A New Clustering Method to Generate Training Samples for Supervised Monitoring of Long-Term Water Surface Dynamics Using Landsat Data through Google Earth Engine

Alireza Taheri Dehkordi [1], Mohammad Javad Valadan Zoej [1], Hani Ghasemi [2], Ebrahim Ghaderpour [3,4,*] and Quazi K. Hassan [3]

1   Department of Photogrammetry and Remote Sensing, K. N. Toosi University of Technology,
    Tehran 19967-15433, Iran; alireza.tahery@email.kntu.ac.ir (A.T.D.); valadanzouj@kntu.ac.ir (M.J.V.Z.)
2   Department of Civil Engineering, K. N. Toosi University of Technology, Tehran 19967-15433, Iran;
    hanighasemi@email.kntu.ac.ir
3   Department of Geomatics Engineering, University of Calgary, 2500 University Drive NW,
    Calgary, AB T2N 1N4, Canada; qhassan@ucalgary.ca
4   Department of Earth Sciences, Sapienza University of Rome, Piazzale Aldo-Moro, 5, 00185 Rome, Italy
*   Correspondence: ebrahim.ghaderpour@ucalgary.ca or ebrahim.ghaderpour@uniroma1.it

Abstract: Water resources are vital to the survival of living organisms and contribute substantially to the development of various sectors. Climatic diversity, topographic conditions, and uneven distribution of surface water flows have made reservoirs one of the primary water supply resources in Iran. This study used Landsat 5, 7, and 8 data in Google Earth Engine (GEE) for supervised monitoring of surface water dynamics in the reservoir of eight Iranian dams (Karkheh, Karun-1, Karun-3, Karun-4, Dez, UpperGotvand, Zayanderud, and Golpayegan). A novel automated method was proposed for providing training samples based on an iterative K-means refinement procedure. The proposed method used the Function of the Mask (Fmask) initial water map to generate final training samples. Then, Support Vector Machines (SVM) and Random Forest (RF) models were trained with the generated samples and used for water mapping. Results demonstrated the satisfactory performance of the trained RF model with the samples of the proposed refinement procedure (with overall accuracies of 95.13%) in comparison to the trained RF with direct samples of Fmask initial water map (with overall accuracies of 78.91%), indicating the proposed approach's success in producing training samples. The performance of three feature sets was also evaluated. Tasseled-Cap (TC) achieved higher overall accuracies than Spectral Indices (SI) and Principal Component Transformation of Image Bands (PCA). However, simultaneous use of all features (TC, SI, and PCA) boosted classification overall accuracy. Moreover, long-term surface water changes showed a downward trend in five study sites. Comparing the latest year's water surface area (2021) with the maximum long-term extent showed that all study sites experienced a significant reduction (16–62%). Analysis of climate factors' impacts also revealed that precipitation ($0.51 \leq R^2 \leq 0.79$) was more correlated than the temperature ($0.22 \leq R^2 \leq 0.39$) with water surface area changes.

Keywords: k-means; clustering; water; classification; random forests; support vector machines; Iranian dams; reservoirs; long-term

1. Introduction

Unquestionably, water resources are vital for the survival of humans and other creatures and contribute substantially to the development of various agricultural, industrial, recreational, and environmental activities worldwide [1]. Population growth, urbanization, and industrialization have increased water demand, requiring effective management and monitoring of water resources to ensure food security, especially in countries with arid or semi-arid climates such as Iran [2,3]. There have been many problems in different regions of

Iran in the last few years, mainly in the central and southern provinces, due to the decrease in the average annual precipitation, mismanagement, and misuse of water resources [4,5]. Iran's climatic and hydrological conditions and the uneven rainfall distribution make the dam's reservoirs the primary solution for storing surface water flows to supply water for drinking, agricultural and industrial sectors, suppress floods, and generate electricity [6]. Thus, long-term monitoring of water surface dynamics in the reservoirs can inform authorities about the overall change, which is a prerequisite for timely and effective management of water resources [7].

Images acquired by Earth observation satellites have a wide coverage area, potentially a high revisit frequency, and rich spectral information. Thus, they have widely been utilized to map, monitor, and detect spatiotemporal changes in surface water resources in different studies [8]. Several active and passive missions have been used, obtaining data in different wavelengths of the electromagnetic (EM) spectrum, including Advanced Very-High-Resolution Radiometer (AVHRR) [9], Moderate Resolution Imaging Spectroradiometer (MODIS) [10], Visible Infrared Imaging Radiometer Suite (VIIRS) [11], Landsat [12,13], Sentinel [14–16], Spot [17], and IKONOS [18].

There are currently petabytes of remote sensing data available to scientists, researchers, engineers, and decision-makers that can be used for a wide range of applications. Especially for large-scale and long-term studies, these data require huge storage spaces and powerful hardware processing systems [19]. Today, with the development of cloud-based platforms such as Google Earth Engine (GEE), it has been possible to process remote sensing data online without downloading [20]. Various studies showed the effectiveness of GEE in different remote sensing applications such as landcover classification [21,22], wetland detection [23], water quality monitoring [24], flood mapping [19], and impact analysis of drought and floods on croplands [25].

Water mapping algorithms can generally be categorized as threshold-based (TH) or machine learning-based (ML) methods [26]. TH methods apply thresholds to spectral indices derived from remote sensing images. These indices are based on the spectral characteristics of water in different EM wavelengths. There are some popular water-related indices, including Normalized Difference Water Index (NDWI) [27], Normalized Difference Water Index (MNDWI) [28], and (Automated Water Extraction Index) AWEI [29]. MNDWI and AWEI were utilized in GEE to map and monitor surface water, lakes, and dam reservoirs in [30]. Meanwhile, other studies combined several indices to make a knowledge-based decision tree classifier for water mapping [26,31–34]. In TH methods, threshold selection directly affects the results. Furthermore, selecting an appropriate threshold is time-consuming and challenging for different images and study areas [35]. Since water and non-water classes have high inter-class variability, their spectral characteristics may vary spatially and temporally [36,37]. Moreover, TH methods might misclassify low-albedo surfaces such as shadows and asphalts as well [29].

In ML methods, machine learning-based supervised classifiers such as SVM [38], RF [39], and neural networks [40] are used to map surface water. These classifiers learn to automatically distinguish water from non-water areas using training data [41]. The provision of training data, especially in long-term and large-scale studies, is costly and time-consuming [42]. Collecting reference samples through ground field surveys is not possible for the past years. Additionally, some regions might be inaccessible, leading to a non-uniform distribution of samples over the study area. Therefore, some studies developed methods to produce training samples. For example, Refs. [43,44] used global reference maps to generate training samples. They proposed threshold-based methods, which relied on the existence of reference maps at the desired time. It is not guaranteed that their utilized reference maps will continue to be produced in the future. Thus, an automatic method to provide training data that is spatially and temporally robust is still lacking.

Function of the Mask (Fmask) algorithm is a widely used pre-processing technique for cloud, snow, ice, and cloud shadow removal in Landsat imageries, which also provides an initial water map [45]. The Fmask water map can be generated over each Landsat scene,

which can address the challenges mentioned earlier in the previous paragraph. However, Fmask's initial water map may contain errors in both water and non-water classes [46]. This study aimed to propose a novel method that provides training samples for water mapping using supervised classification techniques based on the capabilities of the GEE. The proposed methodology deployed an automated iterative-based K-means clustering refinement procedure on the Fmask initial water map and generated reliable training data. Generated samples were used to train machine learning-based classifiers, RF and SVM, with different features derived from Landsat imageries. Water surface changes of eight dam reservoirs located in three Iran provinces experiencing severe water challenges were investigated using the proposed approach.

The remainder of the paper is structured as follows: first, in Section 2 (Materials and Methods), a complete description of study sites, satellite-based data, and the framework of the article are provided. In Section 3, obtained results are presented and will be further discussed in Section 4 (Discussion). Finally, Section 5 states the conclusions of the article.

2. Materials and Methods

2.1. Study Sites

This study examined long-term changes in the water surface area in eight important and prominent dams in Khuzestan (KHZ), Isfahan (ISF), and Chaharmahal and Bakhtiari (CHB) provinces of Iran. About 11 million people live in these three provinces, which is approximately 13% of Iran's population. Multiple rivers, such as the Karun, Zayandehrood, Karkheh, Dez, Jarahi, Marun, and Golpayegan rise and flow in CHB, ISF, and KHZ. However, these provinces face severe problems in supplying drinking and agricultural water [5]. In addition, these three provinces have a significant role in transferring water to the central provinces of Iran, such as Yazd, Qom, and Kerman. Therefore, water problems in these three provinces directly affect other provinces in the central regions of Iran. Figure 1 shows Iran's location and ISF, CHB, and KHZ in the central and southern regions. It also depicts the location, Shuttle Radar Topographic Mission's (SRTM's) Digital Elevation Model (DEM), the corresponding basin of each study site, the latest high-resolution Google Earth image of dams, and some photos taken during extensive filed visits in 2019, 2020, and 2021 [47,48]. A summary of the characteristics of each dam is presented in Table 1. This study investigated changes in water surface area one year after the opening date of each dam until 2021 for each dam. In the case of older dams, due to the limited number of satellite images, changes were not examined before 1990 [26].
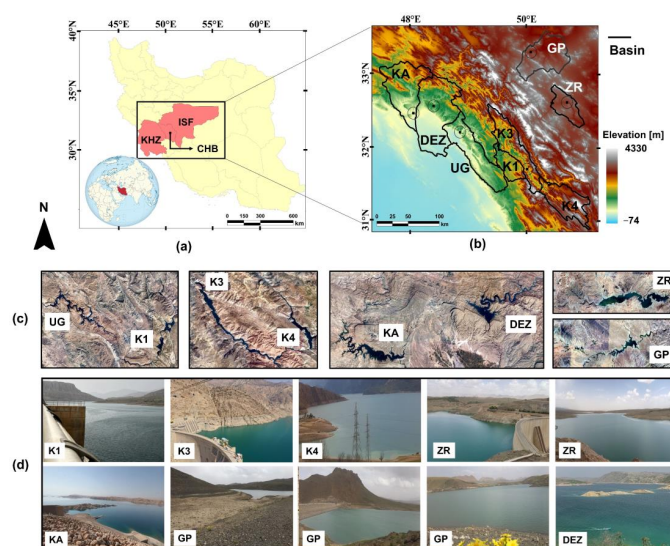

Figure 1. (**a**) Iran's location and ISF, CHB, and KHZ provinces, (**b**) spatial distribution of the study sites and corresponding basins, (**c**) latest Google Earth satellite imagery of each study site, and (**d**) some photos taken (excluding UG) during the extensive field visits in 2019, 2020, and 2021.

Table 1. Characteristics of the study sites.

Site	River	Opening Date	Catchment Area (km²)
Karun-1 (K1)		1976	1436.6
Karun-3 (K3)	Karun	2005	3292.3
Karun-4 (K4)		2010	1177.5
Upper Gotvand (UG)		2012	3904.1
KArkhe (KA)	Karkhe	2001	4318.5
DEZ	Dez	1963	6268.6
GolPayegan (GP)	Anaarbar	1970	3533.7
ZayandeRud (ZR)	Zayenderud	1971	1599.3

2.2. Satellite-Based Data

National Aeronautics and Space Administration's (NASA's) Landsat program is the longest-running Earth observation optical satellite. This study analyzed long-term water extent changes using orthorectified reflectance images of Operational Land Imager (OLI), Enhance Thematic Mapper-plus (ETM+), and Thematic Mapper (TM) sensors from the Landsat 8 (L8), 7 (L7), and 5 (L5) missions, respectively. The Landsat Ecosystem Disturbance Adaptive Processing System (LEDAPS) algorithm was used to obtain reflectance in TM and ETM+ sensor images [49]. In the case of OLI sensor images, atmospheric corrections were applied using the Land Surface Reflectance Code (LaSRC) algorithm [50]. The current study used six bands of Blue, Green, Red, NIR, SWIR1, and SWIR2 of these images. Six Landsat scenes covered the study sites, as shown in Figure 2a. Figure 2b also illustrates the number of images used by each mission from 1990 to 2021 (with cloud coverage of less than 10%). We monitored water changes in eight study sites using about 800 Landsat satellite images in this study. This large number of images can only be processed with cloud-based platforms such as GEE or powerful hardware systems.
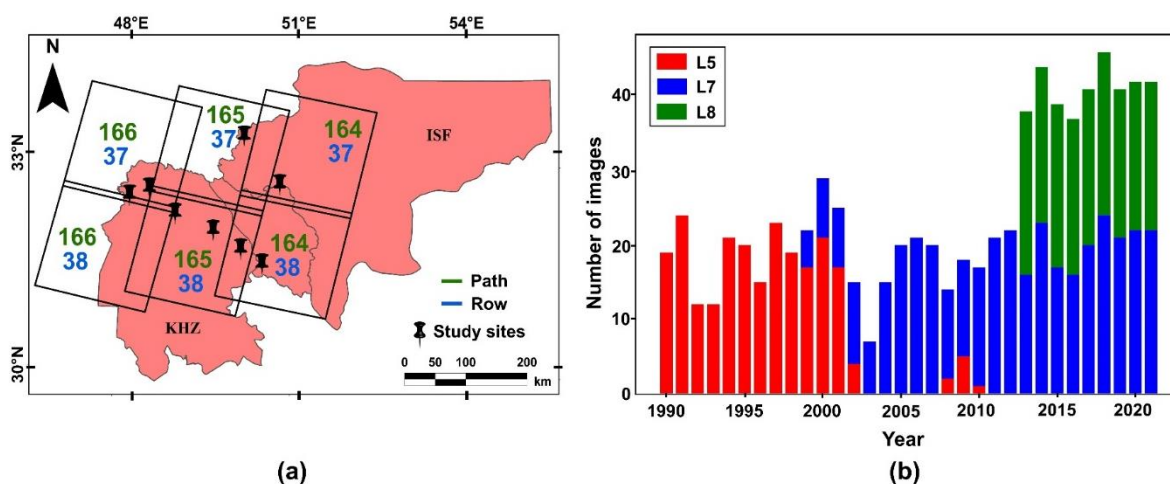

Figure 2. (**a**) Landsat scenes over the study sites, and (**b**) number of used Landsat scenes (could cover <10%) in each year.

The amount of water stored in dams is influenced by two important climate factors, precipitation, and temperature [26,33,51]. The Famine Early Warning Systems Network (FEWS NET) Land Data Assimilation System (FLDAS) data was used to examine the relationship between surface water area changes in each dam, Mean Annual Temperature (MAT), and Mean Annual Precipitation (MAP). FLDAS is a monthly freely available product that uses Noah version 3.6.1 surface model with CHIRPS-6 hourly rainfall and is downscaled using the NASA Land Surface Data Toolkit [52]. It has provided information

on climate-related variables since 1982. The Shuttle Radar Topography Mission (SRTM) 30-m Digital Elevation Model (DEM), "USGS/SRTMGL1_003", was also used to produce a hill shadow mask of the study sites [48,53]. Furthermore, we investigated changes in built-up areas based on MODIS Land Cover Type Yearly Global (MLCTY) 500 m data. MLCTY was the only freely available yearly global landcover map with a long history record (from 2001 to 2020) at the time of conducting the research [54]. Changes in built-up areas can indicate population growth, directly affecting water demand [55].

2.3. Framework

Figure 3 gives an overview of this study's framework, designed based on GEE capabilities and fully implemented in this platform. Each step is described in the following sections.
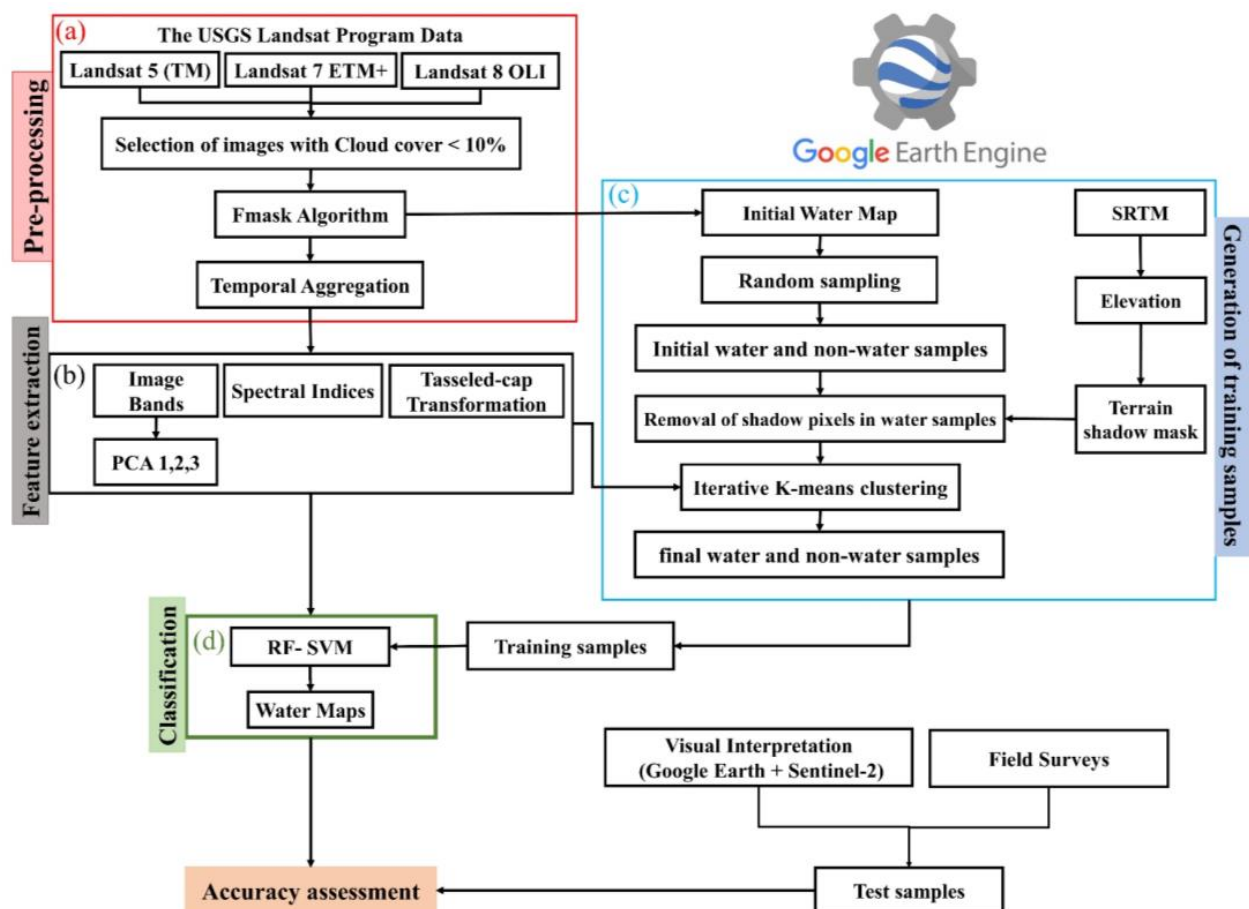

Figure 3. Schematic diagram of the framework of this study: (**a**) Pre-processing, (**b**) Feature Extraction, (**c**) proposed method for generating training samples, and (**d**) Classification.

2.3.1. Pre-Processing of Landsat Data

Since the study area was large and covered by multiple Landsat scenes, cloud-free composite images had to be generated first. High cloud-covered optical images were practically unusable, so we used images with less than 10% cloud coverage. In order to have more cloudless images, we used images acquired from August to September each year. This period, coinciding with the end of summer, prevented the wrong estimations caused by snow and ice. It was also the shortest period for generating a complete image mosaic from study sites in all years (1990–2021). In addition, since the lowest water levels were mainly reported at the end of summer in the reservoirs, this period could also estimate the annual permanent water surface area in each reservoir [56]. For each Landsat scene, low-quality pixels, clouds, cloud shadows, snow, ice, and Scan-Line Corrector (SLC)-Off Gaps were masked out using the Fmask algorithm [45]. Then, we used the temporal

aggregation technique with a median filter to produce cloud-free image mosaics. In this technique, the median value of each image band is kept, and the possible noise in the initial images can be reduced in the output composite [57]. This process was conducted for all years between 1990 to 2021, resulting in a cloud-free composite over study sites with six different bands (Red, Green, Blue, NIR, SWIR1, and SWIR2) for each year.

2.3.2. Feature Extraction

Three feature sets were extracted from generated cloud-free composites in the previous step. Using different features could help boost the accuracy of water mapping because they were used directly to generate training samples and detect water in the study sites. The first feature set consisted of the top 3 components after the Principal Component Analysis (PCA) transformation of image bands, containing more than 95% of initial Landsat spectral bands information [58]. Additionally, we derived four spectral indices (SIs), Normalized Difference Vegetation Index (NDVI) [59,60], MNDWI [28], (Enhanced Vegetation Index) EVI [60,61], and AWEI [29]. Many studies extensively used these indices to map water in remote sensing images [33,62,63]. It should be highlighted that several combinations of different spectral indices were tested, among which the combination of NDVI, MNDWI, EVI, and AWEI could achieve higher classification accuracies. Moreover, the Wetness, Greenness, and Brightness components of the TC transformation were used as the third feature set [64,65].

2.3.3. Proposed Method for Generating Training Samples

This study developed an automated method to generate training samples from both water and non-water classes. As mentioned, the Fmask algorithm was used to remove cloud, cloud shadow, snow/ice, and low-quality pixels in Landsat images. FMask is a multi-pass algorithm that uses decision trees to label pixels in the scene prospectively; it then validates or discards those labels according to scene-wide statistics. The algorithm provides a binary map of each image's water and non-water areas, which was used to provide initial candidate training samples. However, based on visual interpretation and quantitative assessments (provided in Section 3.1), this initial binary map contained some inevitable errors in both classes. Consequently, refinement processes were required to obtain accurate and reliable final samples. First, we randomly selected initial samples from both water and non-water classes from the Fmask water map. It is worth noting that initial samples in each class were randomly selected with different initial seeds (random number generators) to contain diverse land cover classes, which avoided bias in the classification model. Due to the possibility of confusion between water and hill shadows, SRTM elevation data were used to produce a hill shadow mask [66]. The produced mask was used to remove shadow pixels that were considered water samples in the initial training dataset. Water and non-water samples have distinct spectral behaviors, so in the case of clustering, samples of each class must also be grouped together in the same cluster. Otherwise, they are wrong and have to be removed from the process. In other words, there were two groups of initial water and non-water samples. Due to the different spectral signatures of the water and non-water samples, these two groups must represent two separate clusters. Thus, in the case of clustering, water samples that were clustered in the group of non-water samples were mistakenly detected as water by the Fmask algorithm. Non-water samples clustered in the water group were also incorrectly identified as non-water samples by the Fmask algorithm. As a result, they must be excluded from the initial training set. We used the iterative K-means clustering method to refine the selected sample [67]. Inputs to K-means were the initial sample set and features of Landsat images (PCA, SI, and TC) for each year. Initial samples were clustered in each feature space by K-means in the first iteration, and the wrong samples were removed. In the second iteration, the filtered samples from the previous step were clustered again, and the wrong samples were removed. The process continued until no other sample was removed from each cluster.

2.3.4. Classification

In this step, extracted features were classified using the training samples of each class from the last step and two machine learning algorithms, Random Forest (RF) and Support Vector Machines (SVM). Both methods showed satisfactory performance in related research [38,39]. RF is an ensemble classifier consisting of many decision trees generated during training, and the final class of a sample is determined based on majority voting among these trees [68]. SVM is also a linear classification method that uses the training data to find the optimal separating hyperplane between classes. If target classes are not linearly separable, the kernel trick is used to project the input feature space to higher dimensions, in which the data can be separated linearly [69]. In order to determine the input parameters of RF and SVM methods, we combined the suggestions of previous studies and quantitative analysis to achieve the highest classification accuracy on the test samples. In the RF method, two parameters of "number of trees" and "number of features per split" were considered equal to 100 and the square root of input features, respectively [70,71]. Additionally, in the SVM method, we used the Radial Basis Function (RBF) as the kernel and set gamma to 1 and regularization parameter (C) to 10 [72]. Classification algorithms were trained on 5000 training samples from each class in each year, produced by the proposed automated methodology. Training samples were equally distributed in all the study sites, which led to training one classification model (RF or SVM) rather than eight classification models each year for eight study sites. A trial-and-error strategy was employed to determine the appropriate number of training samples (for values between 1000 and 10,000, with a step of 1000, to get the highest overall classification accuracy on the validation samples). Increasing the number of training samples to more than 5000 did not lead to a significant increase in classification accuracy [73,74]. It should be highlighted that the same number of samples (5000) derived from the initial Fmask water map were also directly used to train classification models to investigate the effectiveness of the iterative clustering-based refinement procedure.

2.3.5. Accuracy Assessment

Finally, the performance of the adapted RF and SVM models was evaluated using test data. Test data were obtained from extensive field visits and visual interpretation techniques using high-resolution satellite imageries such as Google Earth and Sentinel-2 data in three recent years (2019, 2020, and 2021) [33,75,76]. Field visits were conducted in August and September of each year (see Figure 1d). A total of 2000 evaluation samples were provided per class each year. As test samples were prepared by a completely different procedure than training samples, it was possible to carefully examine how the proposed methodology performed in generating training samples. Test samples had no involvement in the training process and were also used for the accuracy assessment of the Fmask initial water map and comparison of our results with global reference maps.

Different parameters of classification's confusion matrix (Table 2), including Overall Accuracy (OA), kappa coefficient, and User accuracy of each class (UA_w and UA_{nw}), were used to quantitatively evaluate the results (Equations (1)–(3)) [77,78]. It should be highlighted that in a 2-class confusion matrix, UA of one class is equal to the producer accuracy (PA) of the other class. True positives (TP) and true negatives (TN) represent the correct classification of water and non-water pixels. A false positive (FP) is the incorrect classification of water as non-water, whereas a false negative (FN) is the incorrect assignment of non-water to the water class.

$$OA = \frac{TP + TN}{TP + FP + FN + TN} \tag{1}$$

$$UA_w = \frac{TP}{TP + FP} \, , \, UA_{nw} = \frac{TN}{TN + FN} \tag{2}$$

$$Kappa = \frac{OA - P_e}{1 - P_e} \, , \, P_e = \frac{(TP + FN)(TP + FP) + (TN + FN)(TN + FP)}{TN + FN} \tag{3}$$

Table 2. Confusion matrix of classification (TP = True Positive, TN = True Negative, FP = False Positive, and FN = False Negative).

		Test Data	
		Water	**Non-Water**
Classification Result	**Water**	TP	FP
	Non-Water	FN	TN

We also employed McNemar's distribution-free statistical analysis to examine the efficiency of various input features in water mapping. McNemar's test compares the error matrices of two classification results to determine how they have been improved [79]. In doing so, the McNemar test calculates the z value:

$$z = \frac{f_{12} - f_{21}}{\sqrt{f_{12} + f_{21}}} \tag{4}$$

where f_{12} is the number of correctly classified pixels by the first classifier while incorrectly classified by the second classifier and f_{21} is the number of correctly classified pixels by the second classifier while incorrectly classified by the first classifier. z follows an χ^2 distribution with a corresponding probability (*p*-Value). The more χ^2 values and lower *p*-Value, indicates the significant improvement between the two classifiers.

T-test statistical analysis was also applied to investigate the statistical significance of surface water, Mean Annual Precipitation (MAP), and Mean Annual Temperature (MAT) change slopes at three confidence levels (* or 90% or $\alpha < 0.1$, ** or 95% or $\alpha < 0.05$, and *** or 99% or $\alpha < 0.01$).

3. Results

3.1. Accuracy of Water Mapping

Figure 4 illustrates the RGB Landsat composites, initial Fmask water maps, and classification maps obtained from training RF model with initial samples of Fmask without conducting the proposed iterative refinement procedure for 2021 in KA, DEZ, K1, K3, and K4 reservoirs. It also depicts the quantitative accuracy assessment of the Fmask initial map (Figure 4d) and RF classification map (Figure 4e) using test data (see Section 2.3.5) in 2021, 2020, and 2019 in all study sites. As can be seen, the Fmask initial map and classification map of RF trained with initial Fmask samples contained significant misclassifications. There was approximately an equivalent performance of both maps in all years. Moreover, UAs of water and non-water classes did not exceed 82% in all years. Thus, if Fmask initial water map was directly used to get final reference samples, there might be wrong samples belonging to the other class. As a result, low performance of RF classification model was observed (OA < 79%) when using the Fmask initial samples without refinement procedures. Thus, the iterative K-means refinement procedure was utilized to refine candidate water and non-water samples to obtain high-reliability training samples in water and non-water classes.

Figure 5 depicts the distribution of the initial and final training samples of water and non-water classes in two three-dimensional feature spaces, SI and TC. Water and non-water samples were expected to be separated in different feature spaces because of their distinct spectral signatures (left scatter plot in Figure 5a,b). However, some samples belonging to the water class behaved similarly to non-water samples and vice versa because of initial water map errors (see Figure 4). As a result, there was a low performance of the RF model trained with the initial Fmask samples (Figure 4), indicating that direct use of the Fmask initial water map could not generate accurate maps. After deploying the proposed iterative clustering-based procedure, the incorrect samples were removed by using the iterative K-means method, and the final samples were completely separated in both spaces (right scatter plot in Figure 5a,b).
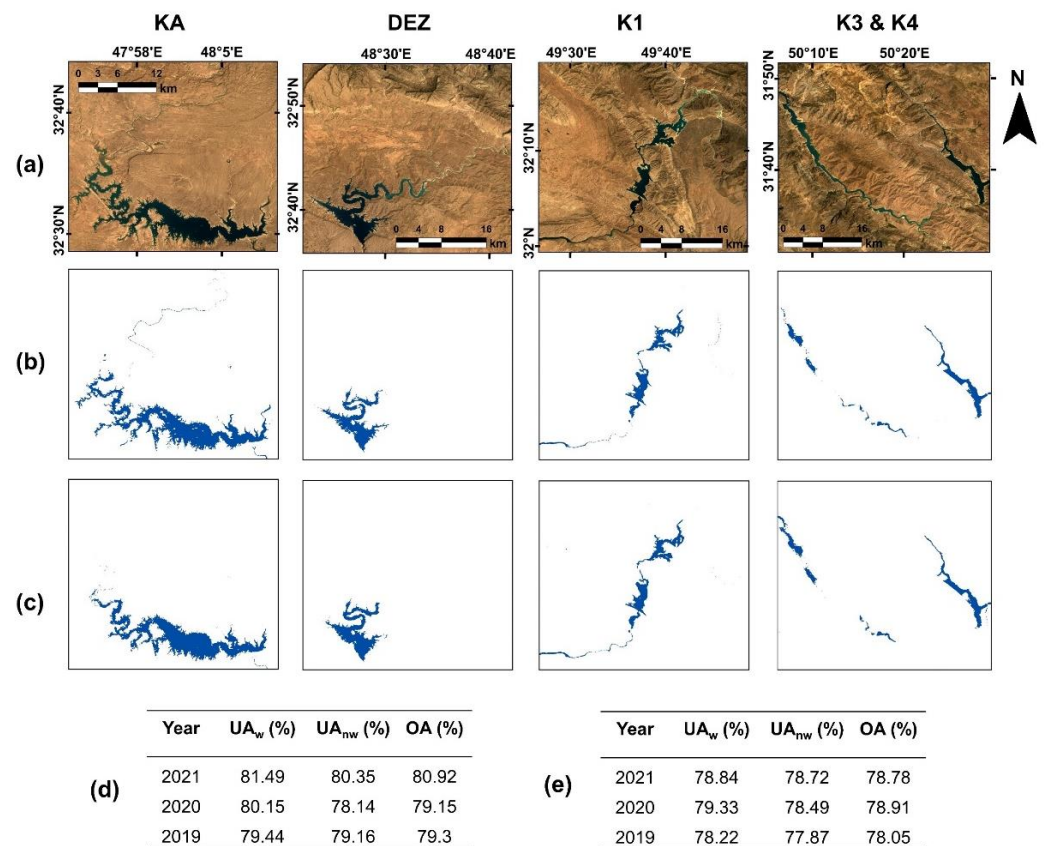
Year	UA_w (%)	UA_{nw} (%)	OA (%)
2021	81.49	80.35	80.92
2020	80.15	78.14	79.15
2019	79.44	79.16	79.3

Year	UA_w (%)	UA_{nw} (%)	OA (%)
2021	78.84	78.72	78.78
2020	79.33	78.49	78.91
2019	78.22	77.87	78.05

Figure 4. KA, DEZ, K1, K3, and K4 study sites (**a**) Landsat RGB composite, (**b**) Fmask initial water map, (**c**) Classification maps of RF model trained with all features and initial Fmask samples without conducting the proposed iterative clustering-based refinement procedure. The values of UA_w, UA_{nw}, and OA for (**d**) Fmask initial water map and (**e**) classification map of adapted RF in 8 study sites using test samples (see Section 2.3.5).
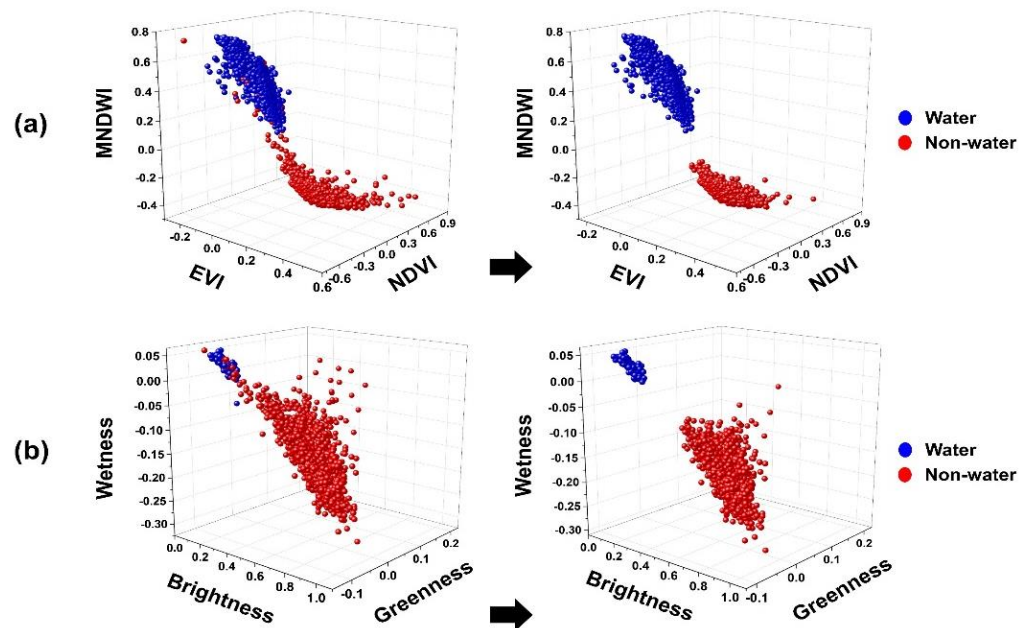

Figure 5. Distribution of the initial and final training samples in two 3D feature spaces, (**a**) x: EVI, y: NDVI, z: MNDWI, (**b**) x: Brightness, y: Greenness, z: Wetness.

Generated training samples were used to train RF and SVM models. Trained classifiers were validated for three consecutive years, 2021, 2020, and 2019 using test samples (see Section 2.3.5). As test samples were prepared by a completely different procedure than training samples, it was possible to carefully examine how the proposed methodology performs in generating training samples. Figure 6 shows the quantitative accuracy parameters (UA_w, UA_{nw}, OA, and kappa) for classifying different input properties (SI-, PCA-, TC-only, and all features: SI, PCA, and TC) using two classification techniques, RF and SVM, in 2021, 2020, and 2019. The followings can be concluded.
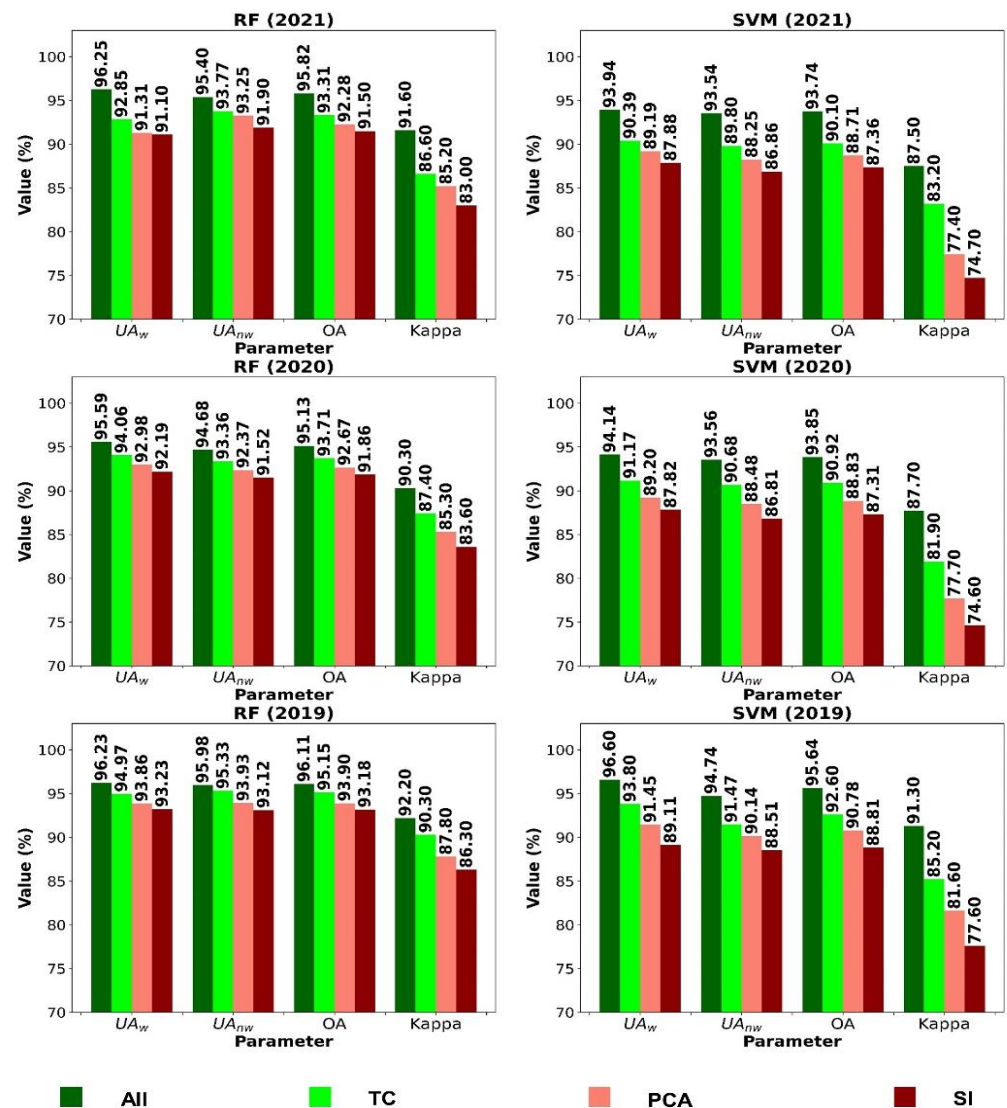
Figure 6. Accuracy assessment of the adapted RF and SVM models trained with generated samples of the proposed methodology in 2021, 2020, and 2019.

(1)   High classification OAs of RF and SVM in all three years (at least 93.74% by SVM in 2021) indicated the satisfactory performance of the proposed methodology for the generation of training samples. The adapted RF using all features and direct initial samples of Fmask resulted in OAs of around 80% (Figure 4). After deploying the proposed methodology, OAs have increased by about 13%. As a result, the proposed iterative clustering-based refinement procedure could effectively remove outliers of initial samples from the Fmask water map and generate reliable training samples.

(2)   Using all input features for classification led to the highest classification accuracy in each year. For example, the OA values for the RF classifier were 95.82, 95.13, and 96.11% in 2021, 2020, and 2019, respectively. The TC features achieved higher accuracy

than PCA and SI in all three years. In other words, the TC transformation resulted in a better separation of target classes. Figure 5 also illustrated a better separation of reference samples of both classes in the TC feature space relative to SI. SI showed the weakest performance in all three years compared to the other two input features. PCA transformation ranked second. To sum up, using all features led to an increase of at least 1% (2019) and a maximum increase of 2.5% in OA (2021) compared to the best feature set (TC).

(3) The RF model outperformed the SVM method in all parameters for each feature set. For example, in the case of classifying all features in 2021, RF achieved an overall accuracy of at least 0.5% greater than SVM in the same year. For the RF method, the highest OA was related to 2019 with 96.11%, and the lowest was related to 2020 with 95.13. Meanwhile, SVM's highest and lowest OAs were 95.64 % and 93.74% in 2019 and 2021, respectively. Also, RF achieved higher UA in both water and non-water classes than SVM. RF's UA was 96.25 in 2021 for water classes, whereas it was 95.4 for non-water classes. They were 93.94% and 93.54% for SVM.

Furthermore, Table 3 highlights the McNemar's test results between classification maps obtained from all features (All) against TC-, PCA-, and SI-only using the RF model based on test samples. In most cases, utilizing all features significantly improved the classification results (p-Values < 0.001). Comparing TC features to other input features (PCA and SI), TC features produced the closest results to the "All" case, as lower χ^2 values were reported (2021: 25.89, 2020: 9.97, and 2019: 7.48). Due to the better performance compared to the other cases, our results were derived by the trained RF classifier using all features and training samples from the proposed methodology in the following sections.

Table 3. Results of the McNermar's test between classification results of different input features (All vs. TC, PCA, and SI) for RF classifier.

	All vs. TC		All vs. PCA		All vs. SI	
Year	χ^2	*p*-Value	χ^2	*p*-Value	χ^2	*p*-Value
2021	25.89	<0.001	41.88	<0.001	106.51	<0.001
2020	9.97	<0.01	23.14	<0.001	57.63	<0.001
2019	7.48	<0.01	19.66	<0.001	37.15	<0.001

3.2. Long-Term Change Analysis

The proposed methodology utilized the initial water map of Fmask to generate training samples. An initial Fmask water map can be generated from any Landsat image using the Fmask algorithm. Hence, the proposed method was temporally transferable to generate training samples, meaning it could be applied at different years. Thus, the proposed framework was deployed to examine the long-term surface water area of study sites in this section. Figure 7 presents the computed long-term surface water area in our study sites. Results indicated that study sites could be ranked as follows based on the average area of the lake surface: 1-KA, 2-UG, 3-DEZ, 4-K1, 5-K3, 6-ZR, 7-K4, and 8-GP. KA and GP reservoirs with an average area of 72.71 km^2 and 1.8 km^2 were the largest and smallest, respectively. As can be seen, out of eight study sites, five dams (K1, K3, DEZ, ZR, and GP) showed a general downward trend, among which ZR and GP experienced a more dramatic decline. Their area was estimated at about 15 km^2 and 1 km^2 in 2021, respectively. In comparison, it was about 36 km^2 and 2.7 km^2 in 1990, which showed an approximately 60% fall. The area of the ZR reservoir has always been less than the long-term average since 2008. Moreover, the current area (2021) of K1, K3, DEZ, ZR, and GP has decreased by about 5, 5.5, 12, 9.5, and 0.8 km^2 compared to their long-term averages (μ). In the case of older dams (K1, DEZ, ZR, and GP), the largest surface area of the reservoir occurred before 2000, and in the last 20 years, the water surface area has always been less than the wettest year (the year with the largest water surface area, which is shown by the green bar in Figure 7).

For example, the largest area of the K1 and DEZ was observed in 1997 and 1992, while it was 1993 for ZR and GP.
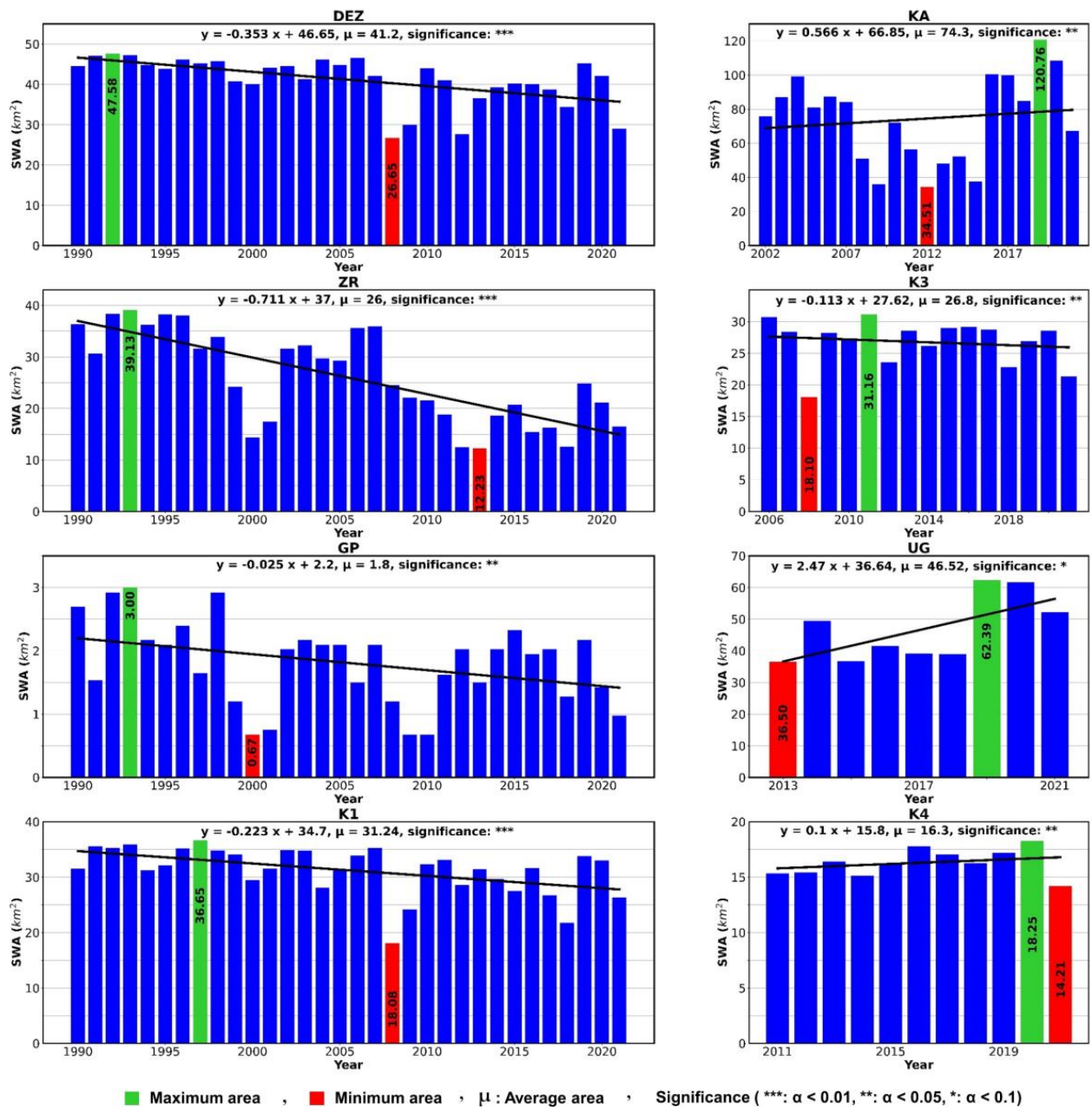

Figure 7. Long-term change analysis of water surface area in the study sites.

The overall trend for UG, K4, and KA dams, built after 2000, was upward. This uptrend was slight for the K4 dam (trend line slope: 0.1). In comparison, the UG dam showed the sharpest upward trend among our study sites. However, the water area in all dams experienced a significant downward trend in the last three years. Comparing the latest year (2021) and the wettest year of each dam, our study sites showed 39% (DEZ), 16% (UG), 62% (GP), 28% (K1), 31% (K3), 22% (K4), 44% (KA), and 58% (ZR) reduction in the reservoir area, respectively. Furthermore, the t-test statistical analysis also indicated that all dams have a changing slope of at least 90% confidence. DEZ, ZR, and K1 change slopes were statistically significant at 99% ($\alpha < 0.01$) confidence level, indicating that surface water area would decrease by about 3.5, 7.1, and 2.25 km^2 in the upcoming decade, respectively.

UG change slope had the lowest confidence compared to the other dams. The level of confidence in other reservoirs was 95% ($\alpha < 0.05$).

As mentioned earlier, all dams had a lower water surface area in 2021 than in the wettest year. There was at least a 16% decrease (UG dam) in the reservoir's current area in all study sites compared to the maximum. The dam's reservoir area decreased from the outer edges of the lakes, which were usually shallow, and the areas that have remained unchanged are the deepest [63,76]. Comparisons showed the current critical condition of some dams. In ZR and GP dams, about 58% and 62% of the reservoir vanished compared to the wettest year. KA and DEZ also experienced a dramatic decrease of about 40%.

ZR had the sharpest downward trend among study sites. Taking ZR as an example, Figure 8 demonstrates the surface water extent changes between the maximum (1993), minimum (2013), and latest year closest to the long-term average area (2019). As can be seen, 12.23 km^2 of water was detected in all three dates and there was 26.9 km^2 disappearance of the water extent in 2013 with respect to 1993.
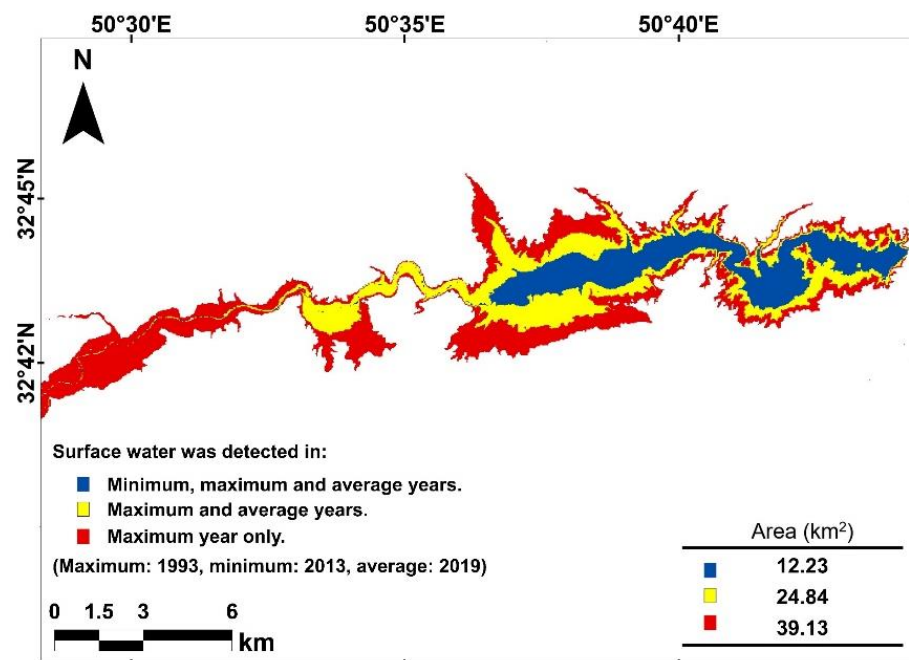

Figure 8. Surface water extent changes in ZR study site between maximum area in 1993, minimum area in 2013, and the latest year closest to the average long-term area in 2019.

3.3. Influence of Precipitation and Temperature

As mentioned in the satellite-based data Section 2.2, we used FLDAS data to get the Mean Annual Precipitation (MAP) and Mean Annual Temperature (MAT) every year for each dam's basin from September to September (Figure 9). Here, we investigated the relationship between MAP and MAT with the reservoir's water surface area. MAP trended downward in K1, DEZ, ZR, K3, and GP dams, which were exactly the ones experiencing a decline in water surface area. In the case of KA, UG, and K4 dams, showing an upward trend in the water surface area, the MAP overall trend was also growing. Therefore, the overall trend of MAP and water surface area was similar in our study sites. Additionally, there was mainly an increase in the water surface area where the MAP peaks. For example, 2019 was one of the rainiest years, causing destructive floods in Iran's southern and central regions [80]. The increase in rainfall was directly reflected in the water surface area in 2019. As shown in Figure 7, most dams in 2019 had the highest water surface area in recent years. It should be noted that MAP slopes were statistically significant at a confidence level of at least 90%.
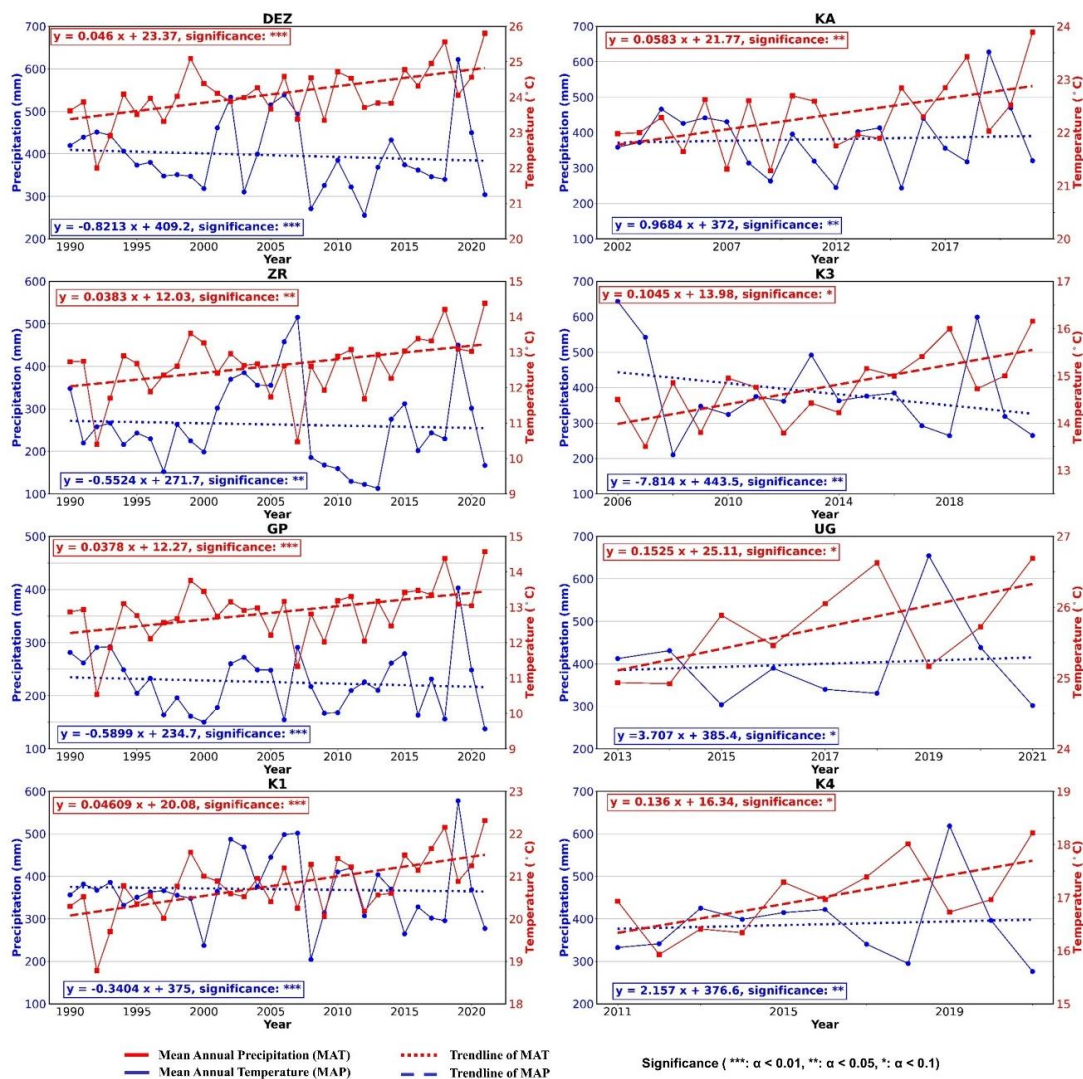
Figure 9. Mean Annual Precipitation (MAP) and Mean Annual Temperature (MAT) in the corresponding basins of the study sites.

MAT showed an increasing trend in both growing and declining dams. K1, DEZ, ZR, K3, and GP dams, having a downward trend in water surface area, experienced an upward trend in MAT as well. Similarly, MAT increased in KA, UG, and K4 dams with an upward water surface area trend. Moreover, ZR and GP dams, located in ISF, had lower MAT than other study sites, mostly located in KHZ, one of the warmest regions in Iran. In Table 4, we used the coefficient of determination (R^2) parameter to investigate the correlation between MAP and MAT with the water surface (R^2_{MAP} and R^2_{MAT}). As can be seen, there was a higher correlation between MAP and water surface area, which ranges between 0.51 (K4) to 0.79 (DEZ). In comparison, MAT showed less correlation with water surface area, ranging between 0.22 (KA and K3) to 0.39 (ZR).

Table 4. Long-term correlation analysis between MAP and MAT with water surface area in the study sites.

Site	R^2_{MAP}	R^2_{MAT}	Site	R^2_{MAP}	R^2_{MAT}
K1	**0.69**	**0.29**	**K3**	**0.54**	0.22
DEZ	0.79	0.23	**UG**	0.73	0.29
ZR	0.63	0.39	**K4**	0.51	0.27
KA	0.6	0.22	**GP**	0.65	0.31

4. Discussion

This study developed an approach for providing reliable training data for supervised machine learning models. Our method did not rely on the existence of reference maps and used Fmask initial water map to generate training samples. Due to the feasibility of applying Fmask on any Landsat image, it could provide training data for supervised classifiers for any Landsat scene. Both quantitative and visual analysis showed that Fmask initial map contained some errors in both classes (Figure 4). Moreover, the direct use of Fmask samples to train classification models resulted in low classification accuracies (Figure 4). Therefore, it was essential to use refining processes to generate accurate training samples. Iterative K-means clustering was used to refine initial random samples of both classes. Initial and final training samples were compared in two 3D feature spaces, SI and TC (Figure 5). Due to the errors in the initial map, some samples of each class behaved similarly to the other. However, the wrong samples were eliminated by iterative K-means, and final training samples were separated after refining the initial points using the proposed methodology in 3D feature spaces. RF and SVM classifiers were trained using the generated samples and achieved high quantitative accuracies in three recent years (OA $\geq$ 93.39%) (Figure 6), indicating the satisfactory performance of the proposed methodology for training samples generation.

The performance of RF and SVM methods in water mapping was also compared (Figure 6). Both methods achieved high OAs and kappa coefficients according to the results. However, the RF method outperformed SVM. Other studies also showed the better performance of the RF [81,82]. We also evaluated the efficiency of different input features in water mapping (Figure 6). The results show that simultaneous use of all features (TC, PCA, and SIs) increased the classification accuracy. TC-only features achieved the highest classification accuracy. Additionally, PCA-only features outperformed SI-only, the same as other studies [83]. The superiority of TC features to the SIs was also reported in [84]. A better separation was also observed between water and non-water samples in the 3D feature space of TC than in SIs space (Figure 5).

The long-term change analysis of water surface area in study sites demonstrated a general downward trend for five dams (K1, K3, DEZ, ZR, GP), among which ZR and GP experienced a more severe decrease (Figure 7). K4, KA, and UG dams had an increasing overall trend. However, comparing the latest year (2021) and the wettest year of each dam, all study sites showed significant reductions in the reservoir area, ranging from 16% (UG) to 62% (GP).

Relationship analysis between MAT and MAP with water surface area indicated that MAT is less correlated with water surface area, which revealed that precipitation had a more significant influence on the reservoir water surface area than the temperature [33,76] (Table 4). The results showed that the overall trend of MAP acted the same as the overall trend of water surface area changes (compare Figures 7 and 9). In addition, an increase or decrease in surface water area was the same as an increase or decrease in MAP. Note that the heaviest annual precipitation did not necessarily lead to the greatest water extent. For example, the heaviest annual precipitation in GP and DEZ basins occurred in 2019 (Figure 9). However, the maximum surface water extent was not observed in 2019 (Figure 7) because there were other factors to consider, such as water demand and dam topography.

Floods and drought events directly affected the dam lake area. For example, 2019 floods in Iran's southern and central regions resulted in a sharp increase in water surface area in most study sites [80]. Additionally, between 1999 to 2001, the water surface area of ZR and GP dams, located in ISF, experienced a dramatic decline. In 2000, GP dam had the lowest water surface area in the recent 31 years and the lowest lake areas of ZR occurred in 2000, 2012, and 2013. Some studies pointed out the most severe droughts in ISF in the mentioned years [85–87]. A severe drought in KHZ province was also reported in 2008 by [88], when K1, K3, DEZ, and KA Dams had low water surface area.

As depicted in Figure 1b, four K3, K4, K1, and UG reservoirs have been constructed on the Karun River. However, they showed different overall trends. Mentioned dams

were arranged sequentially, so the amount of water they contained was directly affected by previous reservoirs [89]. Therefore, the upward slope of K4 could be a result of its location atop other dams. K4 dynamics were tracked only from 2011 to 2021 since it was opened in 2010. The short 10-year period of monitoring and flood events of 2019 may be other contributing factors. K4 and K1, located after K3, showed a downward trend. Since the Khersan River also flows into the reservoir of this dam, K4 experienced a more gradual decline than K1. UG, the last dam built on the Karun River, showed an upward trend. It could also be due to the Lali River joining Karun just before UG. However, in the UG, the short period of monitoring water dynamics and 2019 flood events should also be considered.

Reservoir regulation pattern is also an important element for change analysis. Regulating reservoirs, mainly divided into annual and multi-year, have a certain storage capacity that can be used to regulate the inflow process [90]. The regulating performance corresponds to the available storage capacity. The annual regulating reservoir can well regulate the water inequality within the year. The multi-year regulating reservoir can achieve the water distribution between years [91].

Water demand is another element that influences the amount of water stored in dam reservoirs. Our investigations indicated a rising trend at a 99% confidence level in the total area of built-up regions in the three provinces of ISF, KHZ, and CHB using the MLCTY product. A growing population is directly proportional to an increasing built-up area [55]. As a result of population growth, water demand in various drinking, agricultural, and industrial sectors increase. Therefore, the downward trend in precipitation, and the upward trend in temperature and water demand can contribute to reducing water in reservoirs. The situation can pose a danger to society's food security, particularly since ISF, KHZ, and CHB are Iran's critical agricultural centers and affect a large portion of the country's population because agricultural activities mostly depend on the reservoir's water [92].

4.1. Comparison with Similar Studies

The proposed methodology used the Fmask initial water map to generate training samples for supervised classification. The Fmask initial water map could be derived by applying the Fmask algorithm to any Landsat scene. Thus, the proposed method was temporally transferable in providing training data, meaning that it could be used to generate training samples in different times. Thus, the proposed framework gave us the possibility of comparing our results with different reference global and Iranian maps. In this section, we compare our results with some reference maps of the same articles.

First, we compared our results in 2017 with the first public Iranian landcover map presented in [93]. The mentioned map was freely accessible in GEE as "Iran Land Cover Map v1 13-class (2017)", last accessed on March 2022. They proposed a threshold-based method that migrates ground truth samples from a reference year to any target year based on similarity measures. Visual comparisons are provided in Appendix A (Figure A1). It was revealed that the results of this article are more accurate, indicating the satisfactory performance of the proposed methodology in the training sample generation. Although they used optical and synthetic aperture radar (SAR) images with higher spatial resolution, their results showed considerable errors in water mapping.

Figure A2 in Appendix A visually compares our results with two global maps, World Cover and Joint Research Center (JRC), for 2020 in three K3, UG, and K1 study sites [13,94]. World Cover map is a publicly available landcover with a spatial resolution of 10 m. JRC is a 30-m long-term global water map [13]. Adapted RF trained with the samples generated by the proposed methodology could produce comparable results with two global reference maps, which were generated using different data sources and complex methods [76]. Quantitative results using the same test data (see Section 2.3.5) also indicated that adapted RF has achieved an OA of 95.13%, higher than JRC and World Cover maps by about 0.5% and 1.7%. The adapted RF's superior performance indicated the proposed method's usefulness for generating training samples.

4.2. Limitations, Uncertainties and Future Trends

This study investigated long-term inter-annual changes of water surface area in different dam lakes using satellite images for 31 years at a fixed time period in each year. Due to the dynamics of water surface area over a year, intra-annual changes can be considered to analyze overall trends and different factors more effectively [26]. In addition, we examined the long-term changes of eight dams in southern and southwestern parts of Iran. Considering more study sites in other parts of the world could provide an accurate view of dam reservoir status. It should be noted that changes were analyzed based on water surface area. Developing methods for calculating the reservoir's water volume not only helps scientists, managers, and policy-makers examine the dam's overall changes in more detail but also helps them study other dam-related issues such as sedimentation [95].

The proposed framework of this research was designed based on GEE capabilities. All results were also obtained in GEE cloud-based platform. Researchers in different applications can perform processes online in GEE without downloading satellite images, remote sensing products, or having powerful local processing systems. However, it had some limitations in computational time and memory capacity [33,51]. As a result, especially in large study areas or long-term studies, a limited number of input features and training samples can be used. Moreover, there were not various ready-to-use algorithms for users, and applying some methods such as Artificial Neural Networks and Deep Learning techniques requires at least offline training [41,93].

In this study, temperature and precipitation parameters were considered as climate factors to investigate their correlation with water surface area. Other climate factors such as evapotranspiration can be examined [76,96]. In addition, other parameters such as anthropogenic activities and industrial sector needs can help identify the most influential factors on water storage in dam reservoirs [26,51]. We used Landsat satellite imageries with a spatial resolution of 30 m. Therefore, they may cause errors in identifying narrow river channels [76]. Mixed pixels at land-water boundaries may affect classification results as well [29,32]. Thus, satellite imageries with higher spatial resolution, such as Sentinel-2, can be more effective. However, Sentinel-2 images cannot be used in long-term studies since it was launched in June 2015. Utilizing multi-source data (optical and SAR imageries) in water mapping can also be followed [97,98]. Moreover, we used traditional machine learning classifiers, RF and SVM. Future research can use ensemble classifiers and deep neural networks (i.e., convolutional networks), which have shown superior performance in different applications [78,99–101].

5. Conclusions

This study proposed a novel automated method of training sample generation for supervised monitoring of surface water extent changes using Landsat images. The framework of the study was developed and implemented based on GEE cloud-processing platform capabilities. An Iterative K-means clustering was deployed on an initial set of training samples derived from the Fmask water map to provide reliable training samples. These samples were used to train SVM and RF supervised classification models. Eight Iranian reservoirs, located in regions that face severe problems in supplying drinking and agricultural water, were selected to evaluate the performance of the proposed framework. Test samples were provided during extensive field surveys and visual interpretation of high-resolution satellite imageries. As test samples were prepared by a completely different procedure than training samples and were not involved in the training phase, they could carefully examine how the proposed methodology generated training samples. Both quantitative and qualitative results revealed that adapted classification models performed well in classifying water and non-water classes, indicating the proposed novel iterative clustering-based method's success in training sample generation. A comparison of RF and SVM classifiers showed better performance of the RF method in water mapping. Moreover, simultaneous use of TC, PCA, and SI features improved classification accuracy. However, TC-only features achieved higher accuracy than PCA-only and SI-only in water mapping.

Long-term change analysis of all study sites showed a downward trend of five dams (K1, K3, DEZ, ZR, GP), among which ZR and GP have experienced a more severe decrease. In addition, the water surface area of all dams in 2021, compared to the long-term maximum, showed significant reductions (39% (DEZ), 16% (UG), 62% (GP), 28% (K1), 31% (K3), 22% (K4), 44% (KA), and 58% (ZR)). We also analyzed the effect of two climate factors, precipitation, and temperature, on the water surface changes. The results showed that precipitation is more correlated with long-term changes in water surface area (R^2 between 0.51 (K4) to 0.79 (DEZ)). Based on the current state and overall trend of the study sites and the possibility of water scarcity in Iran due to increasing temperature, decreasing precipitation, and population growth, it can be said that all forms of consumption, from individual use to the supply chains of large companies and agricultural sectors, have to be reformed as soon as possible.

Software Availability: The GEE code for the proposed methodology will soon be available at https://github.com/ATDehkordi/Sustainability_ICRP (accessed date: 24 June 2022).

Abbreviations

The following abbreviations have been used in this article.

GEE	Google Earth Engine
Fmask	Function of the Mask
AVHRR	Advanced Very-High-Resolution Radiometer
MODIS	Moderate Resolution Imaging Spectroradiometer
VIIRS	Visible Infrared Imaging Radiometer Suite
TH	Threshold-based water mapping approaches
ML	Machine Learning-based water mapping approaches
EM	Electromagnetic
NDVI	Normalized Difference Vegetation Index
MNDWI	Modified Normalized Difference Water Index
EVI	Enhanced Vegetation Index
AWEI	Automated Water Extraction Index
NDWI	Normalized Difference Water Index
SVM	Support Vector Machines
RF	Random Forest
MAT	Mean Annual Temperature
K1, K3, K4	Karun-1, Karun-3, Karun-4
KA	Karkheh
ZR	Zayanderud
GP	Golpayegan

UG	UpperGotvand
NASA	National Aeronautics and Space Administration
MAP	Mean Annual Precipitation
KHZ	Khuzestan
ISF	Isfahan
CHB	Chaharmahal and Bakhtiari
SRTM	Shuttle Radar Topographic Mission
DEM	Digital Elevation Model
OLI	Operational Land Imager
ETM+	Enhanced Thematic Mapper Plus
TM	Thematic Mapper
LEDAPS	Landsat Ecosystem Disturbance Adaptive Processing System
LaSRC	Land Surface Reflectance Code
NIR	Near-infrared
SWIR1	Shortwave infrared 1
SWIR2	Shortwave infrared 1
FLDAS	Famine Early Warning Systems Network (FEWS NET) Land Data Assimilation System (FLDAS)
FEWS NET	Famine Early Warning Systems Network
MLSTY	MODIS Land Cover Type Yearly Global
SLC	Scan-Line Corrector
PCA	Principal Component Analysis
SIs	Spectral Indices
TC	Tasseled-Cap transformation
RBF	Radial Basis Function
OA	Overall Accuracy
UA	User Accuracy
PA	Producer Accuracy
TP	True Positive
TN	True Negative
FP	False Positive
FN	False Negative
JRC	Joint Research Center

Appendix A. Comparison with Similar Studies.

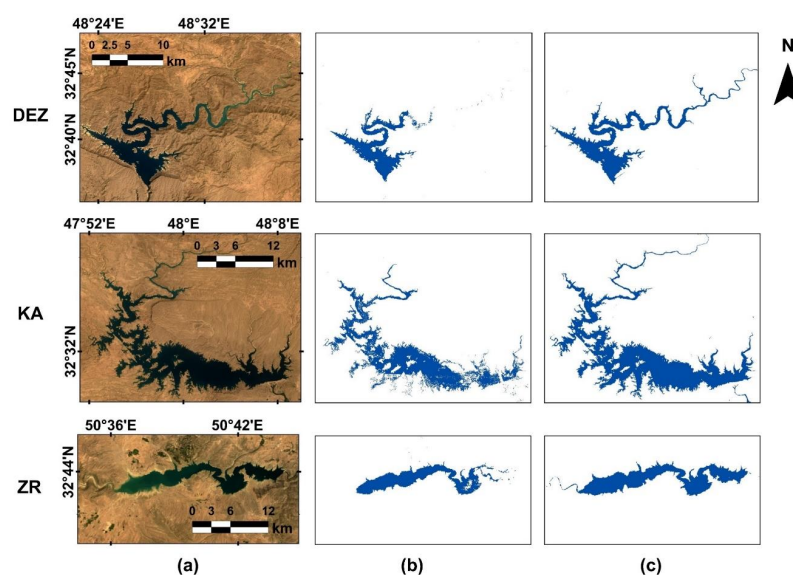**Figure A1.** (**a**) Landsat RGB composites in 2017, (**b**) water map of (source: K. N. Toosi University of Technology LiDAR Lab, entitled "Iran Land Cover Map v1 13-class (2017)" licensed under CC BY 4.0 [93]), and (**c**) our results derived from adapted RF.
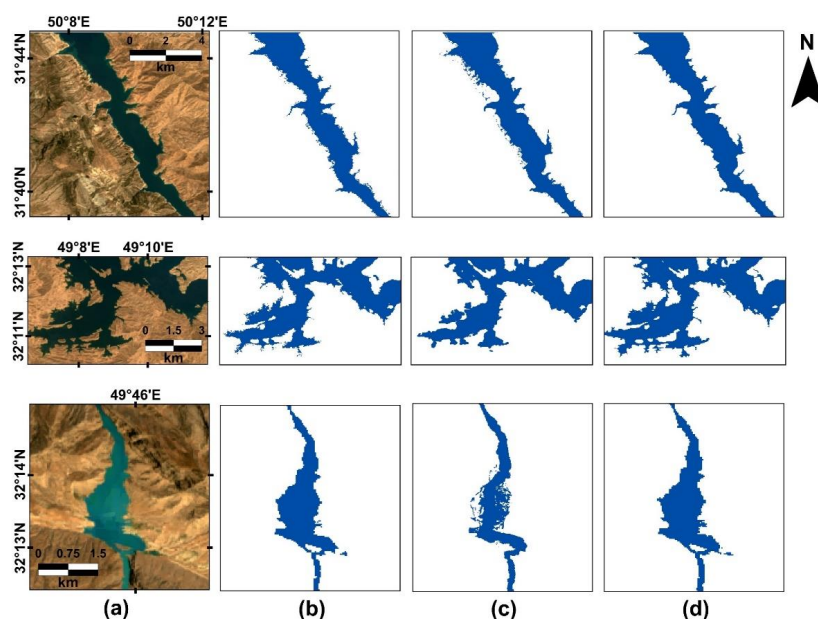
Figure A2. (**a**) Landsat RGB composites, (**b**) 30 m JRC map (source: EC JRC/Google that provides without restriction of use [14]), (**c**) 10 m WorldCover map (source: © ESA WorldCover project 2020/Contains modified Copernicus Sentinel data (2020) processed by ESA WorldCover consortium licensed under CC BY 4.0 [90]), and (**d**) the result of this article. OA of the adapted RF (trained with the samples from the proposed methodology) was 95.13% (see Figure 6). OA of JRC and World cover map was 94.87% and 93.49%, respectively (using the same test samples).

References

1. Niu, W.-j.; Feng, Z.-k. Evaluating the performances of several artificial intelligence methods in forecasting daily streamflow time series for sustainable water resources management. *Sustain. Cities Soc.* **2021**, *64*, 102562. [CrossRef]

2. Luo, P.; Sun, Y.; Wang, S.; Wang, S.; Lyu, J.; Zhou, M.; Nakagami, K.; Takara, K.; Nover, D. Historical assessment and future sustainability challenges of Egyptian water resources management. *J. Clean. Prod.* **2020**, *263*, 121154. [CrossRef]

3. Hassan, Q.K.; Ejiagha, I.R.; Ahmed, M.R.; Gupta, A.; Rangelova, E.; Dewan, A. Remote Sensing of Local Warming Trend in Alberta, Canada during 2001–2020, and Its Relationship with Large-Scale Atmospheric Circulations. *Remote Sens.* **2021**, *13*, 3441. [CrossRef]

4. Neisi, M.; Bijani, M.; Abbasi, E.; Mahmoudi, H.; Azadi, H. Analyzing farmers' drought risk management behavior: Evidence from Iran. *J. Hydrol.* **2020**, *590*, 125243. [CrossRef]

5. Mirzavand, M.; Bagheri, R. The water crisis in Iran: Development or destruction? *World Water Policy* **2020**, *6*, 89–97. [CrossRef]

6. Kornijów, R. Controversies around dam reservoirs: Benefits, costs and future. *Ecohydrol. Hydrobiol.* **2009**, *9*, 141–148. [CrossRef]

7. Sogno, P.; Klein, I.; Kuenzer, C. Remote Sensing of Surface Water Dynamics in the Context of Global Change—A Review. *Remote Sens.* **2022**, *14*, 2475. [CrossRef]

8. Hassan, Q.K.; Rahman, K.M. Applicability of remote sensing-based surface temperature regimes in determining deciduous phenology over boreal forest. *J. Plant Ecol.* **2013**, *6*, 84–91. [CrossRef]

9. Dietz, A.J.; Klein, I.; Gessner, U.; Frey, C.M.; Kuenzer, C.; Dech, S. Detection of water bodies from AVHRR data—A TIMELINE thematic processor. *Remote Sens.* **2017**, *9*, 57. [CrossRef]

10. Li, L.; Vrieling, A.; Skidmore, A.; Wang, T.; Turak, E. Monitoring the dynamics of surface water fraction from MODIS time series in a Mediterranean environment. *Int. J. Appl. Earth Obs. Geoinf.* **2018**, *66*, 135–145. [CrossRef]

11. Huang, C.; Chen, Y.; Wu, J.; Li, L.; Liu, R. An evaluation of Suomi NPP-VIIRS data for surface water detection. *Remote Sens. Lett.* **2015**, *6*, 155–164. [CrossRef]

12. Mueller, N.; Lewis, A.; Roberts, D.; Ring, S.; Melrose, R.; Sixsmith, J.; Lymburner, L.; McIntyre, A.; Tan, P.; Curnow, S. Water observations from space: Mapping surface water from 25 years of Landsat imagery across Australia. *Remote Sens. Environ.* **2016**, *174*, 341–352. [CrossRef]

13. Pekel, J.-F.; Cottam, A.; Gorelick, N.; Belward, A.S. High-resolution mapping of global surface water and its long-term changes. *Nature* **2016**, *540*, 418–422. [CrossRef]

14. Hardy, A.; Ettritch, G.; Cross, D.E.; Bunting, P.; Liywalii, F.; Sakala, J.; Silumesii, A.; Singini, D.; Smith, M.; Willis, T. Automatic detection of open and vegetated water bodies using Sentinel 1 to map African malaria vector mosquito breeding habitats. *Remote Sens.* **2019**, *11*, 593. [CrossRef]

15. Du, Y.; Zhang, Y.; Ling, F.; Wang, Q.; Li, W.; Li, X. Water bodies' mapping from Sentinel-2 imagery with modified normalized difference water index at 10-m spatial resolution produced by sharpening the SWIR band. *Remote Sens.* **2016**, *8*, 354. [CrossRef]

16. Bioresita, F.; Puissant, A.; Stumpf, A.; Malet, J.-P. A method for automatic and rapid mapping of water surfaces from sentinel-1 imagery. *Remote Sens.* **2018**, *10*, 217. [CrossRef]

17. Fu, J.; Wang, J.; Li, J. Study on the automatic extraction of water body from TM image using decision tree algorithm. In Proceedings of the International Symposium on Photoelectronic Detection and Imaging 2007. *Relat. Technol. Appl. SPIE* **2008**, *6625*, 17–25.

18. Sawaya, K.E.; Olmanson, L.G.; Heinert, N.J.; Brezonik, P.L.; Bauer, M.E. Extending satellite remote sensing to local scales: Land and water resource monitoring using high-resolution imagery. *Remote Sens. Environ.* **2003**, *88*, 144–156. [CrossRef]

19. Pandey, A.C.; Kaushik, K.; Parida, B.R. Google Earth Engine for Large-Scale Flood Mapping Using SAR Data and Impact Assessment on Agriculture and Population of Ganga-Brahmaputra Basin. *Sustainability* **2022**, *14*, 4210. [CrossRef]

20. Gorelick, N.; Hancher, M.; Dixon, M.; Ilyushchenko, S.; Thau, D.; Moore, R. Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sens. Environ.* **2017**, *202*, 18–27. [CrossRef]

21. Li, Q.; Qiu, C.; Ma, L.; Schmitt, M.; Zhu, X.X. Mapping the land cover of Africa at 10 m resolution from multi-source remote sensing data with Google Earth Engine. *Remote Sens.* **2020**, *12*, 602. [CrossRef]

22. Loukika, K.N.; Keesara, V.R.; Sridhar, V. Analysis of Land Use and Land Cover Using Machine Learning Algorithms on Google Earth Engine for Munneru River Basin, India. *Sustainability* **2021**, *13*, 13758. [CrossRef]

23. Gxokwe, S.; Dube, T.; Mazvimavi, D. Leveraging Google Earth Engine platform to characterize and map small seasonal wetlands in the semi-arid environments of South Africa. *Sci. Total Environ.* **2022**, *803*, 150139. [CrossRef] [PubMed]

24. Dehkordi, A.T.; Ghasemi, H.; Zoej, M.J.V. Machine Learning-Based Estimation of Suspended Sediment Concentration along Missouri River using Remote Sensing Imageries in Google Earth Engine. In Proceedings of the 2021 7th International Conference on Signal Processing and Intelligent Systems (ICSPIS), Tehran, Iran, 29–30 December 2021; pp. 1–5.

25. Venkatappa, M.; Sasaki, N.; Han, P.; Abe, I. Impacts of droughts and floods on croplands and crop production in Southeast Asia–An application of Google Earth Engine. *Sci. Total Environ.* **2021**, *795*, 148829. [CrossRef] [PubMed]

26. Zhou, Y.; Dong, J.; Xiao, X.; Liu, R.; Zou, Z.; Zhao, G.; Ge, Q. Continuous monitoring of lake dynamics on the Mongolian Plateau using all available Landsat imagery and Google Earth Engine. *Sci. Total Environ.* **2019**, *689*, 366–380. [CrossRef] [PubMed]

27. McFeeters, S.K. The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. *Int. J. Remote Sens.* **1996**, *17*, 1425–1432. [CrossRef]

28. Han-Qiu, X. A study on information extraction of water body with the modified normalized difference water index (MNDWI). *J. Remote Sens.* **2005**, *5*, 589–595.

29. Feyisa, G.L.; Meilby, H.; Fensholt, R.; Proud, S.R. Automated Water Extraction Index: A new technique for surface water mapping using Landsat imagery. *Remote Sens. Environ.* **2014**, *140*, 23–35. [CrossRef]

30. Yang, X.; Chen, Y.; Wang, J. Combined use of Sentinel-2 and Landsat 8 to monitor water surface area dynamics using Google Earth Engine. *Remote Sens. Lett.* **2020**, *11*, 687–696. [CrossRef]

31. Zou, Z.; Dong, J.; Menarguez, M.A.; Xiao, X.; Qin, Y.; Doughty, R.B.; Hooker, K.V.; Hambright, K.D. Continued decrease of open surface water body area in Oklahoma during 1984–2015. *Sci. Total Environ.* **2017**, *595*, 451–460. [CrossRef] [PubMed]

32. Wang, R.; Xia, H.; Qin, Y.; Niu, W.; Pan, L.; Li, R.; Zhao, X.; Bian, X.; Fu, P. Dynamic Monitoring of Surface Water Area during 1989–2019 in the Hetao Plain Using Landsat Data in Google Earth Engine. *Water* **2020**, *12*, 3010. [CrossRef]

33. Deng, Y.; Jiang, W.; Tang, Z.; Ling, Z.; Wu, Z. Long-term changes of open-surface water bodies in the Yangtze River basin based on the Google Earth Engine cloud platform. *Remote Sens.* **2019**, *11*, 2213. [CrossRef]

34. Chen, J.; Kang, T.; Yang, S.; Bu, J.; Cao, K.; Gao, Y. Open-Surface Water Bodies Dynamics Analysis in the Tarim River Basin (North-Western China), Based on Google Earth Engine Cloud Platform. *Water* **2020**, *12*, 2822. [CrossRef]

35. Verpoorter, C.; Kutser, T.; Tranvik, L. Automated mapping of water bodies using Landsat multispectral data. *Limnol. Oceanogr. Methods* **2012**, *10*, 1037–1050. [CrossRef]

36. Ji, L.; Zhang, L.; Wylie, B. Analysis of dynamic thresholds for the normalized difference water index. *Photogramm. Eng. Remote Sens.* **2009**, *75*, 1307–1317. [CrossRef]

37. Fisher, A.; Flood, N.; Danaher, T. Comparing Landsat water index methods for automated water classification in eastern Australia. *Remote Sens. Environ.* **2016**, *175*, 167–182. [CrossRef]

38. Sun, F.; Zhao, Y.; Gong, P.; Ma, R.; Dai, Y. Monitoring dynamic changes of global land cover types: Fluctuations of major lakes in China every 8 days during 2000–2010. *Chin. Sci. Bull.* **2014**, *59*, 171–189. [CrossRef]

39. Ko, B.C.; Kim, H.H.; Nam, J.Y. Classification of potential water bodies using Landsat 8 OLI and a combination of two boosted random forest classifiers. *Sensors* **2015**, *15*, 13763–13777. [CrossRef]

40. Isikdogan, F.; Bovik, A.C.; Passalacqua, P. Surface water mapping by deep learning. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2017**, *10*, 4909–4918. [CrossRef]

41. Wang, Y.; Li, Z.; Zeng, C.; Xia, G.-S.; Shen, H. An urban water extraction method combining deep learning and google earth engine. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2020**, *13*, 769–782. [CrossRef]

42. Friedl, M.A.; Sulla-Menashe, D.; Tan, B.; Schneider, A.; Ramankutty, N.; Sibley, A.; Huang, X. MODIS Collection 5 global land cover: Algorithm refinements and characterization of new datasets. *Remote Sens. Environ.* **2010**, *114*, 168–182. [CrossRef]

43. Li, K.; Xu, E. Cropland data fusion and correction using spatial analysis techniques and the Google Earth Engine. *GIScience Remote Sens.* **2020**, *57*, 1026–1045. [CrossRef]

44. Zhang, H.K.; Roy, D.P. Using the 500 m MODIS land cover product to derive a consistent continental scale 30 m Landsat land cover classification. *Remote Sens. Environ.* **2017**, *197*, 15–34. [CrossRef]

45. Zhu, Z.; Wang, S.; Woodcock, C.E. Improvement and expansion of the Fmask algorithm: Cloud, cloud shadow, and snow detection for Landsats 4–7, 8, and Sentinel 2 images. *Remote Sens. Environ.* **2015**, *159*, 269–277. [CrossRef]

46. Stillinger, T.; Roberts, D.A.; Collar, N.M.; Dozier, J. Cloud masking for Landsat 8 and MODIS Terra over snow-covered terrain: Error analysis and spectral similarity between snow and cloud. *Water Resour. Res.* **2019**, *55*, 6169–6184. [CrossRef] [PubMed]

47. The Statistical Centre of Iran. Iran Statistical Yearbook 1397 (2018–2019); 2019. Available online: https://www.amar.org.ir (accessed on 1 May 2022).

48. Farr, T. The shuttle radar topography mission. *Rev. Geophys.* **2007**, *45*, 1–13. [CrossRef]

49. Masek, J.G.; Vermote, E.F.; Saleous, N.E.; Wolfe, R.; Hall, F.G.; Huemmrich, K.F.; Gao, F.; Kutler, J.; Lim, T.-K. A Landsat surface reflectance dataset for North America, 1990–2000. *IEEE Geosci. Remote Sens. Lett.* **2006**, *3*, 68–72. [CrossRef]

50. Vermote, E.; Justice, C.; Claverie, M.; Franch, B. Preliminary analysis of the performance of the Landsat 8/OLI land surface reflectance product. *Remote Sens. Environ.* **2016**, *185*, 46–56. [CrossRef] [PubMed]

51. Wang, C.; Jia, M.; Chen, N.; Wang, W. Long-term surface water dynamics analysis based on Landsat imagery and the Google Earth Engine platform: A case study in the middle Yangtze River Basin. *Remote Sens.* **2018**, *10*, 1635. [CrossRef]

52. McNally, A. FLDAS noah land surface model L4 global monthly 0.1 × 0.1 degree (MERRA-2 and CHIRPS). *Atmos. Compos. Water Energy Cycles Clim. Var.* **2018**. Available online: https://developers.google.com/earth-engine/datasets/catalog/NASA_FLDAS_NOAH01_C_GL_M_V001 (accessed on 1 April 2022).

53. Jafari, M.; Hasanlou, M.; Arefi, H. SRTM DEM enhancement using a single set of PolSAR data based on the polarimetry-clinometry model. *Int. J. Remote Sens.* **2019**, *40*, 8979–9002. [CrossRef]

54. Sulla-Menashe, D.; Friedl, M.A. *User Guide to Collection 6 MODIS Land Cover (MCD12Q1 and MCD12C1) Product*; USGS: Reston, VA, USA, 2018; pp. 1–18.

55. Zeug, G.; Eckert, S. Population growth and its expression in spatial built-up patterns: The Sana'a, Yemen case study. *Remote Sens.* **2010**, *2*, 1014–1034. [CrossRef]

56. Faragó, S.; Hangya, K. Effects of water level on waterbird abundance and diversity along the middle section of the Danube River. *Hydrobiologia* **2012**, *697*, 15–21. [CrossRef]

57. Carrasco, L.; O'Neil, A.W.; Morton, R.D.; Rowland, C.S. Evaluating combinations of temporally aggregated Sentinel-1, Sentinel-2 and Landsat 8 for land cover mapping with Google Earth Engine. *Remote Sens.* **2019**, *11*, 288. [CrossRef]

58. Yang, J.; Li, W. Feature selection methods in extracting impervious surface from Landsat TM image. *Ann. GIS* **2013**, *19*, 253–261. [CrossRef]

59. Pettorelli, N. *The Normalized Difference Vegetation Index*; Oxford University Press: Oxford, UK, 2013.

60. Ghaderpour, E.; Vujadinovic, T. Change detection within remotely sensed satellite image time series via spectral analysis. *Remote Sens.* **2020**, *12*, 4001. [CrossRef]

61. Liu, H.Q.; Huete, A. A feedback based modification of the NDVI to minimize canopy background and atmospheric noise. *IEEE Trans. Geosci. Remote Sens.* **1995**, *33*, 457–465. [CrossRef]

62. Nguyen, U.N.; Pham, L.T.; Dang, T.D. An automatic water detection approach using Landsat 8 OLI and Google Earth Engine cloud computing to map lakes and reservoirs in New Zealand. *Environ. Monit. Assess.* **2019**, *191*, 1–12. [CrossRef]

63. Wang, Y.; Ma, J.; Xiao, X.; Wang, X.; Dai, S.; Zhao, B. Long-term dynamic of Poyang Lake surface water: A mapping work based on the google earth engine cloud platform. *Remote Sens.* **2019**, *11*, 313. [CrossRef]

64. Baig, M.H.A.; Zhang, L.; Shuai, T.; Tong, Q. Derivation of a tasselled cap transformation based on Landsat 8 at-satellite reflectance. *Remote Sens. Lett.* **2014**, *5*, 423–431. [CrossRef]

65. Huang, C.; Wylie, B.; Yang, L.; Homer, C.; Zylstra, G. Derivation of a tasselled cap transformation based on Landsat 7 at-satellite reflectance. *Int. J. Remote Sens.* **2002**, *23*, 1741–1748. [CrossRef]

66. Li, J.; Sheng, Y. An automated scheme for glacial lake dynamics mapping using Landsat imagery and digital elevation models: A case study in the Himalayas. *Int. J. Remote Sens.* **2012**, *33*, 5194–5213. [CrossRef]

67. Likas, A.; Vlassis, N.; Verbeek, J.J. The global k-means clustering algorithm. *Pattern Recognit.* **2003**, *36*, 451–461. [CrossRef]

68. Belgiu, M.; Drăguţ, L. Random forest in remote sensing: A review of applications and future directions. *ISPRS J. Photogramm. Remote Sens.* **2016**, *114*, 24–31. [CrossRef]

69. Mountrakis, G.; Im, J.; Ogole, C. Support vector machines in remote sensing: A review. *ISPRS J. Photogramm. Remote Sens.* **2011**, *66*, 247–259. [CrossRef]

70. Cánovas-García, F.; Alonso-Sarría, F.; Gomariz-Castillo, F.; Oñate-Valdivieso, F. Modification of the random forest algorithm to avoid statistical dependence problems when classifying remote sensing imagery. *Comput. Geosci.* **2017**, *103*, 1–11. [CrossRef]

71. Ghimire, B.; Rogan, J.; Galiano, V.R.; Panday, P.; Neeti, N. An evaluation of bagging, boosting, and random forests for land-cover classification in Cape Cod, Massachusetts, USA. *GIScience Remote Sens.* **2012**, *49*, 623–643. [CrossRef]

72. Tassi, A.; Vizzari, M. Object-oriented lulc classification in google earth engine combining snic, glcm, and machine learning algorithms. *Remote Sens.* **2020**, *12*, 3776. [CrossRef]

73. Lv, Z.; Li, G.; Jin, Z.; Benediktsson, J.A.; Foody, G.M. Iterative training sample expansion to increase and balance the accuracy of land classification from VHR imagery. *IEEE Trans. Geosci. Remote Sens.* **2020**, *59*, 139–150. [CrossRef]

74. Huang, X.; Weng, C.; Lu, Q.; Feng, T.; Zhang, L. Automatic labelling and selection of training samples for high-resolution remote sensing image classification over urban areas. *Remote Sens.* **2015**, *7*, 16024–16044. [CrossRef]

75. Jiang, W.; He, G.; Pang, Z.; Guo, H.; Long, T.; Ni, Y. Surface water map of China for 2015 (SWMC-2015) derived from Landsat 8 satellite imagery. *Remote Sens. Lett.* **2020**, *11*, 265–273. [CrossRef]

76. Xia, H.; Zhao, J.; Qin, Y.; Yang, J.; Cui, Y.; Song, H.; Ma, L.; Jin, N.; Meng, Q. Changes in water surface area during 1989–2017 in the Huai River Basin using Landsat data and Google earth engine. *Remote Sens.* **2019**, *11*, 1824. [CrossRef]

77. Strahler, A.H.; Boschetti, L.; Foody, G.M.; Friedl, M.A.; Hansen, M.C.; Herold, M.; Mayaux, P.; Morisette, J.T.; Stehman, S.V.; Woodcock, C.E. Global land cover validation: Recommendations for evaluation and accuracy assessment of global land cover maps. *Eur. Communities Luxemb.* **2006**, *51*, 1–60.

78. Taheri Dehkordi, A.; Valadan Zoej, M.J. Classification of croplands using sentinel-2 satellite images and a novel deep 3D convolutional neural network (case study: Shahrekord). *Iran. J. Soil Water Res.* **2021**, *52*, 1941–1953.

79. McNemar, Q. Note on the sampling error of the difference between correlated proportions or percentages. *Psychometrika* **1947**, *12*, 153–157. [CrossRef]

80. Peyravi, M.; Peyvandi, A.A.; Khodadadi, A.; Marzaleh, M.A. Flood in the South-West of Iran in 2019; causes, problems, actions and lesson learned. *Bull. Emerg. Trauma* **2019**, *7*, 199. [CrossRef]

81. Acharya, T.D.; Subedi, A.; Lee, D.H. Evaluation of machine learning algorithms for surface water extraction in a Landsat 8 scene of Nepal. *Sensors* **2019**, *19*, 2769. [CrossRef]

82. Wu, Y.; Duguay, C.R.; Xu, L. Assessment of machine learning classifiers for global lake ice cover mapping from MODIS TOA reflectance data. *Remote Sens. Environ.* **2021**, *253*, 112206. [CrossRef]

83. Balázs, B.; Bíró, T.; Dyke, G.; Singh, S.K.; Szabó, S. Extracting water-related features using reflectance data and principal component analysis of Landsat images. *Hydrol. Sci. J.* **2018**, *63*, 269–284. [CrossRef]

84. Thakkar, A.K.; Desai, V.R.; Patel, A.; Potdar, M.B. An effective hybrid classification approach using tasseled cap transformation (TCT) for improving classification of land use/land cover (LU/LC) in semi-arid region: A case study of Morva-Hadaf watershed, Gujarat, India. *Arab. J. Geosci.* **2016**, *9*, 180. [CrossRef]

85. Eslamian, S.; Jahadi, M. Monitoring and prediction of drought by Markov chain model based on SPI and new index in Isfahan. *Int. J. Hydrol. Sci. Technol.* **2019**, *9*, 355–365. [CrossRef]

86. Moradi, Y.; Said, M.; Bin, I. Drought impacts and vulnerability in Isfahan Province. *World Acad. Sci. Eng. Technol.* **2011**, *53*, 740.

87. Khosravi, I.; Jouybari-Moghaddam, Y.; Sarajian, M.R. The comparison of NN, SVR, LSSVR and ANFIS at modeling meteorological and remotely sensed drought indices over the eastern district of Isfahan, Iran. *Nat. Hazards* **2017**, *87*, 1507–1522. [CrossRef]

88. Dehcheshmeh, M.M.; Ghaedi, S. Climate change and ecological migration: A study of villages in the province of Khuzestan, Iran. *Environ. Res. Eng. Manag.* **2020**, *76*, 6–19. [CrossRef]

89. Hunt, J.D.; Falchetta, G.; Zakeri, B.; Nascimento, A.; Schneider, P.S.; Weber, N.A.B.; Mesquita, A.L.A.; Barbosa, P.S.F.; de Castro, N.J. Hydropower impact on the river flow of a humid regional climate. *Clim. Chang.* **2020**, *163*, 379–393. [CrossRef]

90. Jiang, Z.; Song, P.; Liao, X. Optimization of year-end water level of multi-year regulating reservoir in cascade hydropower system considering the inflow frequency difference. *Energies* **2020**, *13*, 5345. [CrossRef]

91. Liu, Y.; Jiang, Z.; Feng, Z.; Chen, Y.; Zhang, H.; Chen, P. Optimization of energy storage operation chart of cascade reservoirs with multi-year regulating reservoir. *Energies* **2019**, *12*, 3814. [CrossRef]

92. Rocha, J.; Carvalho-Santos, C.; Diogo, P.; Beça, P.; Keizer, J.J.; Nunes, J.P. Impacts of climate change on reservoir water availability, quality and irrigation needs in a water scarce Mediterranean region (southern Portugal). *Sci. Total Environ.* **2020**, *736*, 139477. [CrossRef]

93. Ghorbanian, A.; Kakooei, M.; Amani, M.; Mahdavi, S.; Mohammadzadeh, A.; Hasanlou, M. Improved land cover map of Iran using Sentinel imagery within Google Earth Engine and a novel automatic workflow for land cover classification using migrated training samples. *ISPRS J. Photogramm. Remote Sens.* **2020**, *167*, 276–288. [CrossRef]

94. Van De Kerchove, R.; Zanaga, D.; De Keersmaecker, W.; Souverijns, N.; Wevers, J.; Brockmann, C.; Grosu, A.; Paccini, A.; Cartus, O.; Santoro, M. ESA WorldCover: Global land cover mapping at 10 m resolution for 2020 based on Sentinel-1 and 2 data. In Proceedings of the AGU Fall Meeting 2021, New Orleans, LA, USA, 13–17 December 2021.

95. Schleiss, A.J.; Franca, M.J.; Juez, C.; De Cesare, G. Reservoir sedimentation. *J. Hydraul. Res.* **2016**, *54*, 595–614. [CrossRef]

96. Ghaderpour, E.; Vujadinovic, T.; Hassan, Q.K. Application of the least-squares wavelet software in hydrology: Athabasca River basin. *J. Hydrol. Reg. Stud.* **2021**, *36*, 100847. [CrossRef]

97. Wang, J.; Ding, J.; Li, G.; Liang, J.; Yu, D.; Aishan, T.; Zhang, F.; Yang, J.; Abulimiti, A.; Liu, J. Dynamic detection of water surface area of Ebinur Lake using multi-source satellite data (Landsat and Sentinel-1A) and its responses to changing environment. *Catena* **2019**, *177*, 189–201. [CrossRef]

98. Jafari, M.; Maghsoudi, Y.; Zoej, M.J.V. Analyzing polarimetric signatures for different features in polarimetric SAR data. In Proceedings of the 2014 IEEE Geoscience and Remote Sensing Symposium, Quebec City, QC, Canada, 13–18 July 2014; pp. 2782–2785.

99. Dehkordi, A.T.; Beirami, B.A.; Zoej, M.J.V.; Mokhtarzade, M. Performance Evaluation of Temporal and Spatial-Temporal Convolutional Neural Networks for Land-Cover Classification (A Case Study in Shahrekord, Iran). In Proceedings of the 2021 5th International Conference on Pattern Recognition and Image Analysis (IPRIA), Kashan, Iran, 28–29 April 2021; pp. 1–5.

100. Esfahani, M.M.; Sadati, H. fNIRS Signals Classification with Ensemble Learning and Adaptive Neuro-Fuzzy Inference System. In Proceedings of the 2021 7th International Conference on Signal Processing and Intelligent Systems (ICSPIS), Tehran, Iran, 29–30 December 2021; pp. 1–5.

101. Rahmati, A.; Zoej, M.J.V.; Dehkordi, A.T. Early identification of crop types using Sentinel-2 satellite images and an incremental multi-feature ensemble method (Case study: Shahriar, Iran). *Adv. Space Res.* **2022**. Available online: https://www.sciencedirect.com/science/article/abs/pii/S0273117722004173 (accessed on 25 May 2022). [CrossRef]