Monitoring Large-Scale Inland Water Dynamics by Fusing Sentinel-1 SAR and Sentinel-3 Altimetry Data and by Analyzing Causal Effects of Snowmelt

Ya-Lun S. Tsai 1,*, Igor Klein 1, Andreas Dietz 1 and Natascha Oppelt 2

1 German Remote Sensing Data Center (DFD), German Aerospace Center (DLR), Muenchener Strasse 20, 82234 Wessling, Germany; Igor.Klein@dlr.de (I.K.); Andreas.Dietz@dlr.de (A.D.)
2 Department of Geography, Earth Observation and Modelling, Kiel University, Ludewig-Meyn-Str. 14, 24118 Kiel, Germany; oppelt@geographie.uni-kiel.de

* Correspondence: Ya-Lun.Tsai@dlr.de

Received: 9 November 2020; Accepted: 25 November 2020; Published: 27 November 2020

Abstract: The warming climate is threatening to alter inland water resources on a global scale. Within all waterbody types, lake and river systems are vital not only for natural ecosystems but, also, for human society. Snowmelt phenology is also altered by global warming, and snowmelt is the primary water supply source for many river and lake systems around the globe. Hence, (1) monitoring snowmelt conditions, (2) tracking the dynamics of snowmelt-influenced river and lake systems, and (3) quantifying the causal effect of snowmelt conditions on these waterbodies are critical to understand the cryo-hydrosphere interactions under climate change. Previous studies utilized in-situ or multispectral sensors to track either the surface areas or water levels of waterbodies, which are constrained to small-scale regions and limited by cloud cover, respectively. On the contrary, in the present study, we employed the latest Sentinel-1 synthetic aperture radar (SAR) and Sentinel-3 altimetry data to grant a high-resolution, cloud-free, and illumination-independent comprehensive inland water dynamics monitoring strategy. Moreover, in contrast to previous studies utilizing in-house algorithms, we employed freely available cloud-based services to ensure a broad applicability with high efficiency. Based on altimetry and SAR data, the water level and the water-covered extent (WCE) (surface area of lakes and the flooded area of rivers) can be successfully measured. Furthermore, by fusing the water level and surface area information, for Lake Urmia, we can estimate the hypsometry and derive the water volume change. Additionally, for the Brahmaputra River, the variations of both the water level and the flooded area can be tracked. Last, but not least, together with the wet snow cover extent (WSCE) mapped with SAR imagery, we can analyze the influence of snowmelt conditions on water resource variations. The distributed lag model (DLM) initially developed in the econometrics discipline was employed, and the lagged causal effect of snowmelt conditions on inland water resources was eventually assessed.

Keywords: snow cover; water level; surface area; flooded area; hypsometry; distributed lag model; water resource

1. Introduction

Inland water resources play an essential role in not only the prosperity and stability of human society but, also, the sustainability and balance of various ecosystems. According to the reports of the United Nations Environment Programme (UNEP) [1], UN Food and Agriculture Organization (FAO) [2], and the latest UN World Water Development Report (WWDR) 2019 [3], transboundary surface water resources largely influence the socioeconomic-ecological systems; therefore, it is enlisted as the sixth 2030 Sustainable Development Goal (SDG). Expressly, the importance of local rivers and lakes should...
be noted, as they are critical freshwater sources for many regions [3] and, also, influence the climate [4]. However, as studies suggested [5,6], these vague SDGs provide no dependable guideline on how to achieve them. As a result, to elaborate the practical research gap of SDG target 6.6, i.e., protect and restore water-related ecosystems, an investigation of the main drivers of change in the high-mountain cryosphere is recommended [6]. The importance of snowmelt water to downstream freshwater is widely recognized. Assessing this cryo-hydrosphere interaction is indispensable to understand the future trend of water resources we have in severer climate change scenarios. Due to global warming, the shrinking of snow cover extent (SCE), earlier snowmelt season, and shorter snow cover duration have been observed almost globally, as stated in the Synthesis Report of the Intergovernmental Panel on Climate Change (IPCC) Fifth Assessment Report (AR5) [7]. Moreover, in the same report, it was also indicated that the change of melting snow is altering hydrological systems and affecting water resources in both quantity and quality. Thus, monitoring snowmelt conditions and the dynamics of snowmelt-influenced river and lake systems and quantifying the causal effect of snowmelt water on these waterbodies are critical. It would allow us not only to understand the impact of global warming to inland water resources but, also, to assess the vulnerability and variability of regional freshwater supplies.

For the observation and investigation of the influence of snowmelt on global waterbodies, the employment of spaceborne remote-sensing data is preferred, owing to its high efficiency and broad applicability. As mentioned in the latest report published by UN-Water assessing the progress of SDG indicator 6.6.1 [8], i.e., tracking time series changes in the extent (including quality, quantity, and area) of water-related ecosystems (including rivers and lakes), the utilization of globally transferable satellite data is strongly advocated. For monitoring the snowmelt conditions, as summarized in Tsai et al. [9], it is hardly possible to discriminate wet and dry snow with conventional multi-spectral data sources solely based on the reflectance differences, as the spectral characteristics of both snow types share high similarity. SAR data, on the other hand, are excellent for wet SCE (WSCE) mapping: The snowmelt-caused wetness in the snowpack would change the dielectric constant of the snowpack and, thus, dramatically shorten the penetration depth of the SAR signal, which eventually lead to a significant decrease of the backscatter coefficient when the snowpack starts to melt [9–11].

For monitoring the dynamics of river and lake systems, conventional studies focus on either the water level or surface area. For tracking the surface area, multi-spectral sensor-based band ratio water indices, such as the Normalized Difference Water Index (NDWI) [12], Modified NDWI (MNDWI) [13], and Automated Water Extraction Index (AWEI) [14], are commonly used due to their simplicitics. However, in addition to the difficulty of the selection among these various band ratios, these indices are based on a multi-spectral sensor. Therefore, they are inevitably affected by cloud cover and polar darkness, which frequently happen for alpine regions and high-latitude zones [9]. Moreover, the spectral feature of water in the multi-spectral sensor is also affected by the variation of atmospheric/illumination conditions and water dynamics, such as water depth, sediment load, eutrophication degree, turbidity, sun angle, and sensor view angle [15,16]. On the contrary, SAR data can be utilized for waterbody monitoring thanks to its cloud-penetrating applicability and illumination-independent characteristics. Yet, it might suffer from relatively coarse resolution compared to optical sensors, the higher cost of the imagery, and most importantly, the requirement of computation-intensive processing.

Regarding the monitoring of the water level of lakes or rivers, using in-situ gauging station measurements is the optimal approach. However, although it provides the chance of continuous observations, it is not widely achievable in all lake and river systems because of the cost of instrument setup, intensive maintenance requirements, or the difficult accessibility of remote waterbodies. In fact, the number of available gauging stations has been decreasing globally [17,18]. Additionally, the data might not be publicly available. Another alternative is combining the surface area of a waterbody with the bathymetry of the lake or river channel to derive the water surface height. Nevertheless, accurate information of the bathymetry is still needed, which is only available for very few waterbodies, which were surveyed thoroughly [4,18]. On the contrary, spaceborne altimetry provides a favorable
alternative for waterbody surface height monitoring due to its global availability. Many altimetry databases have therefore been established, including the Database for Hydrological Time Series of Inland Waters (DAHITI) (https://dahiti.dgfi.tum.de/en/) [19], The Global Reservoir and Lake Global Reservoir and Lake Monitor (G-REALM) (https://ipad.fas.usda.gov/cropexplorer/global_reservoir/) [20], and Hydroweb (http://hydroweb.theia-land.fr/) [21]. Unfortunately, because the design of spaceborne altimetry sensors is generally targeted for ocean surface surveying, i.e., a flat, isotropic, and ample reflecting surface, the footprint sizes of their echoes are enormously coarse (>2 km), and the sampling distance between these echo points is vast [22]. These characteristics hinder the viability of inland waterbody monitoring, especially for rivers, which usually have a more inhomogeneous neighboring topography [23]. Thus, to overcome the waveform difference to the standard Brown model [24], a computation-intensive waveform reprocessing (retracking) analysis is usually required for inland waterbodies [4,25–27]. However, the selection of retrackers (ground processing techniques that estimate the range to the point of closest approach on the surface) varies in different studies, and most of them use in-house algorithms or refinements [4,22,28,29], which largely limits the transferability of the method to other regions.

With the aim to overcome these limitations for monitoring the surface area and water level, in the present study, we utilized the new generation of high-resolution, freely accessible spaceborne SAR and altimetry data of Sentinel-1 and Sentinel-3, together with freely available online cloud-based processing services, to achieve a comprehensive inland waterbody dynamics (time series surface area and water level) tracking strategy with broad applicability. Although there have been some studies employing either spaceborne SAR or altimetry sensors for waterbody monitoring, few studies have yet integrated them to achieve an in-situ measurement-free approach. Additionally, in contrast to studies relying on multi-spectral sensors, this altimetry-SAR sensor combination provides a cloud-free and illumination-independent monitoring method. Furthermore, as the Sentinel fleet is the newest generation of spaceborne satellites with a freely accessible data policy, data availability is ensured for the following decades. Last, but not least, based on the depicted comprehensive waterbody dynamics, together with the SAR-based WSCE maps, we can further analyze the relationship between snowmelt conditions and inland water resource variations. By employing a sophisticated regression model adopted from the econometrics domain, we guarantee a quantified and unbiased lagged influence of the snowmelt condition on the waterbody dynamics; hence, the cryo-hydrosphere interaction can be well-identified.

2. Study Areas and Data

2.1. Study Areas

Since the goal of the present study is to (1) monitor the dynamics (both water level and surface area) of inland waterbodies and (2) quantify the lagged influence of snowmelt on inland water dynamics, we selected the study sites based on the following criteria: (1) their spatial scales were wide enough to be detected by spaceborne SAR and altimetry sensors, (2) they demonstrated a considerable magnitude of variations in both water level and surface area, and (3) their watersheds included snow-covered regions. Therefore, in our study, Lake Urmia in Iran and the Brahmaputra River were chosen.

2.1.1. Lake Urmia

Lake Urmia (or Orumiyeh Lake) is located in the northwest of Iran. It is the largest lake in Iran and is the largest perennial hyper-saline lake worldwide [30]. This endorheic lake is situated 1273 m above the mean sea level and is situated in a semi-arid climate zone with average annual precipitation and evaporation of around 341 and 1200 mm/y, respectively [31]. The highest river inflows can be observed in the springtime due to snowmelt [32]. Based on its geological location and climatic characteristics, Lake Urmia and the basin encompass a high biodiversity and are one of the most critical ecological habitats in the Middle East. For instance, Lake Urmia is the largest habitat of brine shrimp Artemia
which leads to frequent morphological changes and affects the natural environment and habitat but, also, the nearby population of seven million inhabitants in both economic, as well as health, aspects [46]. For instance, sediments of the Brahmaputra riverbed are composed of medium-to-fine sand and silt that is uniformly distributed. River braiding, division, and shifting often happen [54]. However, although it is one of the most critical water resources in Asia, there is barely any in-situ river measurement available publicly. Due to the imminent danger of flooding, such in-situ data is regarded as classified information, and even the Global Runoff Data Center (GRDC) keeps no recent observations [55,56]. Together with the fact that Lake Urmia is a terminal lake with a maximum depth of only 16 m, it is even more vulnerable to evaporation [44]. Thus, the salinity has reached a dangerous level for species [42], and the water equilibrium has fallen to a new low standard [43]. According to Abbaspour et al. [45], the lake might eventually dry up within a decade. The lake desiccation endangers not only the natural environment and habitat but, also, the nearby population of seven million inhabitants in both economic, as well as health, aspects [46]. For instance, the depleted lake bed would reveal dissolved salts crusts [47], which would be exposed to the wind and cause salt storms to the surrounding residential areas, the shrinking size of the lake would reduce its function as a mediator in the extreme climate [37], and the dehydrated bed would also lead to species migration [36]. Moreover, the shortage of water resources might also result in a political crisis [48].

![Figure 1](image_url)

**Figure 1.** The locations of the study areas with the Sentinel-3 altimetry satellite pass ground tracks (dotted red lines) and the corresponding virtual stations (VSS) (green points). The investigated water-covered extent (WCE) (black polygon) and the average water-covered frequency of 2003–2018 derived from the Global WaterPack (GWP) are illustrated for (a) Lake Urmia and (b) Brahmaputra River.

2.1.2. Brahmaputra River

The Brahmaputra River originates in the Himalayan mountain range, which is mainly fed by snowmelt water and is the third-largest river in the world by discharge [49–51]. It is a trans-boundary river; by order, it passes by China, India, Bangladesh, and eventually merges with the Ganges river and flows into the Bay of Bengal. It is called by different names in each country, including Yarlung Zangbo, Brahmaputra, and Jamuna in China, India, and Bangladesh, respectively. As the youngest major river among the world [50], the Brahmaputra River has at least three characteristics: abundant tributaries, highly dynamic fluvial activities, and a varying river width, as shown in Figure 1b. Since the sediments of the Brahmaputra riverbed are composed of medium-to-fine sand and silt that is uniformly graded and has poor transport resistance [50], the riverbed and banks area are considerably mobile, which leads to frequent morphological changes affected by fluvial processes [50,52,53]. Therefore, river braiding, division, and shifting often happen [54]. However, although it is one of the most critical
water resources in Asia, there is barely any in-situ river measurement available publicly. Due to the imminent danger of flooding, such in-situ data is regarded as classified information, and even the Global Runoff Data Center (GRDC) keeps no recent observations [55,56].

2.2. Data

2.2.1. Sentinel-1 SAR

The Sentinel-1 mission consists of two-satellites, with Sentinel-1A/B launched in April 2014 and April 2016, respectively. Its advanced terrain observation with progressive scans in the azimuth (TOPS) image acquisition technique enables a high signal-to-noise (SNR) ratio in the long-track direction without a scalloping effect while maintaining a wide coverage like conventional ScanSAR mode [57]. As in the present study, we only utilized the intensity information, the Level-1 Ground Range Detected (GRD) product, for which all the sub-swath have already been merged and de-bursting was selected. To increase the efficiency of the image analysis, we processed the images on Google Earth Engine (GEE). It provides the archive of Sentinel-1 GRD scenes and stores them in decibel (dB) units. Further calibrations, such as the orbit file application, noise removal, radiometric calibration, and terrain corrections, are processed as well. Furthermore, this cloud-based geospatial processing platform enables user-defined processing, which largely accelerates the time series SAR data analysis in a wide spatial scale [58].

2.2.2. Sentinel-3 Altimetry

Similar to Sentinel-1, Sentinel-3 is also composed of two satellites, with Sentinel-3A/B launched in February 2016 and April 2018, respectively. When it comes to inland water monitoring, compared to previous spaceborne altimetry sensors, the biggest advantage provided by the dual-frequency SAR radar altimeter (SRAL) instrument aboard Sentinel-3 is the much finer along-track resolution. Inherited from CryoSat, Sentinel-3 uses the high pulse repetition frequency (PRF) SAR mode (or delay Doppler) processing, which greatly improves the along-track resolution of the commonly used low-resolution mode (LRM) altimeters. It significantly increases the SNR ratio and, thus, enables the detection of much smaller targets [59]. Moreover, the echo reception window positioning method is also upgraded. Thanks to these improvements, Sentinel-3 is capable of monitoring finer inland waterbodies. To process the altimetry data via the European Space Agency (ESA) online service, Level-1 Non-Time Critical (NTC) products are selected.

2.2.3. Auxiliary Data

To include hydrological factors, including rainfall and evapotranspiration, the latest European Centre for Medium-Range Weather Forecasts (ECMWF) ERA5-Land data are used. This newly released reanalysis dataset provides global meteorological variables in a much higher spatial resolution (nine kilometers) compared to previous ERA-Interim (79 km) and ERA5 (31 km) datasets. Currently, it covers data from 1981 to the present, and eventually, it would cover the same period as ERA, i.e., 1950 to near real time.

For the examination of our water level results, the data provided by DAHITI, G-REALM, and Hydroweb are used. To compare our water-covered extent estimations, the Global WaterPack (GWP) was selected. It utilizes daily Moderate Resolution Imaging Spectroradiometer (MODIS) reflectance data at a 250-m spatial resolution and dynamically decides the threshold values, as well as the temporal interpolation techniques, to achieve the daily global cloud-free water/no water classification [15].
3. Methodology

3.1. Water Level Calculation with Sentinel-3 Altimetry

To derive the water level from Sentinel-3 altimetry data, we utilize the ESA’s SAR Versatile Altimetric Toolkit for Ocean Research & Exploitation (SARvatore) service. It is based on the G-POD (Grid Processing on Demand) distributed platform, which guarantees high-speed processing and timely delivery. Another highlight of the SARvatore service is that it provides the inland water-customized processing configuration, which utilizes the higher posting rate (HPR) of 80-Hz data and processes with the new SAMOSA+ analytical retracker algorithm [60,61]. By utilizing this advanced setting mode, we can reduce the separation between two observations in along-track from the original 300 m to 80 m and, thus, provide more valid points over the targeted waterbody. The output files include not only L2 data in NetCDF format but a KML file containing the satellite pass ground track location. The KML file facilitates the identification of the interested regions, as shown in Figure 1, i.e., virtual station (VS) (the intersection of the altimetry’s ground track and the waterbody), in the following analysis of the Multi-Mission Radar Altimetry Toolbox (BRAT). It must be noted that, because both of our study targets are characterized by significant seasonal WCE change, we first utilized the GWP to identify the subregion of a waterbody having (near) permanent water and, then, analyzed the altimetry data of that region to minimize the signal pollution. This step largely enhanced the accuracy by avoiding taking account river/lakebed-caused signals in the dry season. To derive the unbiased orthometric height over the inland waterbody $H_{ORTHO}$, i.e., the water level referred to the geoid, the following correction equation was used:

$$H_{ORTHO} = H_{SAT} - R_{OBS} - (\Delta R_{WTC} + \Delta R_{DTC} + \Delta R_{IONO} + \Delta R_{ET} + \Delta R_{PT}) - N_{GEOID}$$  (1)

where $H_{SAT}$ is the satellite altitude above the reference geoid, $R_{OBS}$ is the observed range, $\Delta R_{DTC}/\Delta R_{WTC}$ is the dry/wet tropospheric correction, $\Delta R_{IONO}$ is the ionospheric correction, $\Delta R_{ET}/\Delta R_{PT}$ is the earth tide and pole tide, respectively, and $N_{GEOID}$ is the geoid height. For a full explanation of each correction, refer to [59]. Based on this formula, both propagation and geophysical biases can be compensated. Other geophysical terms, including lake tides, hydrostatic variations, thermal expansion, and wind piling-up effects are neglected, as suggested by a previous study [25]. For detailed information of each correction, refer to [59,62]. Finally, based on the corrected range, we calculated the average and standard deviation of the water level over the VS.

3.2. Water-Covered Extent (WCE) Calculation with Sentinel-1 SAR

To map the water-covered extent, i.e., the lake’s surface area or river’s channels and flooded area, the sensitivity of the SAR data to the surface wetness was utilized. As the backscatter coefficient would be much lower in wet and smooth surface compared to dry ground, it is viable to depict the water surfaces. Since our goal is to maximize the differentiation of water/nonwater regions, i.e., a binary classification, the task can be separated into two parts: (1) selection of the targeted image and (2) determination of the threshold. For the first point, based on the previous studies [63,64] and our testing, the value difference between waterbody and the land region is more significant in VV polarization (co-polarization) than VH (cross-polarization), so VV imagery was selected. Furthermore, we used Otsu’s algorithm to decide the threshold [65]. It is favored for its simplicity and suitability for SAR image-based binary classification [63,66]. Practically, the critical point is to ensure that the manually selected samples include an equal number of water/nonwater pixels, as Otsu’s algorithm can only handle bimodal distribution. However, sometimes, it is difficult to distinguish lake/river boundaries visually in some flood-caused vague scenes (as discussed in Discussions 5.1); we only used scenes having clear waterbody boundaries for training the Otsu’s algorithm.

To define the targeted WCE for investigation, for Lake Urmia, we used the waterbody boundary provided by the Global Lakes and Wetlands Database (GLWD) (https://www.worldwildlife.org/pages/
global-lakes-and-wetlands-database) [67]. For the Brahmaputra River, we targeted the part of the watershed of the VS where the terrain was a flat plain (elevation lower than 200 m) and frequent river braiding and shifting events happened, as illustrated in Figure 1.

### 3.3. Lake Water Volume Estimation by Hypsometry Calculation and Detrended Volume Retrieval

To estimate the water volume variation of a waterbody, it is common to use either the simplest truncated pyramid model or the power-function model; however, they are only suitable for waterbodies characterized by a regular morphology (such as reservoirs) or a bowl-shaped morphology, respectively [36,68]. We utilized a universally applicable approach by modeling the water volume-water level relationship from the observed surface area-water level relationship [18]. The processing steps included: (1) calculation of the water level difference (or water depth) relative to the minimum water level in the sensing period ($\Delta L = L - L_{\text{min}}$); (2) establishing a scatter plot of surface area-water level difference and modeling it with a polynomial function (typically, the second, third, or quadratic order is chosen) ($A = f(\Delta L)$); (3) integration of the surface area-water level difference function to derive the water volume-water level difference function ($V = \int_{L_{\text{min}}}^{L} A d\Delta L$); and (4) the ingestion of water level observations to the resultant function to derive the corresponding water volume. Details about these steps can be found in [18,69]. In short, by using this approach, we can estimate the “dynamic” water volume above the “static” water volume, i.e., the water volume of the time when the water level is the lowest for any waterbody having a horizontal water surface [70].

Nevertheless, it must be noted that, compared to any single point of the river, the lake is a “container” that can store the water over time instead of merely letting the water flow by. Namely, the results of each previous hydrological year’s water budget balance (increase with snowmelt and rainfall and decrease with evapotranspiration) would affect the current “stock” of the water volume. Hence, if we directly use the original time series water volume as the interannual variation, the intra-annual “container effect” could lead to a misinterpretation. Therefore, in the present study, we employed the Seasonal-trend decomposition procedures based on Local Regression (Loess) (STL) [71] to remove the intra-annual trend. This approach is suggested by previous studies [72–74]. STL is an iterative nonparametric filtering procedure that uses repeated Loess (Local Regression) smoothing. The strength of STL is its robustness and computational efficiency, as well as the capacity to depict a nonlinear pattern in time series data. For details about STL, refer to [71,75]. Practically based on the inner and outer loops with Loess processing, STL can decompose time series data into three components: a low-frequency long-term trend, a high-frequency seasonal variation, and residuals (or remainder). After processing STL, we subtracted the decomposed long-term trend from the original time series water volume to derive the detrended water volume, which represents the real inter-annual water volume variation.

### 3.4. WSCE Mapping with Sentinel-1 SAR and Hydrological Factors Areal Calculation

Based on the finding that the backscatter coefficient of the SAR signal decreases significantly when the snowpack starts to melt and, thus, increases the containing liquid water [9], Nagler and Rott [10] proposed a ratio-thresholding approach in 2000. It utilizes two SAR scenes (one is a wet snow-covered period image, and the other is a referenced snow-free image) and calculates their ratio image, which is then thresholded using a fixed value to derive the binary WSCE. For a detailed description of the processing steps and value setting, refer to [9,76,77].

Since the final regression analysis is watershed-based, all hydrological factors, including WSCE, rainfall, and evapotranspiration, need to be converted. Firstly, we delineated the watershed based on the Shuttle Radar Topography Mission Digital Elevation Model (SRTM DEM). For the Brahmaputra River, the VS where the Sentinel-3 track passes was used as the pour point, and the watershed of that point of that tributary was delineated. Based on the watershed, the WSCE was ratioed to derive the WSCE%; the rainfall and evapotranspiration were summed up to derive the mean amount.
3.5. Hydrological Econometrics Regression Analysis

To quantify the influence of the investigated hydrological factors (WSCE%, mean rainfall, and mean evapotranspiration) on the river/lake water amount, the utilization of a regression analysis is an optimized approach. However, some factors, especially the snowmelt, would pose a lagged effect on the downstream water level/volume as the snowmelt water would not only form a direct surface streamflow but, also, infiltrate into the ground and soil [78–80]. Thus, a conventional simple regression is not viable. Consequently, we employed the advanced distributed lag model (DLM) (or finite distributed model (FDL)) in the present study for handling multiple variables’ dynamic influences with lags. DLM was firstly developed in the econometrics discipline [81] and was recently widely applied in biomedical or environmental exposure-caused mortality studies [82,83]. Yet, DLM is barely utilized in the remote sensing and cryo-hydrology domains. DLM is a dynamic model that assumes that the effect of each independent factor on the dependent variable spreads over some time instead of a single time point. To quantify it, it defines the temporal factor by assigning a lag dimension, which is an equally spaced and ordered time point series [84] (in the present study, we defined the minimum unit of the lag dimension as ten days). The lag dimension enables the effect of a single exposure event to be distributed over a specific period of time; therefore, the holistic understanding of the cross-temporal exposure-lag-response relationship can be revealed [84–86]. The multivariable DLM can be written in the formula:

$$y_t = \alpha + \sum_{s=0}^{q_1} \beta_1 x_{1,t-s} + \sum_{s=0}^{q_2} \beta_2 x_{2,t-s} + \cdots + \sum_{s=0}^{q_N} \beta_N x_{N,t-s} + u_t$$

(2)

where \(\alpha\) is the intercept, \(q\) and \(\beta\) are the lag length and lag coefficient (or lag weight or short-run multiplier) of each independent variable, respectively, and \(u_t\) is the error term. The lag coefficient is solved by the ordinary least squares (OLS) technique and represents the expected change in \(y_t\) stems from the change of \(x_{N,t-s}\) by one unit, holding constant the other independent variables [87]. The sum of each independent variable’s lag coefficient, i.e., \(\sum_{s=0}^{q_N} \beta_N\), is called the long-run multiplier or long-run propensity (LRP), which is the cumulative effect of \(xN\) on \(y\) [88]. In the present study, we focused on the LRP value to quantify the causal effect of each factor on the water resource variations.

Practically, the most critical issue when implementing DLM is selecting a suitable lag length for each independent variable. As in the present study, we included three variables (WSCE%, mean rainfall, and mean evapotranspiration), and based on the background knowledge, it was assumed that their lag lengths should be different; therefore, it is impractical to use the conventional length decision method, i.e., successively adding/reducing lags. Previous studies utilized the prior knowledge to define the plausible lag lengths when multi-regressors were included [89,90]. However, in our study, the lagged effect of the same hydrological factor on the water amount varied from place to place as it was affected by the distance and the regional characteristics. Thus, there is no universal predefined lag length estimation available. Instead, we applied the following processing steps to decide the proper lag length of each hydrological factor: (1) building of the regression of the water amount (water level for the river and water volume for the lake) with each hydrological factor individually with different lag lengths and recording the resultant adjusted \(R^2\) (\(\bar{R}^2\)) value and the direction of the coefficient (positive or negative) and (2) selection of the final lag length (starting and ending lags) of each hydrological factor based on two criteria: (a) having the coefficient direction fitting the knowledge (the coefficient of WSCE% and rainfall should be positive, while the coefficient of evapotranspiration should be negative, because more snowmelt and rainfall would increase the downstream water amount, while more evapotranspiration would reduce the water supply) and (b) having the ascending value of \(\bar{R}^2\) (until the highest value). The theory behind our two-step approach is that the use of the \(\bar{R}^2\) value helps to identify the appropriate set of regressors that can explain the variations of the dependent variable.
well, and only selecting the lags having ascending $R^2$ values avoids selecting an unnecessarily long lag length.

Based on the techniques and processing steps mentioned above, by employing Sentinel-1 images, as well as Sentinel-3 data, together with a DEM and ERA5-Land records, we can quantify the causal effect of snowmelt conditions on any large-scale inland waterbody. The overall workflow of our study is illustrated in Figure 2.

**Figure 2.** The overall workflow of the present study, including the processing of the Sentinel-1 synthetic aperture radar (SAR) and Sentinel-3 altimetry data for deriving the wet snow-covered extent (WSCE), surface area, and water level. The water volume estimated from polynomial fitting and integration is detrended by the Seasonal-trend decomposition procedures based on Loess (STL). Together with the hydrological factors derived from the ERA5-Land dataset, the distributed lag model (DLM) is eventually conducted. GRD: Ground Range Detected. The digital elevation model is abbreviated as DEM.

### 4. Results

#### 4.1. Lake Urmia

**4.1.1. Water Level Retrieval**

Two passes of Sentinel-3 over Lake Urmia are processed via SARvatore to estimate the time series average water level of VSs. To examine the results, the water level records provided by altimetry databases are plotted together in Figure 3, with the error bar representing the standard deviation of the water level among the virtual stations (VSs). As our goal is to compare their long-term trends and short-term noisy levels, and because different altimeter sensors have different instrumental biases [91], in this figure, we show the water level variation relative to each record’s mean water level during our sensing period. Firstly, we can observe that there is a high consistency between the two passes’ results (the locations of two passes are illustrated in Figure 1). Most of the time, the differences of their values are within one standard deviation. It is reasonable that two passes’ results still have some differences owing to the reasons including (1) different wind-caused lake surface wave conditions, as their sensing times are not identical, (2) varying water depths on the altimetry signal penetration depth, and (3) and the different surrounding topography-caused signal pollution conditions. Nevertheless, the internal
consistency of each pass and their cross-consistency remain high when comparing to DAHITI and G-REALM’s records. In Figure 3, it is evident that both of our two passes’ raw results have a much smoother trend compared to the smoothed DAHITI result (no raw data is provided by the DAHITI database; thus, no uncertainty is plotted) and the raw G-REALM results, as these databases’ results show more short-term high-frequency fluctuations. The reason is that the along-track resolution of the HPR Sentinel-3 (80 m) we employed is much higher than the Jason-3 altimetry sensor (>2000 m), which is used for DAHITI and G-REALM. Hence, based on these comparisons, it is confirmed that our Sentinel-3 results generally have a higher reliability. To fuse two passes’ observations for the following analysis, we utilized the weighted average calculation and then used the Gaussian smooth method to filter out the noises.

![Figure 3](image)

**Figure 3.** The time series water level variation of Lake Urmia estimated with Sentinel-3. The internal cross-comparison with different Sentinel-3 passes and the external comparison with altimetry databases, including DAHITI and Hydroweb (based on Jason-3 altimetry), are illustrated. The error bars represent the standard deviation of the water level among the virtual stations (VSs). Note: each of the variations is relative to each record’s average height in the studying period.

4.1.2. Surface Area Estimation

To define the optimized Otsu’s threshold value for mapping the surface area of Lake Urmia, we manually selected water/nonwater samples on each VV polarization scene, which showed a clear waterbody boundary. With the aim to densify the usable scenes, both ascending and descending Sentinel-1 image sets were used and processed individually. Their overall backscatter coefficient distributions and the resultant Otsu’s threshold values are shown in Figure 4. Firstly, we can observe that both the flight direction image sets’ sample intensity distributions are bimodal, which confirms the suitability of the Otsu’s method. Secondly, it is found that both sets have nearly the same threshold values; therefore, we average two values to select a fixed value, i.e., −20.8 dB, to be the waterbody classification threshold for Lake Urmia.

Based on the defined threshold value, we can classify the water-covered area of Lake Urmia, i.e., surface area, from both stacks of ascending and descending Sentinel-1 images. For cross- and external comparisons, we plotted the same date’s surface areas estimated with ascending/descending Sentinel-1 images and the GWP product in Figure 5. It is evident that both of our ascending and descending results match perfectly with GWP with an R-square value of around 0.98. These high values prove that the surface area classified with our SAR-Otsu’s approach is highly reliable. Additionally, it is also interesting to realize that there is a systematic offset between our SAR-based estimation and the GWP product, which is based on a coarse optical sensor. Based on a further analysis, we can confirm that it is because of an underestimation of the GWP product due to a high reflectance of lake soil when the water level is low, as well as due to a mixed pixel effect of the coarser spatial resolution (see Discussion 5.1). Thus, the reliability of our SAR Otsu’s approach for accurately delineating the surface area of Lake Urmia is ensured and the time series surface area dynamics of Lake Urmia can be mapped. It is obvious that the change of the surface area of Lake Urmia is dramatic. In the sensing period, the minimum
and maximum surface areas are around 1126.86 (2017/10/31) and 3569.77 (2019/7/11) km². The areal difference is significant, as the maximum size is almost equal to 3.2 times the minimum surface area.

<table>
<thead>
<tr>
<th>Flight direction</th>
<th>Ascending</th>
<th>Descending</th>
</tr>
</thead>
<tbody>
<tr>
<td>Otsu’s threshold value (dB)</td>
<td>-20.678</td>
<td>-20.962</td>
</tr>
</tbody>
</table>

Distribution of backscatter coefficient

Figure 4. The distributions of the backscatter coefficient and the decided Otsu’s threshold values of the ascending and descending Sentinel-1 SAR imagery for the Lake Urmia waterbody classification.

Figure 5. The internal cross-comparison of Lake Urmia’s surface area detected by ascending and descending imagery of Sentinel-1, and the external comparison with the GWP product.

4.1.3. Hypsometry Estimation and Detrended Volume Retrieval

To derive the time series water volume, the hypsometry estimated from the water level-surface area relationship is necessary. Hence, we first plot the time series water level and surface area (derived from both ascending and descending Sentinel-1 SAR as well as the GWP) in Figure 6. A third-order polynomial function is used for fitting as it provides the best R-squared value. Moreover, previous in-situ surveying [47] and study [36] also suggest the utilization of a third-order function for Lake Urmia’s hypsometry. It is found that both of our Sentinel-1 SAR ascending (0.98) and descending (0.99) show a perfect R-squared value, while the GWP (0.94) shows a slightly poorer matching. These values suggest that both our Sentinel-3 altimetry-based water level and Sentinel-1 SAR-Otsu’ approach-based surface area are reliable as they share a high consistency. The slightly lower R-square value of GWP also agrees with the previous finding of the under-estimation of GWP. Eventually, to densify the available observations, we integrate both ascending and descending SAR stacks-derived surface area to build a surface area-water level relationship, which can be written as:

\[ A = 1164.4 \times dL^3 - 3459.1 \times dL^2 + 4267.1 \times dL + 1102.5 \]  

where A and dL represents the surface area and water level difference, respectively. It yields a high fitting R-squared value around 0.99.
Based on the fact that the dynamic water volume should be equal to zero when the water depth (water level difference relative to the lowest water level during the sensing period) is zero, we can derive the fourth-order water volume-water level function by integrating the surface area-water level formula. Thus, Lake Urmia’s water volume-water level function can be written as:

\[
V = 291.10 \times dL^4 - 1153.03 \times dL^3 + 2133.55 \times dL^2 + 1102.50 \times dL
\]  

(4)

where \( V \) represents the dynamic water volume. Based on this equation, the time series water volume of Lake Urmia can be derived. Nevertheless, to mitigate the bias of the container effect of the lake for the following analysis, as described in the Section 3.3, the STL is then applied to estimate the detrended water volume, as plotted in Figure 7. It is clear to observe that the trend of the water volume was almost stable in the first two years, while a peak can be identified in 2019. By subtracting the trend from the estimated water volume, we can obtain the detrended water volume, which represents the real interannual water volume variation.

4.1.4. WSCE Mapping and Hydrological Factors Calculation

Both ascending and descending stacks of images of Sentinel-1 SAR are used to map the WSCE. The summertime imagery of 2018 is selected as the referenced snow-free image for calculating the ratio
image. By applying the threshold value proposed by Nagler and Rott [9,10], the time series WSCE of each flight direction image stack can be estimated. To examine their consistency, we plotted their time series WSCE% relative to Lake Urmia’s watershed in Figure 8. It is found that both flight directions’ results show a high agreement, which proves the internal robustness of our approach. As there is no solid way to externally validate the WSCE in a comparable high resolution [9,76], this cross-track comparison is used for the examination.

Based on the previous data preparation, before processing the DLM analysis, we needed to convert all hydrological factors to the watershed-based unit, i.e., to derive the mean evapotranspiration/rainfall and WSCE%. Together with the estimated detrended water volume, we illustrated all the time series data in Figure 9 to observe their relationships. Firstly, it was found that excluding rainfall shows two peaks each year; other data, including WSCE%, evapotranspiration, and detrended water volume, have only one peak annually. Additionally, the pattern of the evapotranspiration is nearly identical each year. Thus, we can assume that the variation of water resources of Lake Urmia cannot be perfectly explained by rainfall and evapotranspiration data only. Secondly, by comparing the time of the peak of each data, it is found that, in the temporal aspect, the water volume always reaches each year’s maximum value later than the WSCE%. To quantify these factors’ lagged and overlapping influences on the water volume variation, a hydrological analysis based on DLM is required.

![Figure 8](image1.png)

**Figure 8.** The estimation of the wet snow-covered extent (WSCE) percentage of the Lake Urmia watershed using ascending and descending Sentinel-1 SAR imagery.

![Figure 9](image2.png)

**Figure 9.** The time series detrended water volume of Lake Urmia and the WSCE%, as well as hydrological factors, including liquid rainfall and evapotranspiration.
4.1.5. DLM Hydrological Analysis

To process the DLM for quantifying the causal effects of hydrological factors on the detrended water volume variation, it is necessary to decide the proper lag length for each factor. Based on the steps and criteria mentioned in Section 3.5, we summarized the determined lag length in Table 1. Firstly, it was found that the WSCE% has the longest lag length of 80 days, rainfall has 30 days delay, and evapotranspiration shows no delay longer than ten days. As the directions of the coefficient of WSCE%, rainfall, and evapotranspiration are always positive, positive, and negative in different lag lengths testing, respectively, we selected the lag starting from day zero for all factors. Additionally, we can observe that the WSCE% has the highest $R^2$ of 0.85, followed by 0.23 and 0.61 for rainfall and evapotranspiration, respectively. The high/low $R^2$ values of the WSCE%/rainfall indicate that their influential magnitudes on the water volume differ considerably.

Table 1. The summaries of the distributed lag model (DLM) analysis of Lake Urmia and the Brahmaputra River. WSCE: wet snow cover extent.

<table>
<thead>
<tr>
<th>Study Area</th>
<th>Lake Urmia</th>
<th>Brahmaputra River</th>
</tr>
</thead>
<tbody>
<tr>
<td>WATER VOLUME</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dependent Variable</td>
<td>WATER VOLUME (ORIGINAL)</td>
<td>WATER VOLUME (DETRENDED)</td>
</tr>
<tr>
<td>PRE-ANALYSIS (LAG LENGTH DECISION)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>WSCE%</td>
<td>0.64</td>
<td>0.85</td>
</tr>
<tr>
<td>Mean rainfall</td>
<td>0.25</td>
<td>0.23</td>
</tr>
<tr>
<td>Mean evapotranspiration</td>
<td>0.22</td>
<td>0.61</td>
</tr>
<tr>
<td>Lag length (days)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>WSCE%</td>
<td>0–60</td>
<td>0–80</td>
</tr>
<tr>
<td>Mean rainfall</td>
<td>0–20</td>
<td>0–30</td>
</tr>
<tr>
<td>Mean evapotranspiration</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>DISTRIBUTED LAG MODEL (DLM) REGRESSION ANALYSIS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model adjusted $R^2$ (overall)</td>
<td>0.67</td>
<td>0.91</td>
</tr>
<tr>
<td>Long-run propensity (LRP)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>WSCE%</td>
<td>135.68</td>
<td>108.49</td>
</tr>
<tr>
<td>Mean rainfall</td>
<td>60,695.88</td>
<td>−88,737.6</td>
</tr>
<tr>
<td>Mean evapotranspiration</td>
<td>−76,581.40</td>
<td>−120,327.76</td>
</tr>
</tbody>
</table>

Based on the decided lag lengths, the DLM can thus be processed. The results are summarized in Table 1. A significant regression equation is found ($F(14, 81) = 70.78$, $p < 2.2e^{-16}$) with an $\hat{R}^2$ of 0.91. The high $\hat{R}^2$ indicates that our model contains a clear explanation of the variations of the water volume. The overall F-test value of 70.78 proves that our hydrological factors with different lag lengths can reliably predict the water volume. Eventually, based on the LRP of WSCE% estimated from DLM, we can conclude the causal effect of snowmelt on the water resource as, when increasing one percent of the WSCE% during the period of zero to 80 days before, the detrended water volume of Lake Urmia would increase 108.5 cubic meters while holding constant other independent variables.

4.2. Brahmaputra River

4.2.1. Water Level Retrieval

The same Sentinel-3 altimetry data processing approach is employed to estimate the time series water level of the Brahmaputra River. Since the Brahmaputra River is a long river spreading across
various geomorphologies, and the satellite pass ground track location of each spaceborne altimetry differs, there is no direct way to validate our results, as the water level differs in different river sections. Nevertheless, we considered that the water level of each location of the same river should share a similar pattern, with a higher correlation in shorter distance differences. Thus, in total, three VSs were utilized for comparison, including another Sentinel-3 pass (pass 161, processed by the same approach) and two Jason-3 passes (data provided by DAHITI and Hydroweb individually). Their locations and the time series water level results are illustrated in Figures 1 and 10, respectively. Firstly, it is clear that four time series data have highly similar trends and seasonality. Nevertheless, the magnitude of the water level variation of each location shows an ascending order from the upstream to downstream, i.e., around three (DAHITI), four (Sentinel-3 pass 161’s results), seven (Sentinel-3 pass 324’s result), and eight (Hydroweb) meters, respectively. The reason for the more significant water level change in the downstream may be due to the flatter terrain when compared to the upstream. In addition, as we observed in the Lake Urmia case, the DAHITI and Hydroweb databases contain more noises than our HPR Sentinel-3 results, while the comparable patterns remain.

![Figure 10](https://example.com/figure10.png)

**Figure 10.** The time series water level of the Brahmaputra River estimated with Sentinel-3. The internal cross-comparison with different Sentinel-3 passes, and the external comparison with altimetry databases are illustrated. Note that the VS of each record is different, so the magnitude of water level variations differ, while the similar trends are still identifiable.

### 4.2.2. WSCE Mapping and Hydrological Factors Calculation

Both Sentinel-1 ascending and descending image stacks are used to calculate the WSCE of the Brahmaputra River. The summertime image of 2018 was selected as the referenced snow-free image. To investigate the internal consistency, the WSCE% derived from each stack is plotted in Figure 11. Similar to Lake Urmia’s case, a high agreement between both flight directions is observed. It suggests that the quality of our WSCE results should be trustworthy.

We also analyzed the relationship between the hydrological factors and water level before the DLM processing. Their patterns during the sensing period are illustrated in Figure 12. Compared with the Lake Urmia case, the Brahmaputra River shows some remarkable differences. Firstly, it is observed that all four parameters have only one peak per year. Secondly, the temporal delay between the peak of the water level and WSCE% is longer than the time delay of the water level and rainfall. Based on these two findings, we can assume the importance of the rainfall to water amount variation is higher in the Brahmaputra River than Lake Urmia. Nevertheless, since all hydrological factors and water levels show comparable seasonal patterns, the employment of DLM is indispensable to analyze their lagged causal effects on the water level.
4.2.3. DLM Hydrological Analysis

To decide the suitable lag length for each hydrological factor, we built regressions with each factor individually, and the resultant lag length is summarized in Table 1. Firstly, like Lake Urmia, WSCE% has the longest lag length, and evapotranspiration has the shortest lag length. However, it is found that the lag length of WSCE% is nearly two times longer than for Lake Urmia. Another critical difference is that the starting lag length of WSCE% is not zero days. It is because, in the period of lag length between zero to 60 days, the direction of the WSCE%’s coefficient is negative, which violates the background knowledge. Thus, we only selected the lag length from 70 to 150 days (based on the criteria described in Methodology Section 3.5). Therefore, using the prior knowledge to aid the decision of the lag length is critical. Secondly, when comparing the $R^2$ values, all three hydrological factors show comparable results, i.e., higher than 0.90.

The DLM can then be processed based on the decided lag lengths of each factor; the results are summarized in Table 1. A significant regression equation is found ($F(14, 73) = 198.7, p < 2.2e^{-16}$) with a $R^2$ of 0.97. The high $R^2$ indicates that the variation of the water level can be perfectly explained by our model, and the F-test value of 198.7 suggests that hydrological factors with different lag lengths can reliably predict the water level. Based on these examinations, we can conclude our DLM is robust. Hence, on the basis of the estimated LRP of WSCE%, the causal effect of snowmelt to the water resource can be concluded as, when increasing one percent of the WSCE% during the period of 70 to 150 days...
before, the water level of our VS of the Brahmaputra River would increase 0.1 m while holding constant other independent variables.

4.2.4. Flooded Area Estimation

The SAR-Otsu’s approach is also utilized in the Brahmaputra River to delineate the flooding area (or the change of river channels). Both ascending and descending image stacks of Sentinel-1 are processed individually, with their backscatter coefficient distributions and the resultant Otsu’s threshold values shown in Figure 13. Like in Lake Urmia’s case, we can find comparable Otsu’s threshold values. Hence, we average two values to select \(-16.5\) dB as the fixed threshold for the waterbody classification.

<table>
<thead>
<tr>
<th>Flight direction</th>
<th>Otsu’s threshold value (dB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ascending</td>
<td>(-16.375)</td>
</tr>
<tr>
<td>Descending</td>
<td>(-16.389)</td>
</tr>
</tbody>
</table>

![Figure 13](image-url) The distributions of the backscatter coefficient and the decided Otsu’s threshold values of the ascending and descending of Sentinel-1 SAR imagery for mapping the flooded area of the Brahmaputra River.

For Lake Urmia, we combined the surface area with the water level to derive the hypsometry; on the contrary, for the Brahmaputra River, we can use the water level as the external data to examine our mapped flooded area. It is based on the fact that, when the flooded area is enlarging, the water level will rise simultaneously. Thus, we plotted the time series flooded area, water level, and WSCE% in Figure 14. Firstly, it is found that, generally, the trends of the flooded area derived from both ascending and descending Sentinel-1 SAR imagery stacks are agreeable, although internal noises and the differences between them are also clear. Moreover, their noisy levels are higher when compared to the same descending/ascending stack-based WSCE%, as illustrated in Figure 11. We consider it is because (1) the short-term variation of the flooded area is more abrupt and dynamic compared to the long-term gradient snowmelt process, and (2) the WSCE% mapping is by thresholding the ratio image relative to a referenced image (as described in Methodology Section 3.4), while the flooded area is mapped by directly applying the Otsu’s threshold to each imagery, so the influence of the local incidence angle variation of each image is severer. Hence, to reduce the fluctuations, we average the ascending and descending stacks’ results to derive a smoother trend. Secondly, by comparing the average flooded area with the water level, we can observe a temporally highly correlated pattern. It agrees with our prior knowledge and, also, indirectly proves the credibility of our Sentinel-1 SAR-based flooded area estimations, as well as Sentinel-3 altimetry-based water level records. Finally, the synchronized flooded area and water level trend temporally lag when compared to the WSCE% trend. Around 110 days of the temporal gap between the peak of WSCE% and water level/flooded area are noted, which fall within the estimated lag length of WSCE% (70 to 150 days) for the DLM analysis (Table 1). This agreement also strongly proves the reliability of our DLM lag length decision.
5. Discussions

5.1. The Cause of the Different Lake Surface Areas Detected by Sentinel-1 SAR and GWP Product

To compare the surface area mapped by SAR (Sentinel-1) and the GWP, we plotted their estimations for the same dates in Figures 5 and 16. It can be observed that the GWP systematically underestimates the surface area compared to SAR-derived estimations. To examine that this effect is not caused by the overestimation of our SAR-based results, we plot the same date’s high-resolution Sentinel-2 imagery in Figure 16 for comparison. It must be noted that, due to the reasons that (1) there is no absolute definition of the lake boundary as stated in [4], (2) the calibration procedures of optical imagery would also affect
the waterbody delineation, and (3) our intention is to not include another threshold-based classification approach (such as NDWI), as it would cause another uncertainty; we compare them visually. It is clear that our SAR-based surface area agrees well with the blueish waterbody boundary observed in the Sentinel-2 scene. On the contrary, the GWP underestimates the surface area. Especially on 2017/10/31, the upper and lower parts of the lake obviously remained connected in the Sentinel-2 image, which was also delineated correctly in our Sentinel-1-based results. Nevertheless, the GWP suggests the lake is separated into two parts. This underestimation-caused separation frequently happens in the dry season (as marked in Figure 16b). The reason for this underestimation is due to the high reflectance signal in MODIS near-infrared (NIR) and red bands when water is very shallow. During the low water level period, Lake Urmia’s bright lake soil with a high salt content dominates the spectral reflectance sensed by the optical MODIS sensor. Furthermore, the global operating GWP approach is treating mixed pixels that result from coarser spatial resolutions rather conservatively. Therefore, these shallow waters remain undetected, which results in systematic bias during low water levels, because the GWP is based on a universal approach operating on a global scale and cannot fit all waterbodies perfectly. On the contrary, as mentioned in Methodology, we manually selected water/nonwater samples to decide the customized Otsu’s threshold value for Lake Urmia.

Another finding is that the SAR-based approach can only be implemented during the lake water depletion period, i.e., the dry season. As shown in Figure 16b, we found that all valid SAR-based surface area estimations occur between May and October. To investigate the reasons, firstly, we plotted the rainfall records provided by the ERA5-Land dataset. It is found that the water depletion period is highly correlated but slightly delayed due to the rainfall decreasing period. Hence, we assume the reason why the SAR imagery cannot map the lake boundary properly in the lake water accumulation period, i.e., rainfall-dominated wet season, is that a series of rainfall events would cause the surrounding regions of the lake to have a wet surface due to (1) sudden flooding events of the lake and/or (2) direct ground dampening caused by the rainfall. As we classify the waterbody based on the backscatter coefficient of the SAR signal, which is mainly influenced by the surface roughness and wetness, we can hardly distinguish whether the water is in the lake or on the surface of the surrounding dampened areas. Thus, in the lake water accumulation period, i.e., wet season, the SAR-based observations would dramatically overestimate the surface area. For instance, given the example of 2019/4/7 shown in Figure 16a, in the Sentinel-2 scene, it is found that the surrounding area of the lake is much wetter (more blueish) and shows more small flood-caused tributaries and braids compared to the other three dates of the dry season. On the contrary, the GWP estimations during the wet season reveal the increasing surface water area more accurately and do not overestimate due to the wet soil.

Based on the above-mentioned discussions, we summarized the pros and cons of using SAR and multispectral sensors for waterbody mapping in Table 2.

Table 2. The comparison of the advantages and disadvantages of using SAR and the multispectral sensor for waterbody detection.

<table>
<thead>
<tr>
<th>Sensor</th>
<th>SAR</th>
<th>Multi-Spectral</th>
</tr>
</thead>
<tbody>
<tr>
<td>Classification basis</td>
<td>Backscatter coefficient</td>
<td>Illumination reflectance</td>
</tr>
<tr>
<td>Primary sensing factor</td>
<td>Surface roughness</td>
<td>Surface dielectric property</td>
</tr>
<tr>
<td></td>
<td>Surface dielectric property</td>
<td>Imagery calibration</td>
</tr>
<tr>
<td>Possible noise source</td>
<td>Winds, flooding, soil moisture, vegetation, ice, and snow cover</td>
<td>Cloud, mist, floating vegetation, turbidity, highly eutrophicated waters, high sediment load, ice, and snow cover</td>
</tr>
<tr>
<td>Advantage</td>
<td>All-weather and illumination-independent sensing ability</td>
<td>Cloud penetration</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Straightforward to interpret The abundance of openly accessible sensors and images less sensitive to soil moisture and usable during wet season</td>
</tr>
<tr>
<td>Disadvantage</td>
<td>Affected by wind, flooding, soil moisture</td>
<td>Affected by cloud and illumination conditions</td>
</tr>
</tbody>
</table>
Figure 16. (a) The comparison of the surface area of Lake Urmia detected by Sentinel-1 SAR imagery and the GWP product and the same date high-resolution Sentinel-2 imagery. The mean rainfall volume is plotted in (b) to identify the dry season, with the lake water depletion periods marked in blue boxes. The date that shows lake disconnection in the GWP is marked with black points in (b).
5.2. The Necessity of the Detrending Process for the Lake Water Volume before a Hydrological Analysis

In the present study, we employed the STL seasonal decomposition technique for detrending, i.e., removing the annual aggregation effect. The necessity of detrending can be proven by the much-higher $R^2$ value of detrended water volume (0.91) when compared to the original data (0.67), as summarized in Table 1. The significant enhancement of the water volume variation explanation agrees with our hypothesis: the “container effect” of the lake, i.e., the intra-annual trend of water volume, would hinder the accuracy of the DLM analysis between the water volume and hydrological factors.

5.3. The Lag Length Difference between Each Hydrological Variable and Study Area

As mentioned in Methodology Section 3.5, based on the $R^2$ value, we can decide the suitable lag length of each hydrological factor for the lake and river DLM analyses. The final lag lengths of each case are summarized in Table 1. It is found that, for both cases, the lag length of the WSCE% is always the longest among the three factors, and evapotranspiration is the shortest (nearly instant). This finding agrees with the prior knowledge that the snowmelt water takes a much longer time to aggregate and eventually influences the downstream waterbody, because snow-covered regions are usually located at much higher elevation zones, which have longer distances to the downstream waterbody. Moreover, snowmelt water would not only form a direct surface streamflow but, also, infiltrate into the ground and soil [78–80]. In contrast, although the rainfall would also infiltrate into the ground, its distribution is usually broader but not limited in the high elevation zone. It might also spread over the whole watershed of the waterbody, including the waterbody itself. Hence, the lag length of the rainfall is much shorter than the snowmelt. For evapotranspiration, it is straightforward that its influence is rapid, as it would immediately reduce the amount of water of the waterbody.

Additionally, it is also found that the lag length of the WSCE% of the Brahmaputra River is nearly two-fold of Lake Urmia. We assume this is due to the Brahmaputra River being characterized by a much narrower and longer watershed along the tributary when compared to Lake Urmia. Consequently, the distance between the snow-covered area and the waterbody/VIS is much more distant, and thus, the lagged effect of the snowmelt is more significant.

5.4. The Cause of the Long Lag Length of the SAR Sensor-Based WSCE%

Another reason for the long lag length of SAR sensor-based WSCE% is because of the unique way that SAR detects snowmelt processing and shows the temporal offset to the real condition. Based on the in-situ measurements and modeling simulation [11], it is known that the whole snow melting process can be divided into three phases, i.e., moistening, ripening, and runoff. In each phase, the contents of the snow water equivalent (SWE) (the total mass of water, including both liquid and solid water, stored in the form of snow) and the liquid water content (LWC) (the mass of liquid water inside the snowpack) differ, and the backscatter coefficient of SAR also changes accordingly and results in a U-shaped curve (Figure 17). In detail, firstly, in the moistening stage, the diurnal melting-freezing cycles gradually increase the LWC and lead to the gentle decreasing of the backscatter coefficient; in the ripening stage, the LWC significantly increases (while the SWE still remains the same), and the backscatter coefficient rapidly decreases and reaches the minimum value at the end of the ripening stage when the snowpack is saturated; and, finally, during the runoff stage, as the snowmelt water releases, both the LWC and SWE decrease, and the backscatter coefficient increases accordingly [11].

It must be noted that, as the backscatter coefficient of SAR would already slightly drop in the middle of the first moistening stage, the snow would already be detected and regarded as wet snow by the backscattering ratio threshold method [9–11]. However, the snowmelt water is actually released only until the final runoff stage starts. Namely, the SAR-detected wet snow-covered period would (1) happen much earlier and (2) last much longer than the real snowmelt water-releasing period, as illustrated in Figure 17. As a result, the lag length of the WSCE% would be longer than expected, as it acts more like a leading indicator of the real snowmelt water generation time.
which is reasonable, as more evapotranspiration would reduce the water amount. On the contrary, processing. It reduces the importance of rainfall, which is already low in comparison to the river’s case. The relationship also shows a much weaker linkage. Furthermore, by comparing with the Brahmaputra River case (0.94), the lake’s rainfall-water volume becomes negative in the latter one. However, it must be noted that, in both scenarios, the coefficients are negative, which is reasonable, as more evapotranspiration would reduce the water amount. On the contrary, the coefficient of rainfall should be positive, as more rainfall would cause more water volume. However, only the case of the Brahmaputra River shows a positive value. We conclude the main reason causing the negative coefficient value of rainfall for the Lake Urmia case is the detrending processing. It reduces the importance of rainfall, which is already low in comparison to the river’s case. In detail, by comparing the rainfall’s coefficient in the Lake Urmia case estimated with the original and detrended water volume, we can be observed that the value is positive in the former scenario but becomes negative in the latter one. However, it must be noted that, in both scenarios, the coefficient of the WSCE% and evapotranspiration remain positive and negative, respectively, which agrees with the prior knowledge. It is because the relationship (or importance) between the rainfall and water volume is much weaker than the WSCE% and evapotranspiration-to-water volume in the lake’s case (as described in Section 4.1.4). It can also be proven by the much lower maximum $R^2$ value of rainfall (0.23) compared to the wet SCE% (0.85) and evapotranspiration (0.61), as shown in Table 1. Furthermore, by comparing with the Brahmaputra River case (0.94), the lake’s rainfall-water volume relationship also shows a much weaker linkage.

![Figure 17](image_url)

**Figure 17.** The relationship of SAR-based wet snow detection and the snowmelt processing (modified from [11]). The amount of liquid water content (LWC), snow water-equivalent (SWE), and the released snowmelt runoff together with the SAR signal backscatter coefficient in different snowmelt phases are illustrated.

### 5.5. The Reason of the Negative LRP of Rainfall in DLM of the Detrended Lake Water Volume

In Sections 4.1.5 and 4.2.3, we only analyzed the LRP of the WSCE%, due to the fact that we focused on depicting the casual effects of increasing the WSCE on the water resource. On the contrary, the other two hydrological factors (mean rainfall and evapotranspiration) are the control variables that are not the investigated targets but help mitigate the omitted variable bias [87] to ensure the robustness of the DLM analysis. Including control variables is indispensable, as there might be correlations between independent variables, which would mislead the estimated lag coefficient and the interested LRP.

Nevertheless, we can still analyze the coefficient of the other two hydrological factors, as shown in Table 1. For evapotranspiration, in both the river and lake cases, the coefficients are negative, which is reasonable, as more evapotranspiration would reduce the water amount. On the contrary, the coefficient of rainfall should be positive, as more rainfall would cause more water volume.

![Diagram](image_url)
5.6. Current Limitations and Future Goals

In the present study, we proposed a cloud-free and illumination-independent comprehensive inland water monitoring approach based on freely accessible and sustainable data/services, which guarantees methodological reproducibility and enables users to holistically and continuously monitor the dynamics of any large-scale waterbody. The econometrics’ DLM further facilitates the accurate analysis of the lagged causal effect of the snowmelt condition on the water amount. Nevertheless, there are still some limitations in our present study. Hence, future improvements are required.

Firstly, as mentioned above, it is found that the high spatial resolution, cloud-free SAR backscatter-Otsu-based surface area detection method can only be implemented in the lake water depletion period, i.e., dry season, due to the sensitivity of the SAR signal to the surface wetness. On the contrary, the GWP provides continuous and daily observations in all seasons, as shown in Figure 16b, although it has some disadvantages, such as systematic underestimation due to the high spectral reflectance of the lake’s bottom when the water is shallow (as shown in Figures 5 and 16) and coarse spatial resolution. Yet, in a future study, we would integrate both multispectral and SAR sensor-based surface areas to achieve a robust, high temporal and spatial resolution continuous surface area mapping strategy.

Secondly, in our study, we used the WSCE% for representing the snowmelt condition. Although it is efficient and straightforward to depict the whole snowmelt dynamics [9,11,76,77], it cannot provide the information of how much snowmelt water is actually generated during the snowmelt period. Instead, the measurement of either SWE or LWC should be used. Unfortunately, although there are many studies aiming at quantifying the SWE or LWC with SAR data, so far, there is no reliable conclusion provided [9]. In addition to the wetness of the snowpack, other factors such as soil moisture, surface roughness, snow grain size, snow density, and temperature would also affect the SAR signal. Consequently, the utilization of either the inversion technique or empirical model is usually required [9]. The employments of multispectral and passive sensors are the alternative; however, they would be affected by either cloud cover and polar darkness or poor spatial resolution, as mentioned in the Introduction. Hence, in the future study, we would explore the SAR-SWE relationship using the inversion technique and compare the results with the passive sensor-based SWE products provided by Copernicus Global Land Service [92] and GlobSnow [92].

In addition, in the current study, we assumed the water balance of Lake Urmia could be explained by the WSCE, rainfall, and evapotranspiration. Since the surface water inflow can be estimated by the lagged effects of the snowmelt and rainfall water, and as Lake Urmia is an endorheic lake, there is no stream outflow. Nevertheless, the influence of groundwater seepage and leakage were not measured in the present study due to (1) the lack of piezometric records around the lake and (2) their influences on the water balance are normally trivial [93]. Thus, the groundwater is commonly ignored in the lake water balance studies [94,95]. However, to enhance the completeness of monitoring the water resource variations of Lake Urmia, the hydrological models such as the MODular three-dimensional finite-difference groundwater flow model (MODFLOW) [96] or a simple groundwater input-output model [97] could be included in the future.

Another improvement that could be considered is using the modified DLM for a hydrological regression analysis. In the present DLM applied in this study, the weight of each lag is set equally. On the contrary, the advanced constrained DLM uses a lag-based smooth function, such as linear declining lag weights or polynomial/spline distributed lag models [85]. These constrains would allow a more accurate estimation of each lag weight of each independent variable, as the potential multicollinearity between different lag of each independent variable might hinder the estimation of each lag coefficient in unconstrained DLM (although the LRP, the main investigated target in the current study, would remain reliable) [88]. Moreover, the autoregressive distributed lag model (ARDL) [98] should also be explored as the past value of the water amount might also influence the current water amount.
The rain-on-snow (ROS) hydrometeorological phenomena should also be investigated for its influence on the snowmelt-led water resource variations. It is found that reactive oxygen species (ROS) events would weaken the snowpack structure and accelerate the melting of the snowpack due to heat transfer, which leads to an earlier and shorter melting period and even causes floods [99–101]. Moreover, in the long run, it can also result in a thinner springtime snowpack with less LWC and SWE and, eventually, reduces the water supply in the dry season [102,103]. Practically, when ROS events take place, they might form an icy crust on the snowpack, as the relatively warm precipitation would melt the surface of the snowpack and seep through the snow, which refreezes during the nighttime [104]. This icy crust would change the surface roughness and the dielectric property and may mislead the SAR-based WSCE mapping results. Hence, the relationship between the WSCE, precipitation, and SAR backscatter coefficient should be examined.

Finally, to utilize the proposed strategy’s broad applicability, establishing an automatic processing pipeline is worth investigating. In detail, the whole workflow (as shown in Figure 2) can be divided into five parts, including water level (volume) retrieval, water-covered area extraction, WSCE estimation, watershed delineation, and hydrological factors calculation, as well as an analysis of the causal effects of snowmelt. Each part can be automatized as follows: the most complex first part consists of a user-defined SARvatore task processed with a BRAT command line-based processing script and STL processing scripts; the second part is composed of a manual selection of training samples and automatic applying of the decided Otsu’s threshold value for classification; the third part is conducted with the stacking of SAR images and batch ratio/thresholding calculations; the fourth part is the watershed delineation, and hydrological factor calculations can be conducted by built-up Geographic Information System (GIS) tools and written scripts, respectively; at last, the fifth part can be automatized with predefined criteria of selecting the proper lag length for the final DLM analysis. Thus, based on part-wise automatization, the proposed strategy of monitoring large-scale waterbody’s dynamics and assessing snowmelt’s causal effects can be implemented more efficiently.

6. Conclusions

The significance of inland water resources has been gradually emphasized, especially in times of global warming. Within all waterbody types, lake and river systems are critical not only for human society but, also, for natural ecosystems. In addition to the temperature and rainfall variability, climate change alters the snowmelt phenology and snow distribution, while runoff originating from snowmelt is the primary water supply source for many river and lake systems. Therefore, monitoring the dynamics and variations of water resources, as well as the snowmelt conditions, is necessary to understand cryo-hydrosphere interactions under the influence of climate change. So far, only a few studies addressing this question relying on remote-sensing techniques exist. Most previous studies utilize either in-situ measurements or multispectral sensors, which are limited to small-scale regions or affected by cloud cover, respectively. In the present study, we employed the latest spaceborne Sentinel-1 SAR and Sentinel-3 altimetry data to achieve a high resolution, cloud-free, and illumination-independent comprehensive inland waterbody dynamics (time series surface area and water level) sensing strategy. Moreover, in contrast to previous studies utilizing in-house SAR and altimetry data-processing algorithms, we employed freely available cloud-based services to ensure the broad applicability and transferability.

Based on the altimetry and SAR data, the water level and the water-covered extent (surface area of lakes and the flooded area of rivers) can be successfully estimated. For the surface area mapping, the SAR backscatter coefficient-based Otsu’s threshold has proven to be capable of accurately classifying the waterbody boundary by comparing it with the Global WaterPack (GWP) and high-resolution Sentinel-2 imagery. Nevertheless, we found that this approach can only be implemented in the lake water depletion period, i.e., dry season, due to the fact that the rainfall-caused surface wetness over nonwaterbody regions would mislead the backscatter-based classification. For the water level estimation, our results show a high consistency between different passes and even out-performs the
well-known altimetry databases owing to the finer along-track resolution of Sentinel-3. Moreover, by fusing both the surface area and water level information, we can achieve comprehensive inland water dynamics monitoring. For our selected study sites, such as Lake Urmia, we could estimate the hypsometry and derive the water volume change. Via utilizing the STL decomposition technique, we can mitigate the “container effect” of the lake and extract the interannual water variation. For the Brahmaputra River, both time series of the water level and the flooded area can be identified and used for cross-comparison. The synchronized trend observed between them also suggests the high reliability of our derived Sentinel-1 and Sentinel-3 results.

Last, but not least, together with the WSCE mapped with SAR imagery, we can analyze the influence of snowmelt on water resource variations. With the aim to handle the lagged causal effect of snowmelt and mitigate the omitted errors in the regression, the DLM initially developed for econometric applications is employed in the present study. In the lag lengths pre-analysis, we found that the WSCE% has the longest lag length compared to rainfall and evapotranspiration. It is due to the infiltration process of snowmelt water and the fact that the snow-covered area is usually located in the high elevation region, which is more distant to downstream water level-investigated VS. The lag length of the WSCE% was also observed to be nearly twice as long in the Brahmaputra River compared to Lake Urmia owing to the shape differences of their watersheds. Furthermore, the temporal offset between the snowmelt water generation and the SAR-based WSCE detection is also discussed, which is one reason for the long lag length. Eventually, the casual effect of snowmelt conditions on inland water resources is unbiasedly quantified with the DLM: for Lake Urmia, when increasing one percent of the WSCE% during the period of zero to 80 days before, the detrended water volume would increase by 108 cubic meters while hold still other factors; for the Brahmaputra River, when increasing one percent of the WSCE% during the period of 70 to 150 days before, the water level of the investigated VS would increase 0.1 m, while the other independent variables remained constant.

Author Contributions: Conceptualization, Y-L.S.T.; Formal analysis, Y.-L.S.T.; Methodology, Y.-L.S.T. and I.K.; Resources, I.K.; Supervision, A.D. and N.O.; Validation, Y.-L.S.T.; Visualization, Y.-L.S.T.; Writing—original draft, Y.-L.S.T. and I.K.; Writing—review & editing, I.K., A.D. and N.O. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Acknowledgments: Support by the German Academic Exchange Service (DAAD) fellowship to Ya-Lun S. Tsai is gratefully acknowledged. Special thanks to Marco Restano (ESA-ESRIN) for supporting the SARvatore service and Yao-Chun Huang (LSE) for guiding the DLM analysis.

Conflicts of Interest: The authors declare no conflict of interest.

References

Remote Sens. 2020, 12, 3896


30. Zarghami, M. Effective watershed management; Case study of Urmia Lake, Iran. Lake Reserv. Manag. 2011, 27, 87–94. [CrossRef]


43. Tourian, M.; Elmi, O.; Chen, Q.; Devaraju, B.; Roohi, S.; Sneeuw, N. A spaceborne multisensor approach to monitor the desiccation of Lake Urmia in Iran. Remote Sens. Environ. 2015, 156, 349–360. [CrossRef]


47. Alipour, S. Hydrogeochemistry of seasonal variation of Urmia Salt Lake, Iran. Saline Syst. 2006, 2, 9. [CrossRef]


51. Immerzeel, W.W.; Van Beek, L.P.H.; Bierkens, M.F.P. Climate change will affect the Asian Water Towers. Science 2010, 328, 1382–1385. [CrossRef]
54. Richardson, W.; Thorne, C.R. Multiple thread flow and channel bifurcation in a braided river: Brahmaputra Jamuna River, Bangladesh. *Geomorphology* 2001, 38, 185–196. [CrossRef]
70. Hayashi, M.; Van Der Kamp, G. Simple equations to represent the volume–area–depth relations of shallow wetlands in small topographic depressions. *J. Hydrol.* 2000, 237, 74–85. [CrossRef]
72. Fathian, F.; Dehghan, Z.; Esfahani, S. Analysis of water level changes in Lake Urmia based on data characteristics and non-parametric test. *Int. J. Hydrol. Sci. Technol.* 2014, 4, 18–38. [CrossRef]
74. Hassan, A.A.; Jin, S. Lake level change and total water discharge in East Africa Rift Valley from satellite-based observations. *Glob. Planet. Chang.* 2014, 117, 79–90. [CrossRef]
Remote Sens. 2020, 12, 3896


77. Tsai, Y.-L.S.; Dietz, A.J.; Oppelt, N.; Kuenzer, C. Tsai A Combination of PROBA-V/MODIS-based Products with Sentinel-1 SAR Data for Detecting Wet and Dry Snow Cover in Mountainous Areas. Remote Sens. 2019, 11, 1904. [CrossRef]


81. Almon, S. The Distributed Lag Between Capital Appropriations and Expenditures. Econ. J. Econ. Soc. 1965, 33, 178. [CrossRef]


83. Allen, M.J.; Sheridan, S.C. Mortality risks during extreme temperature events (ETEs) using a distributed lag non-linear model. Int. J. Biometeorol. 2015, 62, 57–67. [CrossRef]


**Publisher’s Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).