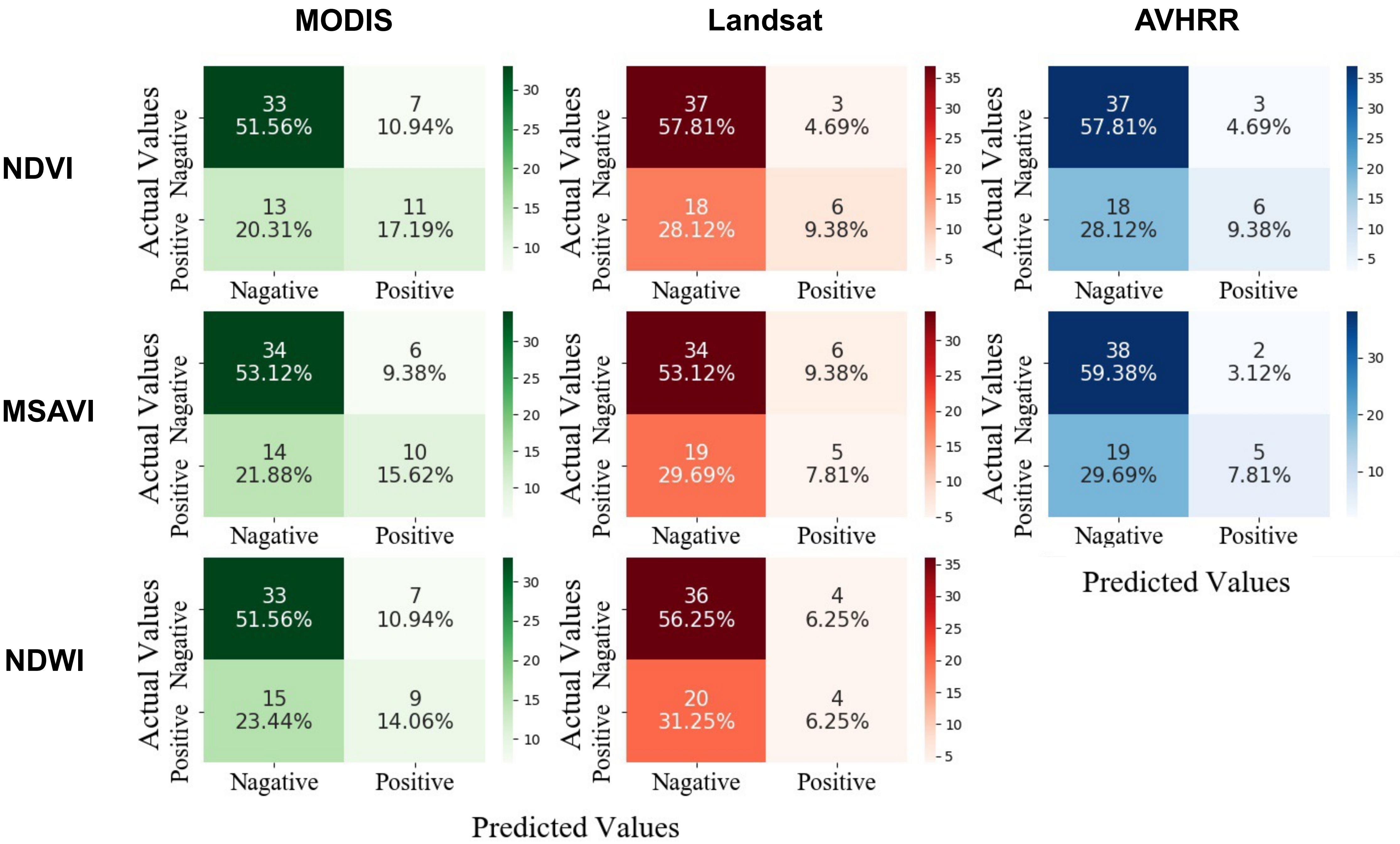
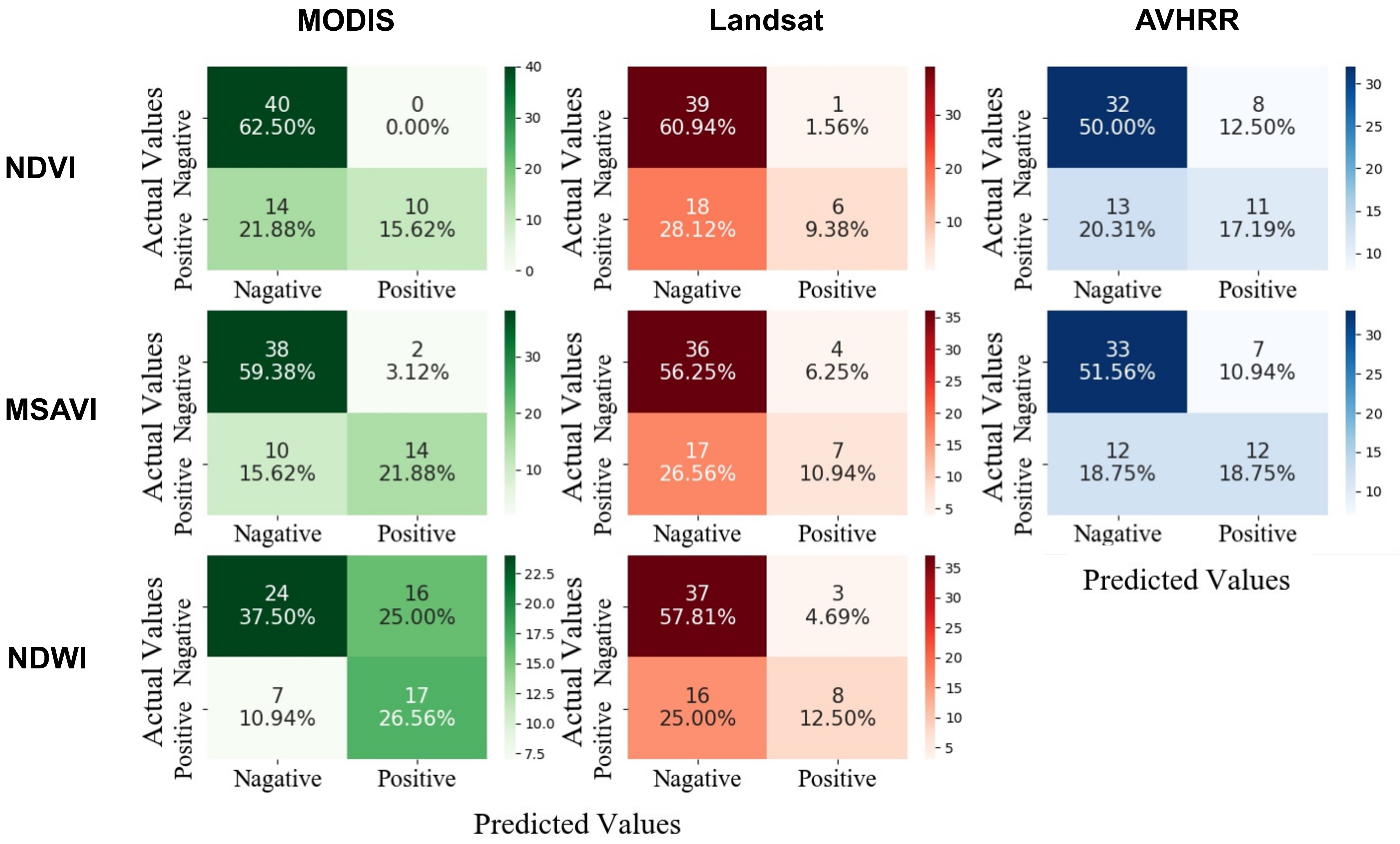

Landsat
AVHRR
NDVI
MSAVI
NDWI
Actual Values
Nagative
Positive
40
58.82%
3
4.41%
16
23.53%
9
13.24%
35
41.67%
22
26.19%
11
13.10%
16
19.05%
39
57.35%
4
5.88%
14
20.59%
11
16.18%
39
46.43%
18
21.43%
12
14.29%
15
17.86%
37
54.41%
6
8.82%
13
19.12%
12
17.65%
Nagative
Positive
Predicted Values

Figure 9.



MODIS
Landsat
AVHRR
NDVI
MSAVI
NDWI
Actual Values
Nagative
Positive
Predicted Values
33
51.56%
7
10.94%
13
20.31%
11
17.19%
37
57.81%
3
4.69%
18
28.12%
6
9.38%
37
57.81%
3
4.69%
18
28.12%
6
9.38%
34
53.12%
6
9.38%
14
21.88%
10
15.62%
34
53.12%
6
9.38%
19
29.69%
5
7.81%
38
59.38%
2
3.12%
19
29.69%
5
7.81%
33
51.56%
7
10.94%
15
23.44%
9
14.06%
36
56.25%
4
6.25%
20
31.25%
4
6.25%

Figure 10.



MODIS
Landsat
AVHRR
NDVI
MSAVI
NDWI
Actual Values
Nagative
Positive
Predicted Values
40
62.50%
0
0.00%
14
21.88%
10
15.62%
39
60.94%
1
1.56%
18
28.12%
6
9.38%
32
50.00%
8
12.50%
13
20.31%
11
17.19%
38
59.38%
2
3.12%
10
15.62%
14
21.88%
36
56.25%
4
6.25%
17
26.56%
7
10.94%
33
51.56%
7
10.94%
12
18.75%
12
18.75%
24
37.50%
16
25.00%
7
10.94%
17
26.56%
37
57.81%
3
4.69%
16
25.00%
8
12.50%

Figure 11.

a **Pre-MODIS**

		Landsat		AVHRR	
		BFAST	**GSI_{anom}**	**BFAST**	**GSI_{anom}**
NDVI	Accuracy	0.69	0.72	0.73	0.6
	Precision	**0.75**	**0.75**	0.62	0.42
	Recall	0.24	0.36	0.37	**0.59**
	F1	0.36	0.48	0.47	0.49
MSAVI	Accuracy	0.66	**0.73**	0.71	0.64
	Precision	0.67	0.73	0.64	0.45
	Recall	0.16	0.44	0.26	0.55
	F1	0.26	**0.55**	0.37	0.5
NDWI	Accuracy	0.63	0.72		
	Precision	0.5	0.66		
	Recall	0.2	0.48		
	F1	0.29	**0.55**		

b **MODIS**

		MODIS		Landsat		AVHRR	
		BFAST	GSI_{anom}	BFAST	GSI_{anom}	BFAST	GSI_{anom}
NDVI	Accuracy	0.69	0.78	0.67	0.7	0.67	0.67
	Precision	0.61	**0.98**	0.67	0.85	0.67	0.57
	Recall	0.46	0.41	0.25	0.25	0.25	0.45
	F1	0.52	0.58	0.36	0.38	0.36	0.5
MSAVI	Accuracy	0.69	**0.81**	0.61	0.67	0.67	0.7
	Precision	0.62	0.87	0.45	0.63	0.71	0.63
	Recall	0.42	0.58	0.21	0.29	0.21	0.5
	F1	0.5	**0.7**	0.29	0.4	0.32	0.55
NDWI	Accuracy	0.66	0.64	0.62	0.7		
	Precision	0.56	0.51	0.5	0.72		
	Recall	0.38	**0.7**	0.17	0.33		
	F1	0.45	0.59	0.25	0.45		

Figure 12.

a

Pre-MODIS

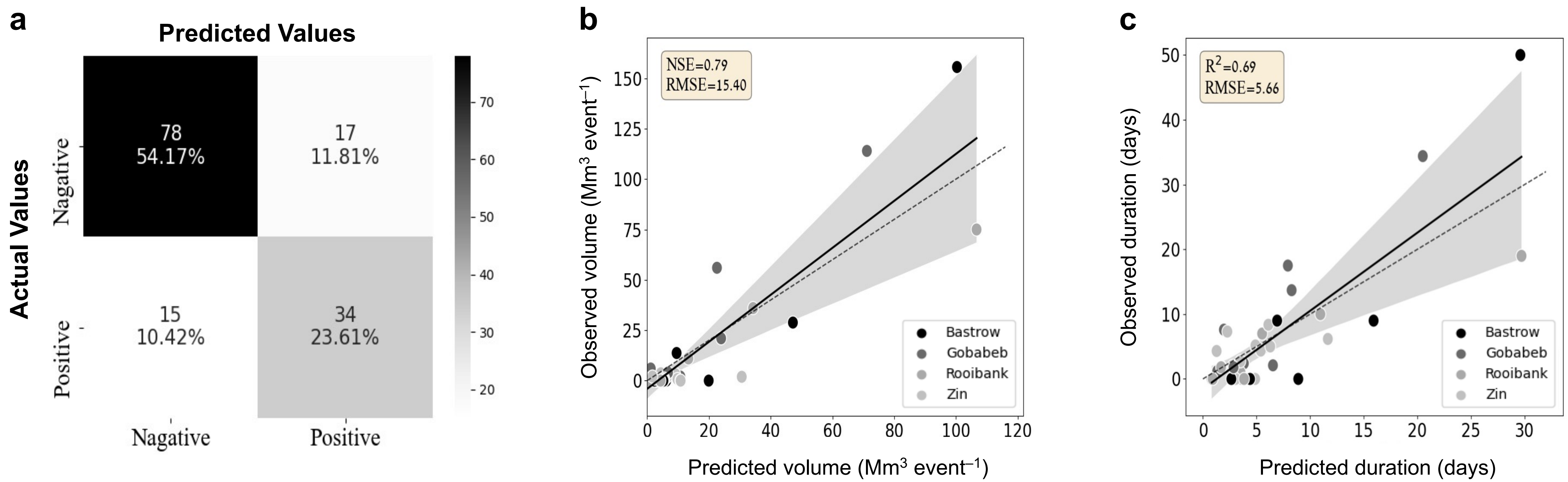
			Landsat		AVHRR	
			BFAST	GSI_{anom}	BFAST	GSI_{anom}
NDVI	Volume (M m^3)	R^2	0.26	0.00	**0.43**	0.12
		RMSE	22.6	17.6	3.25	34.0
	Duration (days)	R^2	0.28	0.00	**0.28**	0.14
		RMSE	5.3	8.58	2.73	15.0
MSAVI	Volume (M m^3)	R^2	0.04	0.02	0.2	0.08
		RMSE	28.0	40.0	26.67	36.7
	Duration (days)	R^2	0.14	0.03	0.15	0.10
		RMSE	6.14	14.9	6.89	16.3
NDWI	Volume (M m^3)	R^2	0.42	0.04		
		RMSE	19.0	36.8		
	Duration (days)	R^2	0.21	0.04		
		RMSE	5.3	13.9		

b

MODIS

			MODIS		Landsat		AVHRR	
			BFAST	GSI_{anom}	BFAST	GSI_{anom}	BFAST	GSI_{anom}
NDVI	Volume (M m^3)	R^2	0.64	0.26	0.22	0.53	0.46	0.36
		RMSE	26.5	41.8	9.55	35.0	40.2	34.6
	Duration (days)	R^2	0.53	0.18	0.25	0.57	0.51	0.39
		RMSE	8.9	12.4	2.33	10.3	11.5	10.2
MSAVI	Volume (M m^3)	R^2	0.52	0.51	0.28	0.38	0.07	0.10
		RMSE	31.9	31.43	37.5	35.2	56.6	41.0
	Duration (days)	R^2	0.41	0.41	0.37	0.42	0.05	0.10
		RMSE	10.2	9.85	10.8	10.5	16.2	12.1
NDWI	Volume (M m^3)	R^2	**0.75**	0.38	0.48	0.01		
		RMSE	16.5	27.6	14.3	44.2		
	Duration (days)	R^2	**0.62**	0.25	0.03	0.01		
		RMSE	5.7	9.8	10.5	14.5		

Figure 13.


a
Predicted Values
Actual Values
Nagative
Positive
78
54.17%
17
11.81%
15
10.42%
34
23.61%
Nagative
Positive
70
60
50
40
30
20
b
NSE=0.79
RMSE=15.40
Observed volume (Mm³ event⁻¹)
Predicted volume (Mm³ event⁻¹)
Bastrow
Gobabeb
Rooibank
Zin
c
R²=0.69
RMSE=5.66
Observed duration (days)
Predicted duration (days)
Bastrow
Gobabeb
Rooibank
Zin