International Journal of Applied Earth Observation and Geoinformation 135 (2024) 104283

Contents lists available at ScienceDirect

International Journal of Applied Earth Observation and Geoinformation

journal homepage: www.elsevier.com/locate/jag

Multi-temporal remote sensing of inland surface waters: A fusion of sentinel-1&2 data applied to small seasonal ponds in semiarid environments

Francesco Valerio [a,b,c,*], Sérgio Godinho [c], Gonçalo Ferraz [a,b,e,g], Ricardo Pita [f], João Gameiro [a,b,d], Bruno Silva [f], Ana Teresa Marques [a,b,d], João Paulo Silva [a,b,d,e]

[a] CIBIO, Centro de Investigação em Biodiversidade e Recursos Genéticos, InBIO Laboratório Associado, Campus de Vairão, Universidade do Porto, 4485-661 Vairão, Portugal
[b] BIOPOLIS Program in Genomics, Biodiversity and Land Planning, CIBIO, Campus de Vairão, 4485-661 Vairão, Portugal
[c] EaRSLab—Earth Remote Sensing Laboratory, University of Évora, 7000-671 Évora, Portugal
[d] CIBIO, Centro de Investigação em Biodiversidade e Recursos Genéticos, InBIO Laboratório Associado, Instituto Superior de Agronomia, Universidade de Lisboa, 1349-017 Lisboa, Portugal
[e] Estação Biológica de Mértola, 7750-329 Mértola, Portugal
[f] MED – Mediterranean Institute for Agriculture, Environment, and Development & CHANGE – Global Change and Sustainability Institute, Institute for Advanced Studies and Research & Unidade de Biologia da Conservação, Departamento de Biologia, Universidade de Évora, Pólo da Mitra, Ap. 94, 7006-554 Évora, Portugal
[g] Departamento de Biologia, Faculdade de Ciências, Universidade do Porto, 4169-007 Porto, Portugal

A R T I C L E I N F O

Keywords:
multispectral instrument (MSI)
Synthetic aperture radar (SAR)
Surface water occurrence
Surface water extent
Small water bodies
Google Earth Engine

A B S T R A C T

Inland freshwaters are essential in maintaining ecological balance and supporting human development. However, comprehensive water data cataloguing remains insufficient, especially for small water bodies (i.e., ponds), which are overlooked despite their ecological importance. To address this gap, remote sensing has emerged as a possible solution for understanding ecohydrological characteristics of water bodies, particularly in water-stressed areas. Here, we propose a novel framework based on a Sentinel-1&2 local surface water (SLSW) model targeting very small (<0.5 ha, $Mdn \approx 0.031$ ha) and seasonal water bodies. We tested this framework in three semiarid regions in SW Iberia, subjected to distinct seasonality and bioclimatic changes. Surface water attributes, including surface water occurrence and extent, were modelled using a Random Forests classifier, and SLSW time series forecasts were generated from 2020 to 2021. Model reliability was first verified through comparative data completeness analyses with the established Landsat-based global surface water (LGSW) model, considering both intra-annual and inter-annual variations. Further, the performance of the SLSW and LGSW models was compared by examining their correlations for specific periods (dry and wet seasons) and against a validation dataset. The SLSW model demonstrated satisfactory results in detecting surface water occurrence ($\mu \approx 72\,\%$), and provided far greater completeness and reconstructed seasonality patterns than the LGSW model. Additionally, SLSW model exhibited a stronger correlation with LGSW during wet seasons ($R^2 = 0.38$) than dry seasons ($R^2 = 0.05$), and aligned more closely with the validation dataset ($R^2 = 0.66$) compared to the LGSW model ($R^2 = 0.24$). These findings underscore the SLSW model's potential to effectively capture surface characteristics of very small and seasonal water bodies, which are challenging to map over broad regions and often beyond the capabilities of conventional global products. Also, given the vulnerability of water resources in semiarid regions to climate fluctuations, the present framework offers advantages for the local reconstruction of continuous, high-resolution time series, useful for identifying surface water trends and anomalies. This information has the potential to better guide regional water management and policy in support of Sustainable Development Goals, focusing on ecosystem resilience and water sustainability.

* Corresponding author.
E-mail address: fvalerio@cibio.up.pt (F. Valerio).

https://doi.org/10.1016/j.jag.2024.104283
Received 4 July 2024; Received in revised form 15 November 2024; Accepted 20 November 2024
Available online 29 November 2024

1. Introduction

Inland freshwater systems play a key role in supporting biological cycles and promoting ecosystem services (De Groot et al., 2018; Gleick, 1998), yet they represent a limited natural resource with scarce representation (~3%) in the Earth's biosphere, of which only a portion (~35 %) represents a renewable resource globally available (Downing et al., 2006). The mounting evidence of water scarcity in many regions and associated socio-economic and ecological impacts has progressively pressed for advancements in freshwater detection and monitoring strategies. Such advancements can mitigate stressors and overexploitation related to natural freshwater systems, and enhance resilience towards climate change (Erwin, 2009). As a result, the spatiotemporal dynamics of global and local freshwater systems have emerged as topics of prime interest, aligning with the development of Sustainable Development Goals (SDGs; UN, 2022).

Spatiotemporal information from Earth Observation (EO) Satellites has opened a new era for assessing qualitative and quantitative dynamic attributes of freshwater resources (Bhaduri et al., 2016; Petropoulos et al., 2015). EO data covering large geographic areas represent an important milestone for sustainable water management, especially in water-stressed areas such as arid and semiarid regions (Heimhuber et al., 2016; Tulbure et al., 2016; Wickens, 1998). In this context, most research has been devoted to inferring water characteristics and dynamics over the long term through quantitative remote sensing methods. Among these, multi-spectral information from MODIS (Moderate Resolution Imaging Spectroradiometer) is notable for its exceptional temporal resolution. However, its spatial resolution (250 m) is relatively coarse, hampering its application for monitoring small and seasonally unstable water bodies. In an attempt to overcome this limitation and improve surface water mapping, substantial progress has been made in employing advanced instruments, such as public sensors affiliated with Landsat missions and their derived products, ensuring higher spatial, radiometric, and spectral resolution (Pekel et al., 2016; Tulbure et al., 2016; Zou et al., 2018).

However, despite their critical importance for numerous ecosystems, small-sized water bodies (< than 5 ha; De Meester et al., 2005), often referred to as 'ponds', are rarely considered in large-scale inventories and mapping programs (Céréghino et al., 2008; Wang et al., 2022). They are hypothesized to be more frequent than large water bodies (i.e., lakes), and to cover a far greater fraction of Earth's surface (Downing et al., 2006). Importantly, ponds and associated freshwater habitats are among the most endangered ecosystems worldwide, and provide refuge to approximately 10 % of known species (Reid et al., 2019). Therefore, the development of satellite-based solutions for detecting and monitoring small-sized water bodies is of crucial importance to improve conservation and sustainable management planning of water resources (Ozesmi and Bauer, 2002; Pekel et al., 2016). This is particularly relevant in semiarid areas where environmental conditions are expected to worsen due to widespread extreme heat events, desertification processes, increased salinity, and agricultural intensification (Reid et al., 2019; Wickens, 1998).

Recently, the high-quality data provided by Sentinel missions have gained ground over Landsat-derived data for delineating and mapping small water bodies, given their better compromise between spectral, temporal, and spatial resolution (Wang et al., 2022). Another emerging practice to assess seasonal dynamics in small-sized water bodies relies on the use of multi-temporal multi-sensor imageries (Ozesmi and Bauer, 2002; Vanderhoof et al., 2023; Tulbure et al., 2016; Tulbure et al., 2022), with focus on Sentinel-1 and Sentinel-2 (Jiang et al., 2022; Mayer et al., 2021; Radoux et al., 2016; Huang et al., 2018), or a combination of both (Bioresita et al., 2019; Chen and Zhao, 2022; Vanderhoof et al., 2023) for classification purposes. Notably, while the combination of Sentinel-1 and Sentinel-2 has the potential to enhance the comprehension of surface water distribution and seasonality, particularly in small water bodies, this remains poorly investigated in the current literature, making it still unclear whether Sentinel-derived products offer higher performance compared to other high-resolution approaches (e.g., Landsat) for detecting and quantifying small-sized water bodies.

In this study, we propose an innovative data-fusion approach using Sentinel-1 and Sentinel-2 information in semiarid environments to address data gaps and improve predictions on the spatiotemporal trends of dynamic surface water, specifically targeting open and very small (<0.5 ha) water bodies. To assess the effectiveness of the approach, water bodies were mapped, and the Sentinel-based local surface water (SLSW) predictions were compared with those from Landsat-based global surface water (LGSW) (Pekel et al., 2016). Specifically, we aimed to: (*i*) identify the best Sentinel-1&2-derived metrics to detect surface water occurrence, and understand the multi-spectral reflectance and backscattering responses involved; (*ii*) utilize SLSW predictions to characterize the intra- and inter-year (2020–2021) dynamic trends (seasonality patterns) in surface water occurrence and extent in three semiarid regions; and (*iii*) compare the SLSW archives' completeness with those derived from the LGSW (Pekel et al., 2016), and evaluate the agreement between the estimated water extent from both data sources (Sentinel and Landsat), with a specific focus on the wet season (January-April) and the dry season (June-September), using a validation dataset.

2. Materials and methods

2.1. Study areas

Three study areas within the Mediterranean biome in southwestern Iberia (Fig. 1a) were selected to test the reliability of Sentinel-based data for accurate water surface mapping: two in Baixo Alentejo (southern Portugal), and one in Extremadura (south-eastern Spain). In these areas, small water bodies were detected (Fig. 1b) and mapped (Fig. 1c).

General similarities between the areas include climatic, topographic, and environmental conditions. In particular, a pronounced climatic variation exists between the wet and dry seasons, inducing severe influences on the water cycle. All areas are influenced by a semiarid climate with hot, dry summers and cool, wet winters, with most of the precipitation occurring from autumn to spring (AEMET, I. P. M. A, 2011). Typical weather parameters across the three areas were averaged yearly between 2012 and 2022 from the E-OBS daily gridded meteorological data (Cornes et al., 2018), resulting into an annual average temperature of 17.42 °C (±SD = 0.52) and a maximum temperature of 23.98 °C (±SD = 0.76), whereas an annual average precipitation amount of 450.69 mm (±SD = 122.24). Besides, soils in these areas are generally poor and, consequently, have limited water-retention capacity. As a result, water bodies are mostly very small (Mdn ≈ 0.031 ha; see Figure SM2.1), and sparsely distributed (Fig. 1). The three areas are characterized by plains and gentle slopes, dominated by pastures and extensive cereal agriculture. Each area is located within a Special Protection Area (Natura 2000 Network): Castro Verde (CV), Vale do Guadiana (VG), and La Serena y Sierras Periféricas (LS). These areas are included within a climate change hotspot, where drought events and water deficit anomalies are expected to increase (Samaniego et al., 2018).

2.2. Study design

The study design is comprehensively depicted in Fig. 2, encompassing a series of distinct steps, ranging from water bodies' inventories and remote sensing data collection and processing, to the subsequent modelling stages for estimating surface water occurrence and extent as high-frequency time series.

2.2.1. Data acquisition and pre-processing

2.2.1.1. Water bodies data collection. The identification and delineation

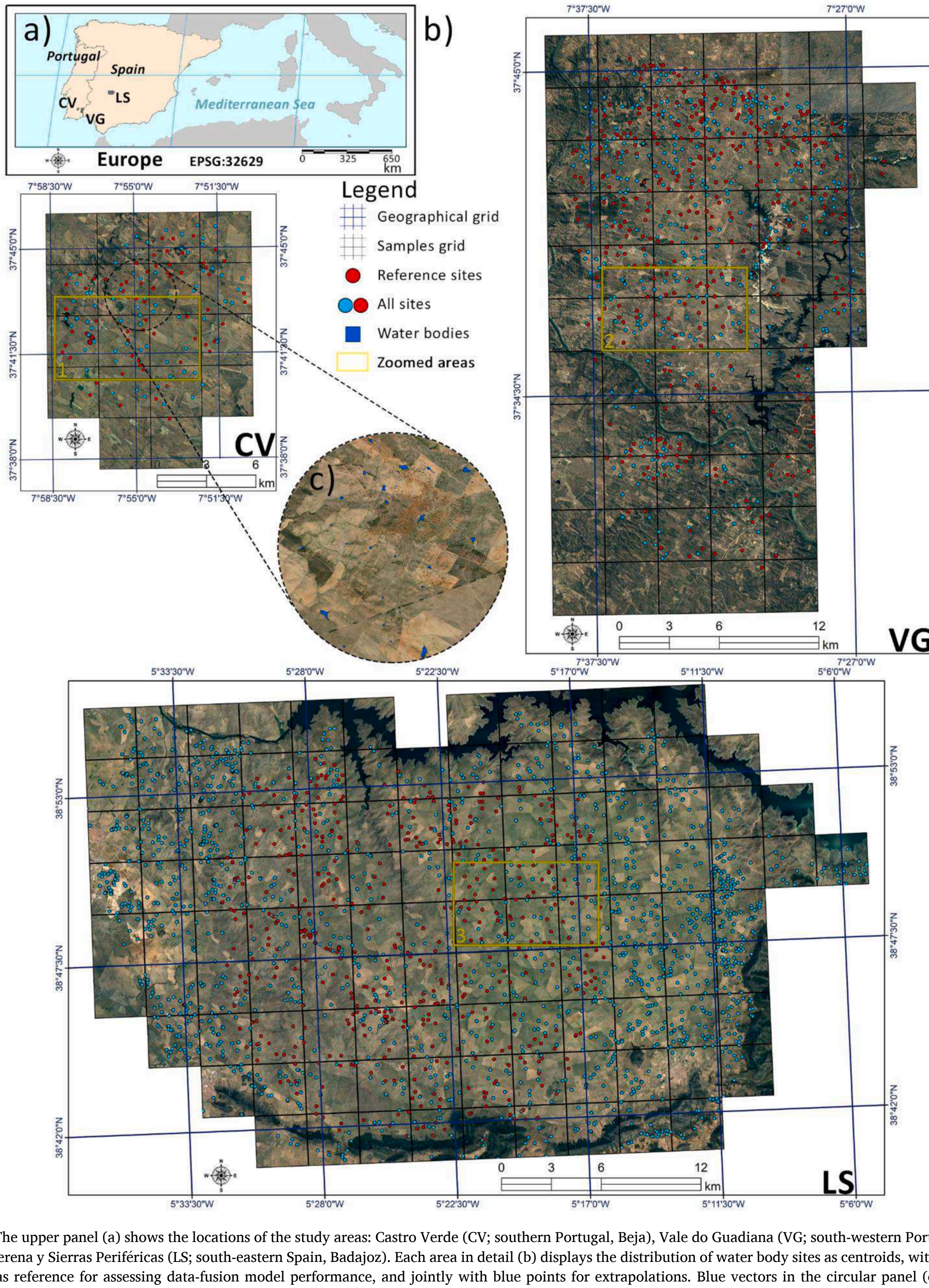

Fig. 1. The upper panel (a) shows the locations of the study areas: Castro Verde (CV; southern Portugal, Beja), Vale do Guadiana (VG; south-western Portugal, Beja), and La Serena y Sierras Periféricas (LS; south-eastern Spain, Badajoz). Each area in detail (b) displays the distribution of water body sites as centroids, with red points serving as reference for assessing data-fusion model performance, and jointly with blue points for extrapolations. Blue vectors in the circular panel (c) represent illustrative water bodies. Each study area is overlaid with an RGB color composite from Sentinel-2 imagery, along with a geographical grid for reference and a sample grid, with 3 × 3 km cells utilized to divide the data into folds. The yellow-highlighted regions (1, 2, and 3) indicate the geographic extents shown in Fig. 6. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

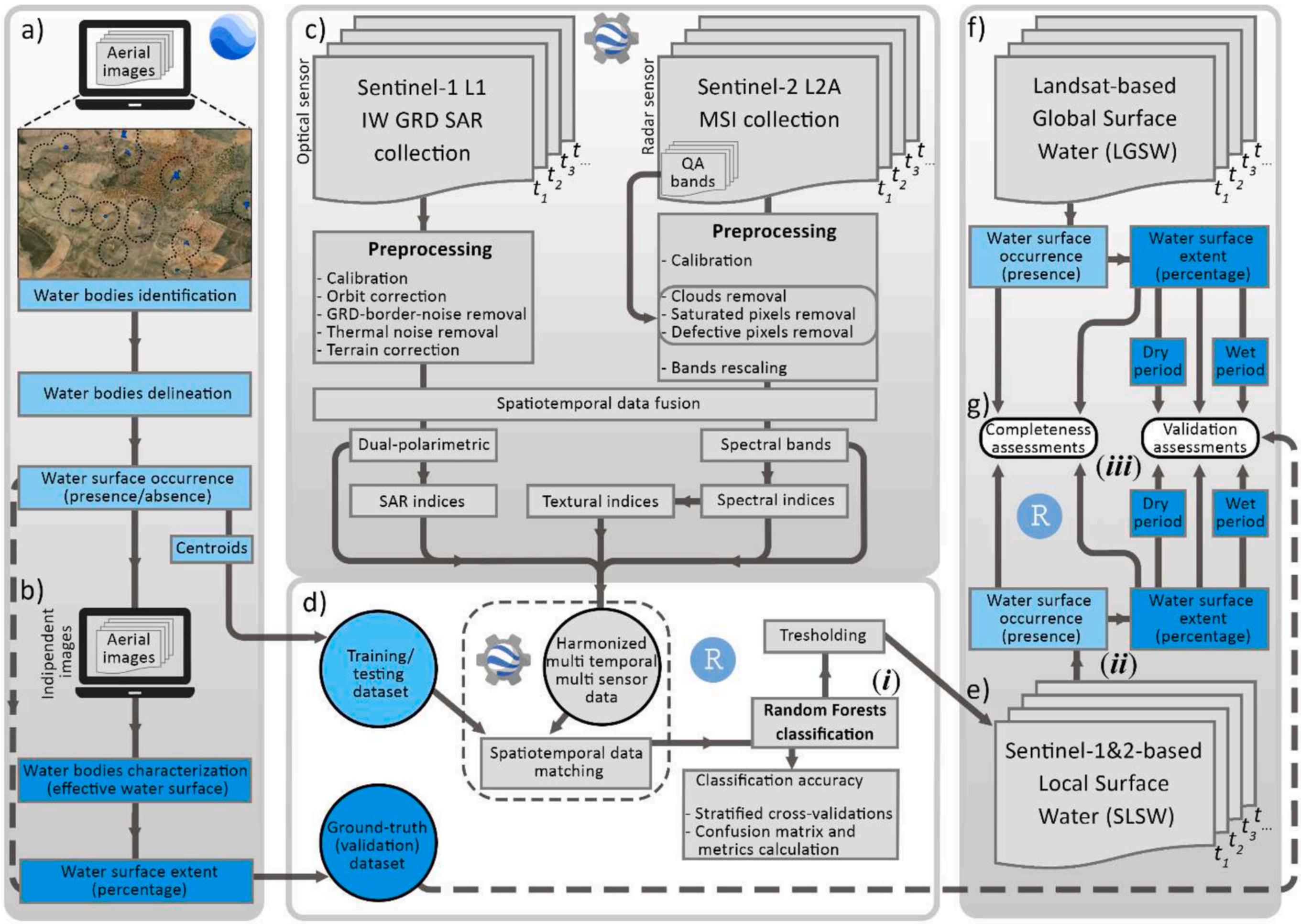

Fig. 2. Framework diagram depicts the analytical steps across four main panels, where arrows represent the operational progressions. The left panel illustrates how water bodies' datasets were prepared, namely surface water occurrence (a) and extent (b), whilst the central-upper panel (c) shows how satellite (Sentinel-1&2) image time series were pre-processed and how variables were calculated. The central-lower panel (d) indicates how data were compiled for subsequent analyses to respond to the first objective (*i*). The right panel depicts the second (*ii*) and third (*iii*) objectives, following the characterization of the SLSW time series (e) and the comparative assessment with the LGSW time series (f).

of water bodies were conducted within a GIS environment through the photointerpretation of high-resolution imagery (Google Earth software; Google) (Fig. 2a), during both dry (August and September) and wet periods (November and March) from 2017 to 2021 (Table SM2.1). The decision to extend the survey period beyond the specific investigative timeframe was motivated by the necessity to comprehensively represent all study areas under consideration. The margins (shorelines) of water bodies were vectorized (Fig. 1) based on the maximum water capacity, determined by observing the separation of evident flood-prone shores from the surrounding landscape. This criterion included the transitional zone (littoral zone) but excluded the vegetative buffer zone. Eligible water bodies included dammed ponds, reservoirs, and natural and semi-natural ponds (temporary, permanent, semi-permanent, and farm ponds), while excavations, pools, and water treatment systems were ignored (Céréghino et al., 2008). A total of 3366 water bodies were vectorized (centroids in Fig. 1), of which 932 were used as a reference dataset (surface water occurrence) for further analysis. This reference dataset was selected using proportional stratified random sampling by parameterizing polygon frequencies within blocks (cells) of 3 × 3 km (Fig. 1) as population '*strata*' (*sensu* Cochran, 1977). This means that the number of water bodies found in each grid block determined the sampling proportion for that block, ensuring that the random sampling was representative of the varying densities of water bodies across the entire study area. Water surface occurrence (presence/absence) was included as an attribute to these polygons, jointly with an ID, and a timestamp (YYYY/MM/DD) reflecting the acquisition date of the imagery used. Additionally, a separate validation dataset was created by characterizing the effective surface water within water bodies, and used to validate the surface water extent model predictions (Fig. 2b; further details in SM1.1).

2.2.1.2. Sentinel-1&2 data pre-processing. The remote sensing data were derived from the Sentinel-1&2 constellations, which are key assets of the Copernicus program administered by the European Space Agency. The Sentinel-1 satellite carries a phase-preserving dual-polarization Synthetic Aperture Radar (SAR) instrument, while the Sentinel-2 satellite carries a multispectral instrument (MSI). These satellites have a swath width of 250 km for Sentinel-1, and 290 km for Sentinel-2. For the analysis, the Level-2A product from the Sentinel-2A and Sentinel-2B generations was used, while, for Sentinel-1A and Sentinel-1B, the Level-1 Ground Range Detected (GRD) product was selected, with the Interferometric Wide (IW) swath as a baseline acquisition modality for both ascending and descending orbits. Both collections underwent several pre-processing steps, including the use of quality assessment (QA) bands (Fig. 2c) to filter out biased pixels; further details are provided in SM1.2. Subsequently, to harmonize the datasets from Sentinel-1&2 satellites, a spatiotemporal data fusion (Fig. 2c) was applied, which involved aligning the two collections both temporally (co-registration;

Chen and Zhao, 2022) and spatially (i.e., resample), resulting into routinely compositing the imageries into 15-day time intervals. This temporal window leverages each sensor's approximate 6-day revisit frequency, ensuring at least one image per sensor in each observation period. The Sentinel-1&2 imageries ingestion and pre-processing analyses were conducted using Google Earth Engine (GEE) (Fig. 2c), a cloud computing platform capable of communicating with different repositories and rapidly process vast amounts of geospatial information (Gorelick et al.,2017).

2.2.2. Environmental factors calculation

To assess the ability of metrics to flag the occurrence of surface water (objective *i*), bands from Sentinel-1&2 were included (Fig. 2c). From the Sentinel-2 reflectance bands, those with a native spatial resolution of 60 m (cirrus band, coastal aerosol, and water vapour) were discarded, resulting in 10 high-resolution bands at 10 m and 20 m, in turn downscaled using default nearest neighbor resampling (Gorelick et al., 2017). These bands covered the visible spectrum (blue, green, red), near-infrared spectrum (NIR1&2, Red-edge1&2&3), and short-wave-infrared spectrum (SWIR1&2) (Drusch et al., 2012). The Sentinel-1 data, originally projected at a resolution of 10 m, was represented by the C-band SAR with dual-polarimetric channels: co-polarized vertical transmit and vertical receive (VV), and cross-polarized vertical transmit and horizontal receive (VH) (Attema et al., 2008). In addition to spectral and SAR information, several water-related indices were considered (Table 1). Among them, the Normalized Difference Vegetation Index (NDVI; Rouse et al., 1974) was selected to further calculate textural indices using the Gray-Level Co-occurrence Matrix (GLCM; Haralick et al., 1973) to address complex surface conditions often found in small water bodies (Bangira et al., 2019). These indices were computed to enhance surface water detection and provide information on landscape biophysical attributes, such as soil and vegetation moisture (Table 1) (Bangira et al., 2019; Huang et al., 2018).

2.2.3. Surface water occurrence and extent analyses

Prior to analyses, the dataset was converted into centroids (Fig. 2a). The ID and timestamp fields from the centroids served as input data for various sequential operations in the web-based tool *GEE_xtract* (Valerio et al., 2024) (Fig. 2d), ensuring the extraction of environmental data with a high spatial match, while providing reasonable temporal alignment in the time series (15-days frequency). To determine water presence (occurrence) within polygons, a binary classification was selected to discriminate between water and non-water classes (response variable), with environmental factors used as explanatory variables. To develop the SLSW (Sentinel-based Local Surface Water) model, a Random Forests (RF; Breiman, 2001) algorithm was used as a supervised machine learning method (Fig. 2d), given its robustness in extracting information about water bodies from space-based remote sensing data (Bangira et al., 2019; Jiang et al., 2022; Peña-Luque et al., 2021). Furthermore, while RF is capable of handling high-dimensional information (Jiang et al., 2022), environmental variables with a Pearson's correlation coefficient $|r|<9$ (Millard and Richardson, 2015) were retained. Hence, our initial dataset resulted into 14 eligible variables for further analysis (Figure SM2.2). The RF algorithm was then configured with a node-splitting parameter equal to the square root of the total of selected variables, and a total of 2000 trees were utilized (Valerio et al., 2020). For the analyses, the reference dataset of 932 water body observations (~28 % of the total dataset) was used. The multivariate RF

Table 1
The portfolio of spectral and SAR bands with derived indices initially included for detecting surface water in small water bodies.

Satellite/sensor	Type	Subtype	Description	Band/equation	Reference
Sentinel-2A/ Sentinel-2B	Spectral bands	Reflectance bands	Blue, Green, Red, Red Edge 1&2&3, NIR1&2, SWIR1&2	B2, B3, B4, B5, B6, B7, B8, B8a, B11, B12	Drusch et al., 2012
	Spectral indices	Water-related indices	Normalized Difference Moisture Index, using NIR1 (NDMI_N1) or NIR2 (NDMI_N2) band	$\frac{(B8-B11)}{(B8+B11)}$	Wilson and Sader, 2002
			Normalised Difference Water Index 2, using NIR1 (NDWI2_N1) or NIR2 (NDWI2_N2) band	$\frac{(B8-B12)}{(B8+B12)}$	Gao, 1996
			Sentinel-2 water index (SWI)	$\frac{(B5-B11)}{(B5+B11)}$	Jiang et al., 2020
		Vegetation-related indices	Normalized Difference Vegetation Index, using NIR1 (NDVI_N1) or NIR2 (NDVI_N2) band	$\frac{(B8-B4)}{(B8+B4)}$	Rouse et al., 1974
			Modified Soil-adjusted Vegetation Index 2, using NIR1 (MSAVI2_N1) or NIR2 (MSAVI2_N2) band	$\frac{(2{*}B8+1-\sqrt{(2{*}B8+1)^2-8{*}(B8-B4)}}{2}$	Richardson and Wiegand, 1977
			Normalized Difference Index 45 (NDI45)	$\frac{(B5-B4)}{(B5+B4)}$	Delegido et al., 2011
	Textural indices	Second-order	Contrast, Variation, Homogeneity, Mean (GLCM-C, GLCM-V, GLCM-H, GLCM-M)	Calculated using NDVI with a 3×3 pixels window in multi-directions (0°, 45°, 90°, and 135°)	Haralick et al., 1973
Sentinel-1A/ Sentinel-1B	Dual-polarimetric descriptors	Co-polarization	Vertical Transmit Vertical Receive	VV	Attema et al., 2008
		Cross-polarization	Vertical Transmit Horizontal Receive	VH	Attema et al., 2008
	SAR indices	Polarization ratios	Cross Ratio (VHVVR or VH/VV)	$\frac{(VH)}{(VV)}$	Alvarez-Mozos et al., 2021
			Normalized Difference Polarization Index (NDPI)	$\frac{(VV-VH)}{(VV+VH)}$	Cao et al., 2008
			Dual-Polarized Radar Vegetation Index VV (RVI)	$\frac{(4{*}VH)}{(VV+VH)}$	Nasirzadehdizaji et al., 2019
			Normalized VH Index (NVHI)	$\frac{(VH)}{(VV+VH)}$	McNairn and Brisco, 2004
			Normalized VV Index (NVVI)	$\frac{(VV)}{(VV+VH)}$	McNairn and Brisco, 2004

models were assessed through cross-validations (Fig. 2d). The data, split into 70 % for training and 30 % for testing, were stratified into 5 k-folds based on spatial blocks of 3 × 3 km (details in Figure SM2.3) to minimize autocorrelation problems and maximize the model's extrapolation power (Valavi et al., 2018). From developed confusion matrices, five accuracy metrics were selected: Area Under the Curve (AUC; Swets, 1988), sensitivity, specificity, accuracy, F1 score, and Matthews' Correlation Coefficient (MCC; Baldi et al., 2000). The importance of variables for classification (water and non-water classes) and for surface water occurrence detection was determined using the Gini index (Breiman, 2001), which measures the effect of each variable in explaining surface water occurrence. The RF multivariate analyses were repeated along the previously selected time series (48 time intervals × 5-fold cross-validations) to infer surface water occurrence probability, at each time interval with retained variables (14), for a total of 672 variables (14 × 48). The continuous classification probability was then subjected to a thresholding procedure to predict 'water' and 'non-water' pixels within the original polygons representing maximum water capacity. The optimal threshold (p = 0.51) was determined using five combined techniques (further details in SM1.3). The surface water occurrence was computed within polygons containing at least one pixel flagged as with water, and then compared proportionally to those lacking any water pixels.

To measure the surface water extent, the cumulative area from the amount of predicted surface water occurrence was calculated within each polygon and compared to the original polygon area, at each time interval. The previously described steps were repeated along the time series, which enabled the investigation of the temporal trends in surface water occurrence and extent of the SLSW model (objective *ii*; Fig. 2e).

2.2.4. Reliability assessment of the SLSW model

The reliability of our modeling approach (SLSW) was tested through a high-quality external model, the JRC Monthly Water History (v1.4) developed by Pekel et al. (2016). Here, a monthly time series coincident with the SLSW period was selected and referred to as LGSW (Landsat-based Global Surface Water; Fig. 2f). The LGSW underwent the same processing steps as SLSW to enable the comparison between models in terms of archive completeness (objective *iii*; Fig. 2g). For each model, the completeness threshold was defined for each timestamp in the time series where median values were above 0 %. The SLSW time series for water surface occurrence and extent were further scrutinized through the comparison with those derived by the E-OBS gridded meteorological parameters (Cornes et al., 2018), focusing on parameters such as mean and maximum temperature, precipitation amount, relative humidity, and surface shortwave downwelling radiation. Additionally, it was investigated how water surface occurrence and extent varied according to the size of water bodies. To accomplish this, the area (ha) of the identified water bodies was categorized into 5 classes, each reflecting a maximum number of associated water pixels, respectively 2, 3, 5, 10, and 100.

Second, the agreement in surface water extent (objective *iii*; Fig. 2g) was assessed through a linear and quadratic regression modeling between SLSW (dependent variable) and LGSW (independent variable) models, under wet (January-April 2021) and dry (June-September 2021) environmental conditions. This statistical procedure was separately repeated by using the independent validation dataset (Fig. 2d), and spatiotemporally coincident predictions (May 2022) from SLSW and LGSW models. For each analysis, the distribution values were compared as density plots, and linear and quadratic regressions were performed. The coefficient of determination (R^2) was used as a measure of relatedness (Peña-Luque et al., 2021), and the slope (B1) was calculated using standardized major axis regression. The statistical analysis and data synthesis were conducted in the R environment (v.4.2.0; R Core Team, 2021). Specifically, data stratification was performed using the '*blockCV*' (v.2.1.4; Valavi et al., 2018) package, while the Random Forests analyses through the '*randomForest*' (v.4.7–1; Liaw and Wiener, 2002) and '*caret*' (v.6.0–90; Kuhn, 2015) packages.

3. Results

3.1. SLSW model outcomes

The analysis of surface water occurrence performance for the SLSW model indicated a low Out-Of-Bag (OOB) error rate (Table SM2.3), with a similar score ($\mu \approx 26$ %) between the water and non-water classifiers. The Specificity metric demonstrated better performance than Sensitivity, and a good level of agreement for Accuracy, AUC, MCC and F1 metrics (Table SM2.3). The most influential variables were the vegetation index NDI45, the SWIR1, BLUE, and NIR1 reflectance bands, followed by the second-order textural index GLCM-C and the SAR index polarization ratio VHVVR (Figure SM2.4).

The probability of surface water occurrence improved with increasing NDI45 index values (Fig. 3). Conversely, for SWIR1 and NIR1 reflectance bands, the probability decreased as the reflectance values increased (Fig. 3). For the GLCM-C and VHVVR variables, highly steep curves (respectively positive and negative) were identified, while for VHVVR, a higher probability was observed from increased backscattering values, followed by an abrupt negative response (Fig. 3).

3.2. SLSW intra- and inter-annual trends

In the three regions combined, a clear seasonal pattern emerged for the SLSW model, characterized by pronounced surface water changes in both 2020 (Fig. 4a) and 2021 (Fig. 4b). Specifically, minimum surface water extent conditions (~20 % to 10 %) were observed from late July to early September, followed by an extensive recovery in November in both years (Fig. 4a and 4b). Both 2020 and 2021 exhibited similar seasonal variability, with high agreement during wetter conditions (November-April); yet, a prolonged anomalous period of water deficit conditions was detected in 2020 (Fig. 4a) between September to early October, where slower recovery compared to 2021 was observed (Fig. 4b). These yearly variations in seasonal trends appear to align with observed meteorological patterns (see Figure SM2.6), supporting the idea that SLSW predictions accurately reflect natural fluctuations rather than being artificially produced (Figure SM2.6).

The three regions exhibited distinct seasonal variability in surface water extent (Fig. 5a and 5b). In particular, during dryer conditions (summer), a larger deficit was consistently observed in the time series, and again, a slower recovery in surface water extent during 2020 (Fig. 5a) with respect to 2021 (Fig. 5b).

The SLSW time series revealed a constant discrepancy between classes throughout the wet season (January-April), which became more pronounced during the dry season (June-September) (Figure SM2.9; SM2.10). In particular, from mid-September to mid-October of 2020, anomalous extended periods of reduced surface water extent were evidenced compared to 2021 (Figure SM2.9), coinciding with higher temperature (mean and maximum), lower humidity and precipitation amount, and markedly lower values in surface shortwave downwelling radiation (Figure SM2.6).

3.3. SLSW and LGSW model comparisons

3.3.1. Archives completeness

For the three regions separately, complete archives in surface water extent emerged for the SLSW model for both 2020 (Fig. 5a) and 2021 (Fig. 5b), contrary to LGSW (Fig. 5c and 5d). Considering the three areas concomitantly, the LGSW time series also reported a lack of completeness (median values equal to 0), corresponding to an underestimation magnitude of approximately 3 times ($\mu \approx 70.5$ %) than SLSW for both 2020 (Fig. 4c) and 2021 (Fig. 4d). Diversely, for the SLSW model, the time series consistently produced median surface water extent values above 0 % (Fig. 4a and 4b). Furthermore, the surface water occurrence

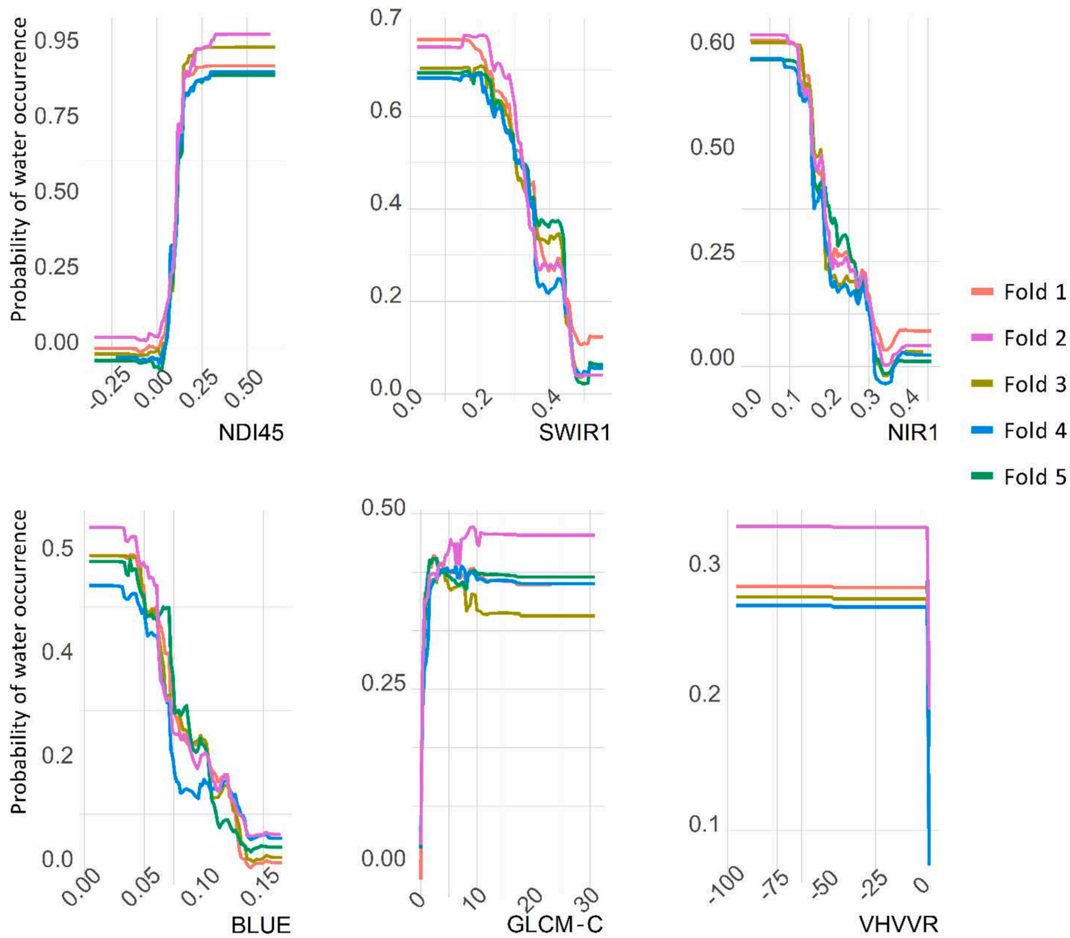

Fig. 3. Partial dependence plots of most important variables explaining surface water occurrence. For each variable, multiple curves are depicted, with each color representing a fold of the stratified RF cross-validations. Additional plots of the remaining variables can be found in Supplementary Materials (Figure SM2.5).

mapping process showed detailed information across the three regions (Fig. 6a). As expected, a higher representativeness of surface water occurrence pixels emerged using SLSW rather than LGSW, in specific during drier periods (e.g., August) compared to wetter periods (e.g., January) (Fig. 6b). The cumulative surface water occurrence along the SLSW time series is also depicted, highlighting pixels subject to higher and lower variability (Fig. 6c).

Regarding surface water occurrence, the results reflected those of surface water extent, indicating lower completeness for LGSW than SLSW models considering all the three areas concomitantly (Figure SM2.7), as well separately (Figure SM2.8, Figure SM2.11).

3.3.2. Comparison between methods and validation dataset

When comparing regression models using SLSW as a function of LGSW for surface water extent, higher correlations (relatedness) were observed during the wet period (Fig. 7a) compared to the dry period (Fig. 7b). When SLSW and LGSW were separately compared with the validation data, SLSW exhibited satisfactory fit and performed well according to the three metrics, with similar results between linear and quadratic regressions (Fig. 7c). Conversely, the LGSW showed poor regression correlations (Fig. 7d). Additionally, the median difference from the density plots between LGSW and the validation data was markedly high, resulting in ~45 % (Fig. 7d), and evidencing an underestimation in surface water extent stronger than the overestimation observed with SLSW (~25 %) (Fig. 7c).

4. Discussion

We propose an innovative multi-temporal multi-sensor data fusion method enhancing the delineation and mapping of small seasonal water bodies. These are often neglected and challenging to monitor, resulting in frequent archive discontinuities. Results shed light on the prospective benefits of combining Sentinel-1 and Sentinel-2 data for surface water detection and quantification within small and complex water bodies, and the added value of time series for inferring intra- and inter-year surface water trends in different areas. The reliability of our findings was supported by comparing local and global products in terms of prediction completeness and agreement during wet and dry periods, and validated using an independent dataset.

4.1. Surface water occurrence inference

The NDI45 index emerged as the most influential variable for explaining the potential occurrence of surface water in our study areas. This is probably due to its strong relationship with chlorophyll content and its high sensitivity to vegetation biophysical properties, particularly in the red edge wavelength interval (Jiang et al., 2022; Delegido et al., 2011). In contrast, previous research on water classification found lower NDI45 importance scores and higher scores for water-related indices such as NDWI and MNDWI, which are possibly more suitable for identifying larger water areas (limnetic zone) (Jiang et al., 2022; Radoux et al., 2016) and clearer waters (Sun et al., 2012). The diminished relative importance of NDWI may also suggest some limitations of traditional water-related indices and associated methods (e.g., NDWI-

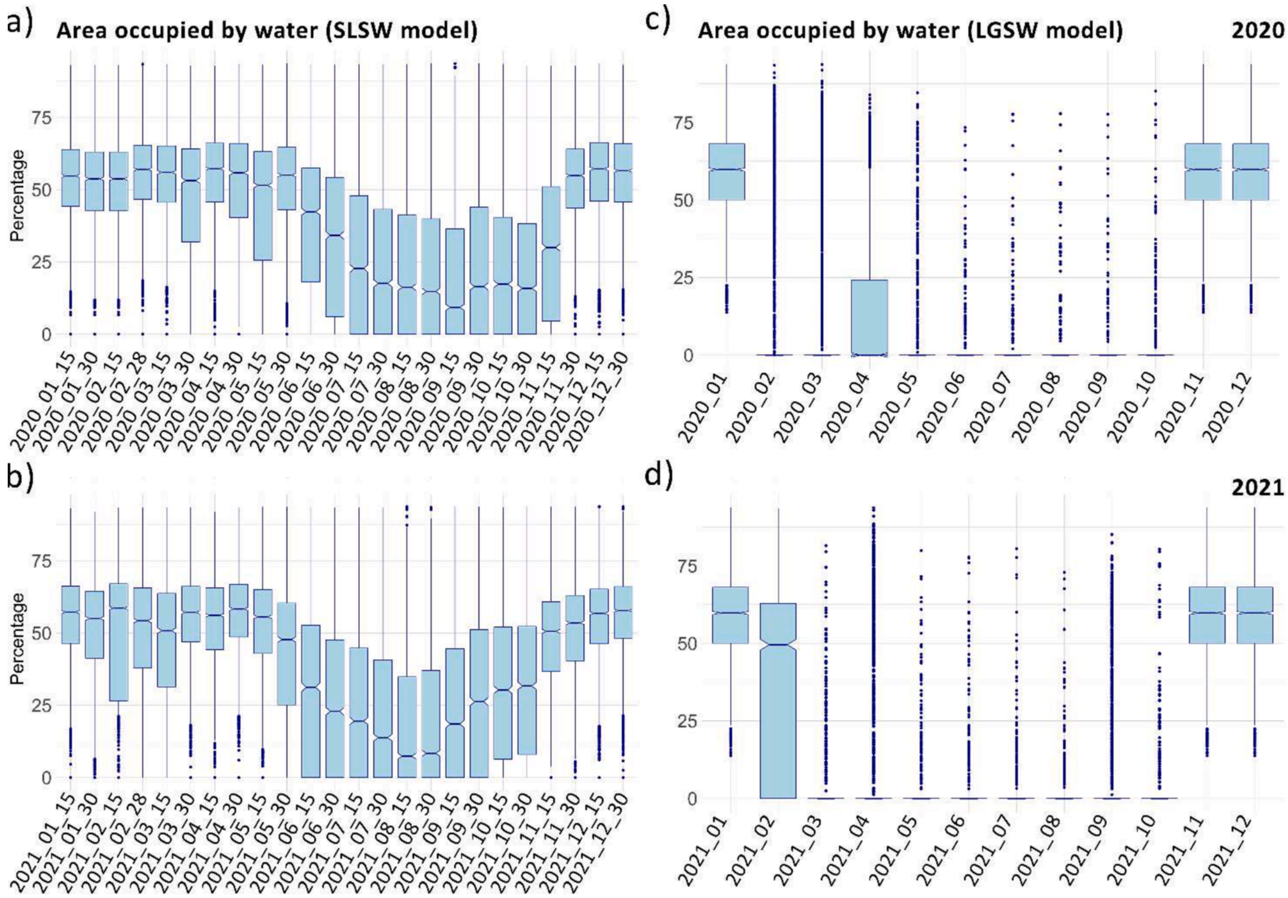

Fig. 4. Surface water extent, considering the three semiarid regions collectively. The time series represent the SLSW model for both 2020 (a) and 2021 (b), and the LGSW model for both 2020 (c) and 2021 (d), with a time interval of 15 days and 30 days, respectively. For each date, the range of values illustrates the water extent percentage within water bodies, with dots depicting the outliers.

based thresholding; Zhou et al., 2017) in specific contexts, such as when targeting small-sized water bodies. The importance of NDI45 relies on its ability to discriminate between different types of vegetation, including both terrestrial and aquatic. For the aquatic vegetation, minimum healthy conditions have been observed at values ≈0.1 in a nearby geographic region in northeastern Spain (Soria et al., 2022), which aligns with our response curve inflexion point. The NIR and SWIR bands were also important as they explicitly detect water signals given the considerable absorption of electromagnetic radiation in the invisible wavelength area, whether non-water landscape attributes typically exhibit higher reflectance (Sun et al., 2012). Interestingly, a band from the visible spectrum, the BLUE band, also demonstrated high relevance, as indicated by the negative response curve, suggesting its association with water-related light absorbability (Jiang et al., 2022; Sun et al., 2012). The horizontal structural GLMC-C variable showed a rapid increase in the probability of surface water occurrence in areas with highly contrasting photosynthetic activity (i.e., central water areas and vegetative buffer zones). This metric successfully captured the landscape ecological properties of water bodies, which may be beneficial in classification, extraction, and segmentation workflows, as well as for accurate mapping of water bodies in uncharted regions (Bioresita et al., 2019; Peña-Luque et al., 2021).

A further key metric explaining water surface occurrence was the cross-polarization SAR metric VHVVR (also known as VH/VV, or VHrVV). The backscattering response is influenced by the synergistic effects of the structure (e.g., volume scatters from vegetation) and the dielectric field of the surfaces, determining the resulting amount of energy received by the sensor (Vreugdenhil et al., 2020). A high probability of surface water occurrence was found with increased backscattered energy, ranging from highly smooth surfaces, indicated by lower backscatter coefficients, to more irregular surfaces (Tang et al., 2022). However, it is important to note that this range of values prior to the inflexion point, roughly around 0, may also include signals from wet soil and scattered vegetation, potentially including the detection of more irregularly inundated areas (Huang et al., 2018; Petropoulos et al., 2015). Interestingly, the VHVVR may hold the potential for conservation studies exclusively targeting temporary ponds, which can be more complex in terms of variant conditions (such as habitat, persistence, and pond bed; Sebastián-González and Green, 2014) compared to permanent water bodies.

4.2. SLSW and LGSW accuracy

4.2.1. Intra and inter-annual surface water trends

The time series comparison demonstrated a better completeness from SLSW in both intra- and inter-annual surface water trends (occurrence and surface water extent), attributable to LGSW archive discontinuities (Pekel et al., 2016). It is relevant to emphasize that during the wet season, that is, when cloud cover is highest, specific dates were identified with missing information in the SLSW time series, although this bias was deemed negligible. Hence, the selected 15-day time windows are a reasonable tradeoff for aggregating information, as found in previous research (Peña-Luque et al., 2021).

The SLSW corroborated high seasonal variability within the three areas, with the Natural Park of Vale do Guadiana (VG) exhibiting the lowest variation, possibly attributed to the local presence of deep-rooted vegetation, likely enhancing infiltration and groundwater recharge. In the other areas (LS and CV), by contrast, higher surface water losses

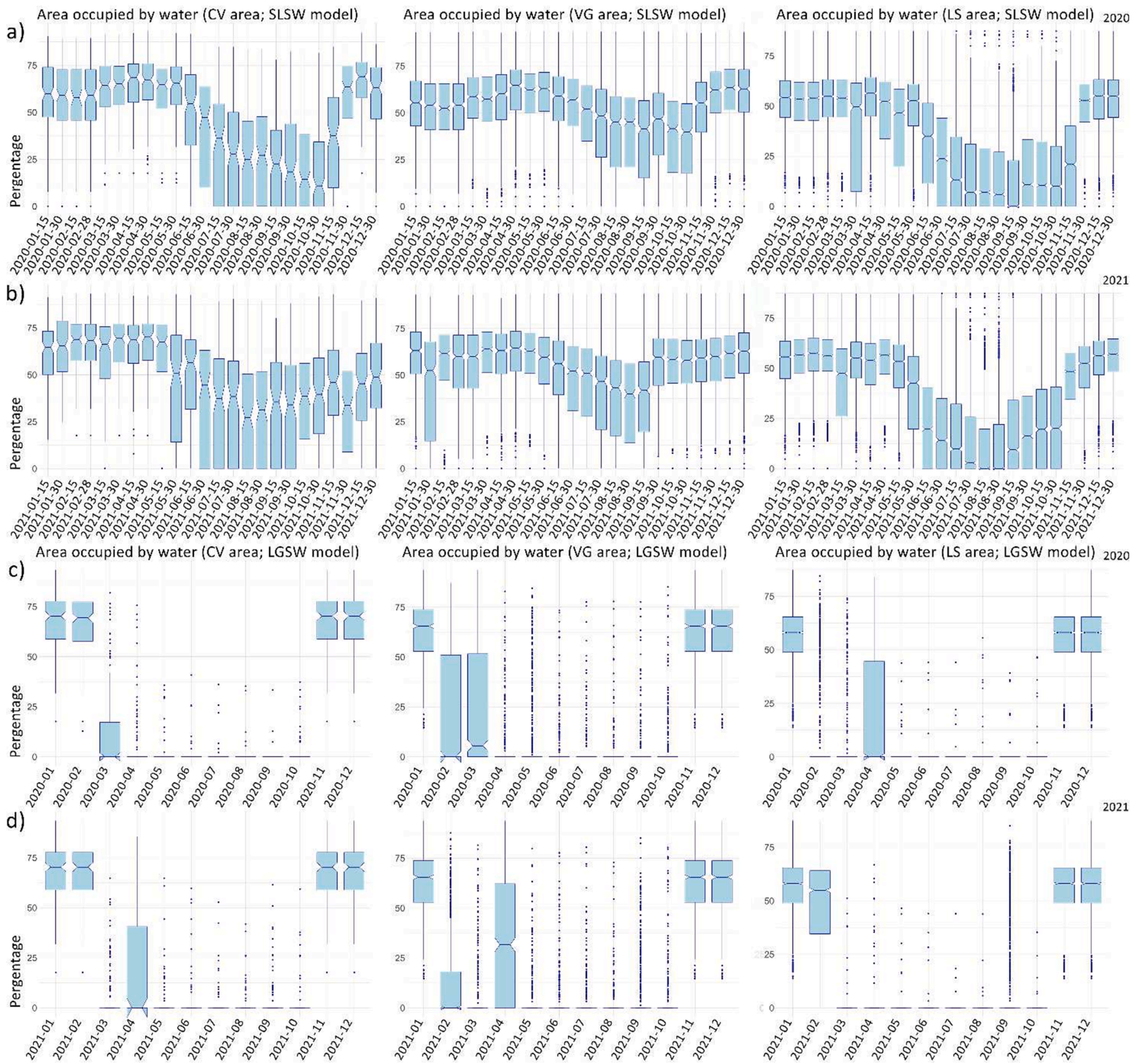

Fig. 5. Surface water extent, considering the three semiarid regions (CV, VG, and LS; Fig. 1) separately. The time series represent the SLSW model for both 2020 (a) and 2021 (b), and the LGSW model for both 2020 (c) and 2021 (d), with a time interval of 15 days and 30 days, respectively. For each date, the range of values illustrates the water extent percentage in water bodies, with dots depicting the outliers.

during summer can be explained by higher hydroclimatic variability influenced by temperature and energy stocks, and typically affecting smaller water bodies (Rachid et al., 2022). Our findings revealed a recurrent pattern in surface water fluctuations, suggesting periodic changes in pond levels, with trends reflecting the period of water bodies recovering, from autumn to early winter (Gómez-Rodríguez et al., 2009). Further, the SLSW analysis indicated that 2020 experienced drier conditions than 2021, concomitantly with E-OBS data showing higher temperatures, lower humidity, and precipitation rates in the former year. Notably, the SLSW time series revealed fine-scale differences between years, uncovering water deficit anomalies most evident during the transition from the late dry season to the early wet season, attributable to prolonged dry conditions. The intra-year variation during this specific period was further evidenced by the E-OBS trends, particularly concerning the substantial amplitude rates found in surface shortwave downwelling radiation. Although additional research is required to fully elucidate the relationship between these drivers and the observed trends, the observed increased amplitude suggests a marked rise in ground radiation absorbability during 2020 (Rachid et al., 2022). These prolonged drier conditions reflect drought events, which are conventionally detected via coarser sensors (Zou et al., 2018). Climatic anomalies are suspected to increase in these regions, with rising temperatures and drought events posing socio-economic concerns, and increasing attention towards sustainable water resource management (Samaniego et al., 2018).

4.2.2. Surface water inference reliability

In both dry and wet seasons, there was minimal agreement between SLSW and LGSW in terms of the extent of identified surface water. During dry periods, the surface water extent decreased dramatically,

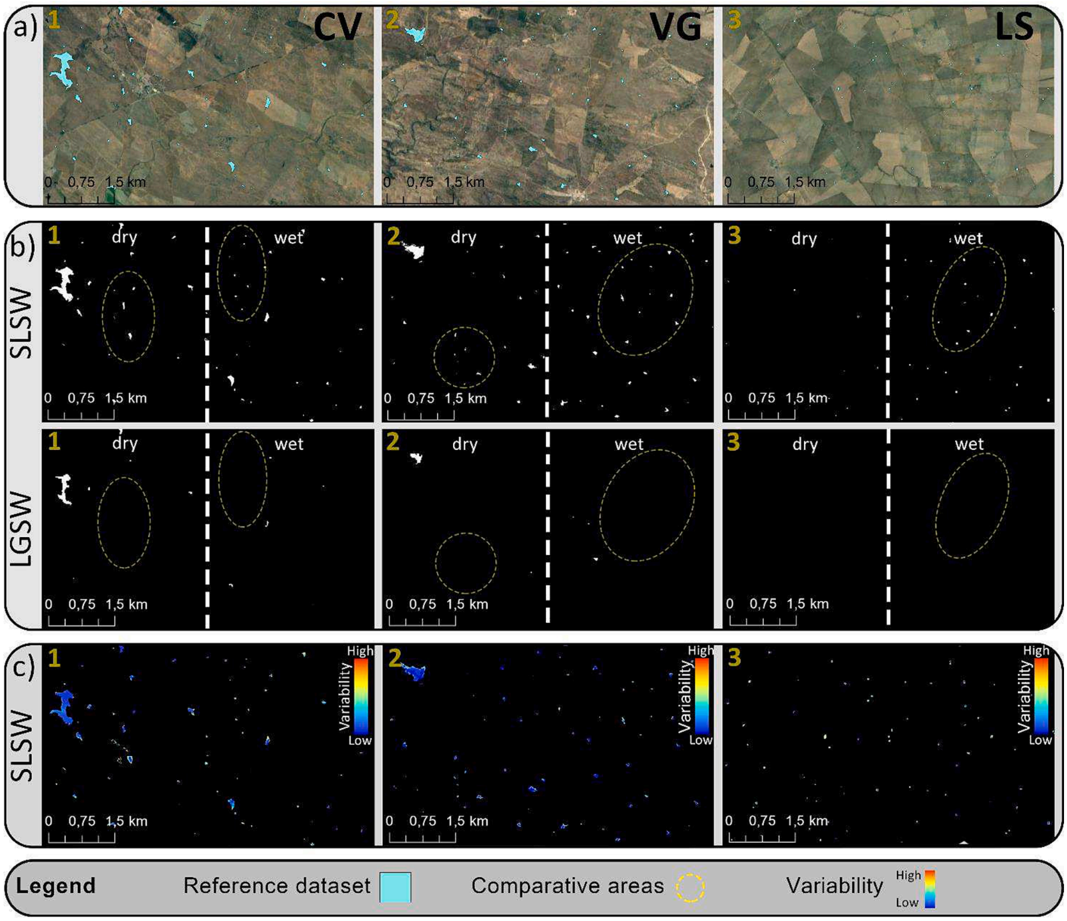

Fig. 6. Surface water distribution in the magnified areas (a), where numbers indicate their representative regions depicted in Fig. 1. The magnified areas were comparable concerning surface water occurrence predictions (in white) by SLSW (upper panels) and LGSW (middle panels) (b), jointly with SLSW-derived variability (lower panels) (c). Upper and middle panels were bisected by a vertical (white) dashed line, to represent dry (August) and wet conditions (January). Ellipses and circles are also depicted in magnified areas highlighting clustered water bodies, to emphasize prediction comparisons between SLSW and LGSW.

hindering the relatedness between SLSW and LGSW, with the latter being less accurate and underestimating surface water. In fact, the relatively limited spatial resolution (30 m) of the Landsat-derived LGSW model is a recognized constraint that warrants adjustment in future EO-based methodologies (Pekel et al., 2016). This supports the advantages of using Sentinel-1&2 data-fusion approaches when targeting very small water bodies in open and seasonal waters, especially in semiarid and water-stressed areas, where frequent water shortages may interfere with sensor detection (Peña-Luque et al., 2021).

The limited ability of LGSW to accurately infer surface water extent in small ponds was further corroborated with the validation data, resulting in very poor agreement and significant underestimation of surface water extent. Furthermore, the higher spectral resolution associated with the Sentinel-derived SLSW model likely enhances its ability to characterize complex water bodies with diverse characteristics (Bangira et al., 2019). On the other hand, it should be noted that while the SLSW exhibited a considerably better performance than LGSW, a slight overestimation tendency emerged in surface water extent values compared to validation data. In this direction, to handle overestimations when extrapolating surface water occurrence in new areas, it might be beneficial the incorporation of topographic predictors (e.g., topographic wetness index; Huang et al., 2018; Tang et al., 2022; Vanderhoof et al. 2023).

4.3. Limitations and opportunities

4.3.1. SLSW limitations

When targeting permanent waters in relatively large reservoirs, high accuracy scores, ranging from 70 % to over 90 %, are often achieved (Bioresita et al., 2019; Jiang et al., 2022; Zhou et al., 2017). Similar outcomes have also been reported for seasonal water bodies (Bangira et al., 2019; Huang et al., 2018). More specifically, recent studies reported high accuracy for detecting surface water in small bodies ranging from 0.5 ha (Wang et al., 2022) to up to 5 ha (DeLancey et al., 2022; Peña-Luque et al., 2021). However, research on very small water bodies ($<$0.5 ha) has been limited, and studies focusing on this size range have shown suboptimal performance when using publicly available EO satellite data (e.g., Cordeiro et al., 2021). Within this range, our review of the literature did not find studies that inventoried and catalogued water bodies comparable to our study ($Mdn \approx 0.031$ ha), while also investigating the dynamics of seasonal and complex waters. Despite some limitations (i.e., sensor resolution, algorithm sensitivity, or technical issues like pond destruction or dislocation), the metric scores, the produced high-frequency time series, and the validation results are generally satisfactory given the complexity of detecting dynamic water surfaces (Bangira et al., 2019; Wang et al., 2022).

To soften classification biases more effectively, Sun et al. (2012) suggested separating water bodies into three classes: clear, green, and turbid water. Also, errors in surface water occurrence analyses might be due to mixed pixels, which cannot be ignored as they frequently provide the sole information for describing environmental characteristics. In this

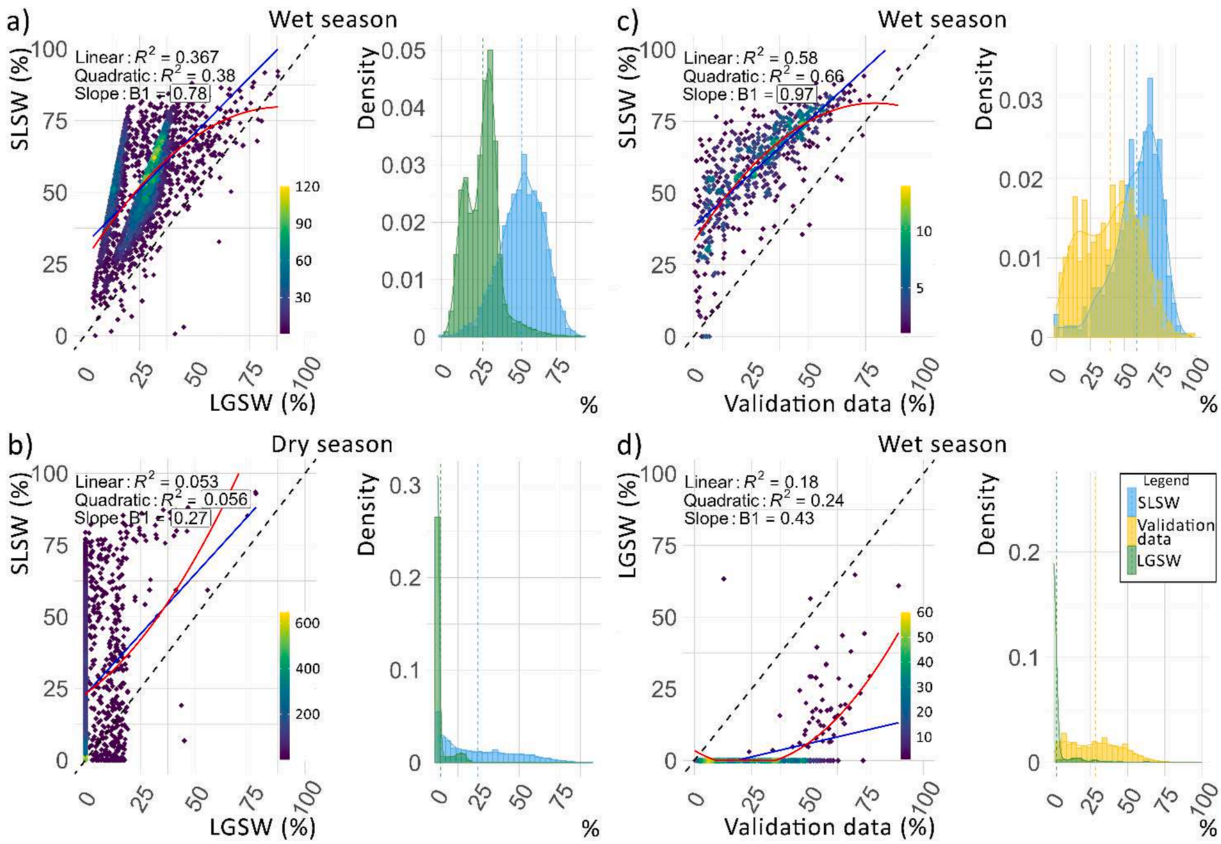

Fig. 7. Density scatterplots are depicted with the linear (blue line) and quadratic (red line) regressions between SLSW and LGSW models concerning surface water extent percentage under both wet (a) and dry conditions (b). Furthermore, independent comparisons of SLSW and LGSW models with the validation dataset are presented in separate panels (c and d). In all panels, each graph point is a representation of a distinct water body. Density plots are coupled with each regression, where dashed lines represent the median value of estimated surface water extent. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

context, the literature widely acknowledges the efficacy of employing sub-pixel approaches (Radoux et al., 2016; Ozesmi and Bauer, 2002). However, caution is needed as multiple distinct classifications to derive spectral signature can significantly increase computation time, which can become unfeasible when reproducing high-frequency time series at high spatial detail. Given the heterogeneity and varying sizes of ponds, it is generally prudent to predefine pond boundaries, though this approach may pose significant limitations when applied at continental scales. Consequently, as an alternative to reduce the time required for data preparation, boundaries could be delineated for a representative subset of detectable ponds for model training, and the results subsequently extrapolated across the entire study area.

4.3.2. *Applications in Ecology*

Water availability is critical for the survival of plant and wildlife populations and the development of human activities, particularly in water-stressed and drought-prone regions. The present framework can be of interest for a range of semiarid research and application fields. It can contribute to the conservation of small wetland areas, as recommended by initiatives such as the Ramsar Convention (De Meester et al., 2005). The predominance of Sentinel-2 data in explaining water surface occurrence reveals that the inclusion of Sentinel-1 data may provide gains in open water scenarios, and that can be advantageous when targeting specific ponds (i.e., temporary ponds). This information is valuable for the protection and restoration of temporary ponds, supported by long-term environmental monitoring using public sensors (Céréghino et al., 2008; De Meester et al., 2005). This is important considering the role of small-sized ponds as habitat for many conservation-priority species (De Meester et al., 2005), especially since these ponds are often poorly managed and severely degraded (Díaz-Paniagua et al., 2024; Gleick, 1998). Interestingly, the observation of inter-annual water deficit anomalies can provide insights into emerging threats (see Reid et al., 2019) and their implications, which are otherwise difficult to detect on ponds with marked hydroperiods, on which local biodiversity depends (De Meester et al., 2005). By enabling the detection of open surface water, this framework advances efforts in measuring the spatiotemporal arrangement and effective connectivity of small-sized ponds (Sebastián-González and Green, 2014). The consideration of time series with high frequency also supports the shifting paradigm from a static unrealistic landscape viewpoint to a dynamic landscape connectivity approach (Martensen et al., 2017). This shift may facilitate better planning solutions to guarantee population viability for different taxonomic groups of plants and animals, based on connectivity constraints imposed by water presence and species dispersal potential (Herceg-Szórádi et al., 2023).

4.3.3. *Applications for water sustainable use*

Sustainable water resources management and development require effective methods for water detection and quality assessment. Ponds, in particular, are highly vulnerable to pollution and overexploitation in many regions, which deepens the challenge of achieving an equilibrium between rural competitiveness and ecological sustainability (Reid et al., 2019). In semiarid regions facing climate change and growing populations, adaptive strategies are becoming increasingly critical to protect water bodies (Gašparović and Singh, 2022; Wickens, 1998), and surrounding habitats, such as grasslands (Shabbir et al., 2019). Remote sensing applications, as the one presented here, may be useful to expand the scope of drought impacts and quality assessments within open water bodies across wide areas. Advances in public sensor technology and cloud-computing environments nowadays offer notable benefits in more reliable and cost-efficient assessments of water quality, with indices such as band ratios and more specific ones such as the Normalized Difference Chlorophyll Index (NDCI) potentially enhancing our understanding of dynamic changes and anomalies in aquatic ecosystems (Deutsch et al., 2018; Mishra and Mishra, 2012). Besides, our modelling procedure can support initiatives aimed at restoring hydrological cycles (De Groot et al., 2018; Gleick, 1998), improving renewable water resources in regions lacking significant reservoirs, and promoting the planning of supplementary or alternative water bodies (Erwin, 2009; Morante-Carballo et al., 2022). In semiarid regions, routinely characterized by poor governance, such initiatives can be valuable for achieving policy objectives (i.e., food security and public health), incorporating the spatiotemporal distribution of social impacts arising from water scarcity, which is expected to increase significantly in Europe (Samaniego et al., 2018). This information is vital in improving water resource management protocols and for advancing SDGs (e.g., SDG 6&15) (Bhaduri et al., 2016), fostering negotiations for conflict resolutions between water governance and environmental sustainability involving institutions, stakeholders, policymakers, and local communities (Homobono et al., 2022; Veldkamp et al., 2015).

5. Conclusions

This research addresses critical gaps in water surface mapping, underscoring the benefits of high spatiotemporal resolution of EO satellite imagery for monitoring inland freshwater systems, specifically small water bodies often overlooked in literature. The utilization of Sentinel-based models revealed numerous advantages in semiarid regions subjected to seasonal changes, particularly in quantifying hydrological dynamics through continuous data analysis. The robustness of our findings was supported through accuracy metrics, model comparison with recent state-of-the-art research, and independent data validation. The produced water surface time series proved competitive with global products of similar resolution, namely from Landsat, demonstrating its utility for regional applications. Sentinel-1 and Sentinel-2 data provide substantial improvements in addressing water-related archive inconsistencies, particularly for very small-sized bodies, while effectively characterizing their complex dynamics, including trends comparisons and the detection of potential anomalies. This is made possible by recent technological advancements of optical and radar sensors, offering high-quality and high-resolution information readily accessible through online repositories and processable using cloud computing platforms (e.g., Google Earth Engine). The proposed framework finally supports innovative pathways for advancing land and water monitoring across a plethora of applications, bringing novel solutions in cross-boundary conservation and sustainability strategies.

Author contributions.

Francesco Valerio: Conceptualization, Data curation, Methodology, Software, Formal Analysis, Writing – Original Draft, Writing – Review & Editing, Visualization. Sérgio Godinho: Conceptualization, Methodology, Investigation, Validation, Supervision, Writing – Review & Editing, Visualization. Gonçalo Ferraz: Data curation, Writing – Review & Editing. Ricardo Pita: Supervision, Writing – Review & Editing. João Gameiro: Supervision, Writing – Review & Editing. Bruno Silva: Software, Validation, Writing – Review & Editing. Ana Teresa Marques: Methodology, Supervision, Writing – Review & Editing, Visualization. João Paulo Silva: Conceptualization, Data curation, Methodology, Writing – Review & Editing, Funding acquisition, Project administration.

CRediT authorship contribution statement

Francesco Valerio: Writing – review & editing, Writing – original draft, Visualization, Software, Methodology, Formal analysis, Data curation, Conceptualization. **Sérgio Godinho:** Writing – review & editing, Visualization, Validation, Supervision, Methodology, Investigation, Conceptualization. **Gonçalo Ferraz:** Writing – review & editing, Data curation. **Ricardo Pita:** Writing – review & editing, Supervision. **João Gameiro:** Writing – review & editing, Supervision. **Bruno Silva:** Writing – review & editing, Validation, Software. **Ana Teresa Marques:** Writing – review & editing, Visualization, Supervision, Methodology. **João Paulo Silva:** Writing – review & editing, Project administration, Methodology, Funding acquisition, Data curation, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We would like to thank Maria de Jesús Palacios and Angel Sanchez from Dirección General de Sostenibilidad / Consejería de Agricultura de la Junta de Extremadura for all their support and encouragement to carry out this work. The work is supported by National Funds through FCT-Fundação para a Ciência e a Tecnologia in the scope of the project UIDP/50027/2020. This work, including the scholarships of FV and ATM, was co-funded by the project NORTE-01-0246-FEDER-000063, supported by Norte Portugal Regional Operational Programme (NORTE2020), under the PORTUGAL 2020 Partnership Agreement, through the European Regional Development Fund (ERDF), in alignment with the European Union's cohesion policy. FCT funded also BS through the grant SFRH/BD/137803/2018 and GF through the grant 2021.04942.BD (https://doi.org/10.54499/2021.04942.BD), while RP through the contract 2022.02878.CEECIND, and JPS through the contract DL57/2019/CP 1440/CT 0021. JG was supported by a post-doc scholarship (BIOPOLIS 2022-13). SG was funded by the FUEL-SAT project (PCIF/GRF/0116/2019) from FCT, and by National Funds through FCT under the project UIDB/05183/2020. Work supported by the European Union's Horizon 2020 Research and Innovation Programme under the Grant Agreement Number 857251.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jag.2024.104283.

Data availability

Data availability Water bodies spatial polygon vectors are available to download at https://figshare.com/articles/dataset/Water_Bodies/23971620.

Code availability: The GEE and R codes are available on GitHub at https://github.com/FrankVal/WatSurf. The framework can be also replicated within GEE using the following paths: the first for data extraction (https://code.earthengine.google.com/04bcf145a2365b1a241aba51ffc412c7) and the second for data extrapolation (https://code.earthengine.google.com/d339e04034f87f53f2831b5918f5fa58).

References

Alvarez-Mozos, J., Villanueva, J., Arias, M., Gonzalez-Audicana, M. (2021, July). Correlation between NDVI and Sentinel-1 derived features for maize. In: 2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS (pp. 6773-6776). IEEE.

Attema, E., Davidson, M., Floury, N., Levrini, G., Rosich, B., Rommen, B., Snoeij, P. (2008, June). Sentinel-1 ESA's new European radar observatory. In *7th European conference on synthetic aperture radar* (pp. 1-4). VDE.

AEMET, I.P.M.A., 2011. Atlas climático ibérico/Iberian climate atlas. In: *Agencia Estatal de Meteorología, Ministerio de Medio Ambiente y Rural y Marino*. Instituto de Meteorologia de Portugal, Madrid.

Baldi, P., Brunak, S., Chauvin, Y., Andersen, C.A., Nielsen, H., 2000. Assessing the accuracy of prediction algorithms for classification: an overview. *Bioinformatics* 16 (5), 412–424.

Bangira, T., Alfieri, S.M., Menenti, M., Van Niekerk, A., 2019. Comparing thresholding with machine learning classifiers for mapping complex water. *Remote Sens. (Basel)* 11 (11), 1351.

Bhaduri, A., Bogardi, J., Siddiqi, A., Voigt, H., Vörösmarty, C., Pahl-Wostl, C., et al (2016). Achieving sustainable development goals from a water perspective. *Frontiers in Environmental Science*, 64.

Bioresita, F., Puissant, A., Stumpf, A., Malet, J.P., 2019. Fusion of Sentinel-1 and Sentinel-2 image time series for permanent and temporary surface water mapping. *Int. J. Remote Sens.* 40 (23), 9026–9049.

Breiman, L., 2001. Random Forests. *Machine Learning* 45, 5–32.

Cao, Y.G., Yan, L.J., Zheng, Z.Z., 2008. Extraction of information on geology hazard from multi-polarization SAR images. Int. Arch. Photogramm., Remote Sens. Spatial Inf. Sci. 1529–1532.

Céréghino, R., Biggs, J., Oertli, B., Declerck, S., 2008. The ecology of European ponds: defining the characteristics of a neglected freshwater habitat. *Hydrobiologia* 597, 1–6.

Chen, Z., Zhao, S., 2022. Automatic monitoring of surface water dynamics using Sentinel-1 and Sentinel-2 data with Google Earth Engine. *Int. J. Appl. Earth Obs. Geoinf.* 113, 103010.

Cochran, W.G., 1977. *Sampling techniques*. John Wiley Sons.

Cordeiro, M.C., Martinez, J.M., Peña-Luque, S., 2021. Automatic water detection from multidimensional hierarchical clustering for Sentinel-2 images and a comparison with Level 2A processors. *Remote Sens. Environ.* 253, 112209.

Cornes, R.C., van der Schrier, G., van den Besselaar, E.J., Jones, P.D., 2018. An ensemble version of the E-OBS temperature and precipitation data sets. J. Geophys. Res. *Atmospheres* 123 (17), 9391–9409.

De Groot, D., Brander, L., & Finlayson, C. M. (2018). Wetland ecosystem services. In *The Wetland book: I: Structure and function, management, and methods* (pp. 323-333). Springer Netherlands.

De Meester, L., Declerck, S., Stoks, R., Louette, G., Van De Meutter, F., De Bie, T., et al., 2005. Ponds and pools as model systems in conservation biology, ecology and evolutionary biology. Aquat. Conserv. Mar. Freshwat. Ecosyst. 15 (6), 715–725.

DeLancey, E.R., Czekajlo, A., Boychuk, L., Gregory, F., Amani, M., Brisco, B., et al., 2022. Creating a Detailed Wetland Inventory with Sentinel-2 Time-Series Data and Google Earth Engine in the Prairie Pothole Region of Canada. *Remote Sens. (Basel)* 14 (14), 3401.

Delegido, J., Verrelst, J., Alonso, L., Moreno, J., 2011. Evaluation of sentinel-2 red-edge bands for empirical estimation of green LAI and chlorophyll content. *Sensors* 11 (7), 7063–7081.

Deutsch, E.S., Alameddine, I., El-Fadel, M., 2018. Monitoring water quality in a hypereutrophic reservoir using Landsat ETM+ and OLI sensors: how transferable are the water quality algorithms? *Environ. Monit. Assess.* 190, 1–19.

Díaz-Paniagua, C., Florencio, M., de Felipe, M., Ramírez-Soto, M., Román, I., Arribas, R., 2024. Groundwater decline has negatively affected the well-preserved amphibian community of Doñana National Park (SW Spain). *Amphibia-Reptilia* 45 (2), 205–217.

Downing, J.A., Prairie, Y.T., Cole, J.J., Duarte, C.M., Tranvik, L.J., Striegl, R.G., et al., 2006. The global abundance and size distribution of lakes, ponds, and impoundments. *Limnol. Oceanogr.* 51 (5), 2388–2397.

Drusch, M., Del Bello, U., Carlier, S., Colin, O., Fernandez, V., Gascon, F., et al., 2012. Sentinel-2: ESA's optical high-resolution mission for GMES operational services. *Remote Sens. Environ.* 120, 25–36.

Erwin, K.L., 2009. Wetlands and global climate change: the role of wetland restoration in a changing world. *Wetl. Ecol. Manag.* 17 (1), 71–84.

Gao, B.C., 1996. NDWI—A normalized difference water index for remote sensing of vegetation liquid water from space. *Remote Sens. Environ.* 58 (3), 257–266.

Gašparović, M., Singh, S.K., 2022. Urban surface water bodies mapping using the automatic k-means based approach and sentinel-2 imagery. *Geocarto Int.* 2148757.

Gleick, P.H., 1998. Water in crisis: paths to sustainable water use. *Ecol. Appl.* 8 (3), 571–579.

Gómez-Rodríguez, C., Díaz-Paniagua, C., Serrano, L., Florencio, M., Portheault, A., 2009. Mediterranean temporary ponds as amphibian breeding habitats: the importance of preserving pond networks. *Aquat. Ecol.* 43, 1179–1191.

Google, Inc., Google Earth software, http://earth.google.com/ [last accessed on March 13, 2022].

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

Haralick, R.M., Shanmugam, K., Dinstein, I.H., 1973. Textural features for image classification. *IEEE Trans. Syst. Man Cybern.* 6, 610–621.

Heimhuber, V., Tulbure, M.G., Broich, M., 2016. Modeling 25 years of spatio-temporal surface water and inundation dynamics on large river basin scale using time series of Earth observation data. *Hydrol. Earth Syst. Sci.* 20 (6), 2227–2250.

Herceg-Szórádi, Z., Demeter, L., Csergő, A.M., 2023. Small area and low connectivity constrain the diversity of plant life strategies in temporary ponds. Divers. Distrib. 29 (5), 629–640.

Homobono, T., Guimarães, M.H., Esgalhado, C., Madureira, L., 2022. Water governance in mediterranean farming systems through the social-ecological systems framework—an empirical case in Southern Portugal. Land 11 (2), 178.

Huang, W., DeVries, B., Huang, C., Lang, M.W., Jones, J.W., Creed, I.F., Carroll, M.L., 2018. Automated extraction of surface water extent from Sentinel-1 data. *Remote Sens. (Basel)* 10 (5), 797.

Jiang, W., Ni, Y., Pang, Z., He, G., Fu, J., Lu, J., et al (2020). A new index for identifying water body from sentinel-2 satellite remote sensing imagery. *ISPRS Annals of Photogrammetry, Remote Sensing & Spatial Information Sciences*, (3).

Jiang, Z., Wen, Y., Zhang, G., Wu, X., 2022. Water Information Extraction Based on Multi-Model RF Algorithm and Sentinel-2 Image Data. *Sustainability* 14 (7), 3797.

Kuhn, M. (2015). Caret: classification and regression training. *Astrophysics Source Code Library*, ascl-1505.

Liaw, A., Wiener, M., 2002. Classification and regression by randomForest. *R News* 2 (3), 18–22.

Martensen, A.C., Saura, S., Fortin, M.J., 2017. Spatio-temporal connectivity: assessing the amount of reachable habitat in dynamic landscapes. Methods Ecol. Evol. 8 (10), 1253–1264.

Mayer, T., Poortinga, A., Bhandari, B., Nicolau, A.P., Markert, K., Thwal, N.S., et al., 2021. Deep learning approach for Sentinel-1 surface water mapping leveraging Google Earth Engine. *ISPRS Open Journal of Photogrammetry and Remote Sensing* 2, 100005.

McNairn, H., Brisco, B., 2004. The application of C-band polarimetric SAR for agriculture: A review. *Can. J. Remote. Sens.* 30 (3), 525–542.

Millard, K., Richardson, M., 2015. On the importance of training data sample selection in random forest image classification: A case study in peatland ecosystem mapping. *Remote Sens. (Basel)* 7 (7), 8489–8515.

Mishra, S., Mishra, D.R., 2012. Normalized difference chlorophyll index: A novel model for remote estimation of chlorophyll-a concentration in turbid productive waters. *Remote Sens. Environ.* 117, 394–406.

Morante-Carballo, F., Montalván-Burbano, N., Quiñonez-Barzola, X., Jaya-Montalvo, M., Carrión-Mero, P., 2022. What do we know about water scarcity in semi-arid zones? A global analysis and research trends. *Water* 14 (17), 2685.

Nasirzadehdizaji, R., Balik Sanli, F., Abdikan, S., Cakir, Z., Sekertekin, A., Ustuner, M., 2019. Sensitivity analysis of multi-temporal Sentinel-1 SAR parameters to crop height and canopy coverage. *Appl. Sci.* 9 (4), 655.

Ozesmi, S.L., Bauer, M.E., 2002. Satellite remote sensing of wetlands. *Wetl. Ecol. Manag.* 10, 381–402.

Pekel, J.F., Cottam, A., Gorelick, N., Belward, A.S., 2016. High-resolution mapping of global surface water and its long-term changes. *Nature* 540 (7633), 418–422.

Peña-Luque, S., Ferrant, S., Cordeiro, M.C., Ledauphin, T., Maxant, J., Martinez, J.M., 2021. Sentinel-1&2 multitemporal water surface detection accuracies, evaluated at regional and reservoirs level. *Remote Sens. (Basel)* 13 (16), 3279.

Petropoulos, G.P., Ireland, G., Barrett, B., 2015. Surface soil moisture retrievals from remote sensing: Current status, products & future trends. *Physics and Chemistry of the Earth, Parts a/b/c* 83, 36–56.

Rachid, N., Issam, N., Abdelkrim, B., Abdelhamid, A., Amina, H., 2022. Pond Energy Dynamics, Evaporation Rate and Ensemble Deep Learning Evaporation Prediction: Case Study of the Thomas Pond—Brenne Natural Regional Park (France). *Water* 14 (6), 923.

Radoux, J., Chomé, G., Jacques, D.C., Waldner, F., Bellemans, N., Matton, N., et al., 2016. Sentinel-2's potential for sub-pixel landscape feature detection. *Remote Sens. (Basel)* 8 (6), 488.

Reid, A.J., Carlson, A.K., Creed, I.F., Eliason, E.J., Gell, P.A., Johnson, P.T., et al., 2019. Emerging threats and persistent conservation challenges for freshwater biodiversity. *Biol. Rev.* 94 (3), 849–873.

Richardson, A.J., Wiegand, C.L., 1977. Distinguishing vegetation from soil background information. *Photogramm. Eng. Remote Sens.* 43 (12), 1541–1552.

Rouse, J.W., Haas, R.H., Schell, J.A., Deering, D.W., 1974. Monitoring vegetation systems in the Great Plains with ERTS. *NASA Spec. Publ* 351 (1), 309.

Samaniego, L., Thober, S., Kumar, R., Wanders, N., Rakovec, O., Pan, M., et al., 2018. Anthropogenic warming exacerbates European soil moisture droughts. *Nat. Clim. Chang.* 8 (5), 421–426.

Sebastián-González, E., Green, A.J., 2014. Habitat use by waterbirds in relation to pond size, water depth, and isolation: lessons from a restoration in southern Spain. Restor. Ecol. 22 (3), 311–318.

Shabbir, A.H., Zhang, J., Liu, X., Lutz, J.A., Valencia, C., Johnston, J.D., 2019. Determining the sensitivity of grassland area burned to climate variation in Xilingol, China, with an autoregressive distributed lag approach. *Int. J. Wildland Fire* 28 (8), 628–639.

Soria, J., Ruiz, M., Morales, S., 2022. Monitoring Subaquatic Vegetation Using Sentinel-2 Imagery in Gallocanta Lake (Aragón, Spain). *Earth* 3 (1), 363–382.

Sun, F., Sun, W., Chen, J., Gong, P., 2012. Comparison and improvement of methods for identifying waterbodies in remotely sensed imagery. *Int. J. Remote Sens.* 33 (21), 6854–6875.

Swets, J.A., 1988. Measuring the accuracy of diagnostic systems. Science 240 (4857), 1285–1293.

Tang, H., Lu, S., Ali Baig, M.H., Li, M., Fang, C., Wang, Y., 2022. Large-scale surface water mapping based on landsat and sentinel-1 images. *Water* 14 (9), 1454.

Team, R. C. (2021). R: A language and environment for statistical computing. Published online 2020.

Tulbure, M.G., Broich, M., Stehman, S.V., Kommareddy, A., 2016. Surface water extent dynamics from three decades of seasonally continuous Landsat time series at subcontinental scale in a semi-arid region. *Remote Sens. Environ.* 178, 142–157.

Tulbure, M.G., Broich, M., Perin, V., Gaines, M., Ju, J., Stehman, S.V., et al., 2022. Can we detect more ephemeral floods with higher density harmonized Landsat Sentinel 2 data compared to Landsat 8 alone? *ISPRS J. Photogramm. Remote Sens.* 185, 232–246.

United Nations. Department of Economic and Social Affairs. (2022). The Sustainable Development Goals: Report 2022. UN.

Valavi, R., Elith, J., Lahoz-Monfort, J.J., Guillera-Arroita, G., 2018. blockCV: An r package for generating spatially or environmentally separated folds for k-fold cross-validation of species distribution models. *Biorxiv* 357798.

Valerio, F., Ferreira, E., Godinho, S., Pita, R., Mira, A., Fernandes, N., Santos, S.M., 2020. Predicting microhabitat suitability for an endangered small mammal using sentinel-2 data. *Remote Sens. (Basel)* 12 (3), 562.

Valerio, F., Godinho, S., Marques, A.T., Crispim-Mendes, T., Pita, R., Silva, J.P., 2024. GEE_xtract: High-quality remote sensing data preparation and extraction for multiple spatio-temporal ecological scaling. *Eco. Inform.* 80, 102502.

Vanderhoof, M.K., Alexander, L., Christensen, J., Solvik, K., Nieuwlandt, P., Sagehorn, M., 2023. High-frequency time series comparison of Sentinel-1 and Sentinel-2 satellites for mapping open and vegetated water across the United States (2017–2021). *Remote Sens. Environ.* 288, 113498.

Veldkamp, T.I., Wada, Y., de Moel, H., Kummu, M., Eisner, S., Aerts, J.C., Ward, P.J., 2015. Changing mechanism of global water scarcity events: Impacts of socioeconomic changes and inter-annual hydro-climatic variability. *Glob. Environ. Chang.* 32, 18–29.

Vreugdenhil, M., Navacchi, C., Bauer-Marschallinger, B., Hahn, S., Steele-Dunne, S., Pfeil, I., et al., 2020. Sentinel-1 cross ratio and vegetation optical depth: A comparison over Europe. *Remote Sens. (Basel)* 12 (20), 3404.

Wang, L., Bie, W., Li, H., Liao, T., Ding, X., Wu, G., Fei, T., 2022. Small Water Body Detection and Water Quality Variations with Changing Human Activity Intensity in Wuhan. *Remote Sens. (Basel)* 14 (1), 200.

Wickens, G.E. (1998). Arid and Semi-arid Environments of the World. In: Ecophysiology of Economic Plants in Arid and Semi-Arid Lands. Adaptations of Desert Organisms. Springer, Berlin, Heidelberg.

Wilson, E.H., Sader, S.A., 2002. Detection of forest harvest type using multiple dates of Landsat TM imagery. *Remote Sens. Environ.* 80 (3), 385–396.

Zhou, Y., Dong, J., Xiao, X., Xiao, T., Yang, Z., Zhao, G., et al., 2017. Open surface water mapping algorithms: A comparison of water-related spectral indices and sensors. *Water* 9 (4), 256.

Zou, Z., Xiao, X., Dong, J., Qin, Y., Doughty, R.B., Menarguez, M.A., et al., 2018. Divergent trends of open-surface water body area in the contiguous United States from 1984 to 2016. *Proc. Natl. Acad. Sci.* 115 (15), 3810–3815.