Data in Brief 43 (2022) 108437

Contents lists available at ScienceDirect

Data in Brief

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

Data Article

Shoreline contour, water level elevation and volumetric dataset (1984-2020) for the Gallocanta Lake (NE Spain)

Jesús Palomar-Vázquez[a], Carlos Cabezas-Rabadán[a,b,*], Alfonso Fernández-Sarría[a], Enrique Priego-de-los-Santos[c], Ramón Pons-Crespo[c], Josep E. Pardo-Pascual[a]

[a] *Geo-Environmental Cartography and Remote Sensing Group[1], Department of Cartographic Engineering, Geodesy and Photogrammetry, Universitat Politècnica de València, Camí de Vera s/n, Valencia 46022, Spain*
[b] *Univ. Bordeaux, CNRS, Bordeaux INP, EPOC, UMR 5805, F-33600 Pessac, France*
[c] *Department of Cartographic Engineering, Geodesy and Photogrammetry, Universitat Politècnica de València, Camí de Vera s/n, Valencia 46022, Spain*

ARTICLE INFO

Article history:
Received 14 April 2022
Revised 20 June 2022
Accepted 27 June 2022
Available online 1 July 2022

Dataset link: Multi-decadal shoreline position, water level elevation and volumetric changes in the Gallocanta Lake (NE Spain) (Original data)

Keywords:
Shoreline changes
Water level fluctuations
Remote sensing
Lake monitoring
Shoreline extraction
Satellite-derived shorelines
Landsat
Sentinel

ABSTRACT

Gallocanta is the largest well-preserved saline lake in Western Europe, included in the Ramsar List. Associated with its shallow morphology, the lake undergoes strong variations in its water surface extent along time that condition the habitat distribution and the ecological functions. Data on the morphology of the lake and its hydrological variations along time may be of paramount ecological importance for the managers of this natural space. Even though its interest for research and management purposes, no accurate and robust dataset of this nature covering large periods of time is available.
This dataset presents a multi-decadal mapping with a sub-weekly frequency (2-5 days) of the contour of the Gallocanta Lake (NE Iberian Peninsula) along the period 1984-2020 (1043 dates with information). The shoreline position appears continuously defined with subpixel accuracy from the freely-available images acquired by the satellites Sentinel-2 (sensor MSI) and Landsat 5 (TM), 7 (EMT+), and 8 (OLI) by

* Corresponding author at: Geo-Environmental Cartography and Remote Sensing Group, Department of Cartographic Engineering, Geodesy and Photogrammetry, Universitat Politècnica de València, Camí de Vera s/n, Valencia 46022, Spain.
E-mail address: carcara4@upv.es (C. Cabezas-Rabadán).

[1] @CGAT_UPV

https://doi.org/10.1016/j.dib.2022.108437

applying the extraction system SHOREX. The satellite-derived shorelines allow the definition of the surface of the lake and are combined with a digital elevation model to assign elevation values to the points defining each shoreline. This allows deducing the mean elevation of the water level and the volumetric changes for those same dates. This data package constitutes a valuable source of information for carrying out robust analyses of the trends of the lake along decades, as well as its response to individual rainfall events.

Specifications Table

Subject	Earth-Surface Processes
Specific subject area	Application of remote sensing techniques for the cartography of environmental processes
Type of data	Geospatial: vector layers of the shoreline contour in shapefile format and the associated attributes.
How the data were acquired	The shoreline contour was obtained from the freely-available images acquired by the satellites Landsat 5, Landsat 7, and Landsat 8 from the U.S. Geological Survey, and Sentinel-2 from the European Space Agency employing the system SHOREX (http://cgat.webs.upv.es/) for shoreline extraction. The resulting area covered by the water was calculated using the QGIS software (www.qgis.org). Using this same software, the elevation values were assigned to the shoreline positions by combining with a digital elevation model from LiDAR data by the National Geographic Institute of Spain, allowing the estimation of the water volume.
Data format	Raw data
Description of data collection	The obtention of shorelines contours was carried out from optical satellite imagery. The images affected by clouds (above 30% for the whole scene) were discarded, while a subsequent manual checking of clouds in the vicinity of the lake was also carried out. Each polygon corresponds to the definition of the area covered by water on each of the dates on which a satellite image free of clouds is available. Each vector entity contains fields quantifying the water surface, the water elevation and the estimated water volume.
Data source location	Satellite optical imagery from the satellites Landsat 5, 7, and 8 is available from the Earth Explorer of the U.S. Geological Survey (USGS) (https://earthexplorer.usgs.gov/), while the Sentinel-2 images are available from the Copernicus Open Access Hub (https://scihub.copernicus.eu/). The elevation of the terrain is available from the digital elevation model from LiDAR data by the National Geographic Institute of Spain (IGN, https://pnoa.ign.es/el-proyecto-pnoa-lidar), and may be downloaded from https://centrodedescargas.cnig.es/CentroDescargas/index.jsp.
Data accessibility	Repository name: Mendeley Data Data identification number: doi:10.17632/9w93mbbzts.2 Direct URL to data: https://data.mendeley.com/datasets/9w93mbbzts/2 Palomar-Vazquez, Jesus; Cabezas-Rabadán, Carlos; Fernandez, Alfonso; Priego-de-los-Santos, Enrique; Pons, Ramón; Pardo-Pascual, Josep E. (2022), “Multi-decadal shoreline position, water level elevation and volumetric changes in the Gallocanta Lake (NE Spain)”, Mendeley Data, V2, doi:10.17632/9w93mbbzts.2

Value of the Data

- The Gallocanta Lake is the largest well-preserved saline lake in Western Europe [1] and part of the List of Ramsar Wetlands of International Importance since 1994. Associated with its shallow morphology, the lake undergoes strong variations in its water content along time that condition the intricate distribution of shallow ponds and other coastal habitats of high environmental value. The changes along time of the hydrological and morphological conditions, and the periods of scarcity of water critically determine the chances of survival for the flora and fauna that inhabits there. Despite the abundance of literature focused on the ecology of Gallocanta and their geoforms [2], there is no accurate database on its morphological and hydrological conditions covering large periods of time. In this type of water body, traditional techniques for the measurement of the elevation as the manual measurements using in-situ rules do not offer robust mid-term measurements. The salinity and the strong water level fluctuations [3] lead to movements and the degradation of the rules, creating data gaps and the disconnection between the derived data. In contrast with that, the employed remote sensing techniques data in many cases offer a limited accuracy, especially at smaller lakes [4]. Considering that, the dataset presented in this work constitutes a novel and appealing resource by offering a continuous characterization of the morphology of Gallocanta Lake and its hydric variations with sub-weekly frequency.
- The availability of a multi-decadal and accurate dataset of the shoreline position together with other representative morphological parameters is of interest to ecologists and environmental and hydric managers. Lakes with strong water surface fluctuations as the Gallocanta Lake may act as indicators of environmental changes at different scales. This is the case of the occurrence of hydrological changes, human actions such as aquifer exploitation, land-use changes, climate change, and other phenomena affecting large territories. On the other hand, there has been a recent bloom of research works attempting to detect lake surface changes [4]. Therefore, any remote sensing scientist or image analyst can benefit from using this dataset as reference data for comparison purposes.
- This is an accessible and free-of-charge dataset that can provide baseline information for further studies aimed at researching the hydric response, and the correspondence between geoforms and lacustrine habitats at lakes in general and at Gallocanta Lake in particular. The analysis of the volumetric changes of the lake permits the estimation of the volume of surface inputs associated with rainfall events. This would enable a better definition of the magnitude of subsurface or groundwater input, and the identification of water losses caused by evapotranspiration, infiltration, or direct subtractions. The shorelines with subpixel accuracy allow a precise definition of the shape of the limits of the water surface. This is of primary interest for ecological studies as shoreline changes in shallow and fluctuating lakes define the distribution of lacustrine habitats and, subsequently are key for the preservation of valuable species. Furthermore, this large package of shorelines with subpixel accuracy may constitute a useful dataset for validation or comparison purposes of other methodologies that attempt the shoreline definition in water bodies from remote sensing. This may be the case of the methods based on the SAR imagery, for which low accuracies have been reported, as well as other alternative methods based on optical imagery so far scarcely used in the characterization and monitoring of inner water bodies.
- The provision of such long series of volume variations provides insights for recognising and quantifying hydrological anomalies (droughts/wet periods), their duration, magnitude and frequency. Also, as the series of records grows, the detection of long-term trends of change that can be associated with the effects of the climate change.

1. Data Description

This dataset consists of a complete mapping of the contour of the Gallocanta Lake at sub-pixel level between June 1984 and June 2020 together with the associated water surface, water elevation and water volume. The shoreline contours are provided as vector polylines in shapefile format that can be imported by most Geographic Information Systems (GIS) software. The coordinate reference system is WGS 1984, projection UTM Zone 30N (EPSG code: 32630). The dataset consists of five files featuring information regarding geometry (.shp), attribute table (.dbf), positional index (.shx), and the projection description (.prj and .qpj).

Fields description:

- Date: acquisition date of the satellite image (YYYYMMDD)
- Sensor: sensor identifier (S2, LT05, LE07, LC08)
- Area_km2: polygon area (square meters)
- Depth_m: height of the water surface (meters)
- Volume_hm3: water volume (cubic hectometers)

2. Experimental Design, Materials and Methods

2.1. Materials

The study considers two data sources:

- Landsat and Sentinel-2 satellite imagery
- Digital elevation model of the study area

2.1.1. Satellite Optical Imagery

The images acquired by the optical satellites Sentinel-2 (sensor MSI) and Landsat 5 (TM), 7 (EMT+), and 8 (OLI). They are available free of charge from the Copernicus Open Access Hub of the European Space Agency (https://scihub.copernicus.eu/) and the Earth Explorer of the U.S. Geological Survey (USGS) (https://earthexplorer.usgs.gov/). The period 1984-2020 is covered by the combination of the different Landsat missions (L5 from 1984 to 2011, L7 from 1999, and L8 from 2013), while Sentinel-2 mission became operational in 2015 (satellite 2A followed by its twin 2B in 2017). These satellites offer a revisit time of 16 and 10 days for Landsat and Sentinel satellites respectively. Because of the progressive overlap of the different missions, the revisit time has been reduced, currently reaching about 2.9 days when all the satellites are combined [5]. From May 31, 2003 onwards Landsat 7 presents data gaps caused by the failure of the Scan Line Corrector (SLC). For this reason, the lowest availability of data occurs between November 2011 (last employed Landsat 5 image) and April 2013 (Landsat 8 images became available). The images contain the middle spatial resolution (10 to 30 m) bands RGB, NIR, SWIR1, and SWIR2. As described in by Sánchez-García et al. [6] the radiometric characteristics of all the satellites enable the extraction of comparable shoreline positions.

2.1.2. Digital Elevation Model

The work relies on the elevation data obtained from a Digital Elevation Model (DEM) covering the entire area that is flooded at some point during the study period. The DEM is defined from LiDAR data acquired by the National Geographic Institute of Spain (IGN, https://pnoa.ign.es/el-proyecto-pnoa-lidar). The point cloud shows an average density of 0.5 points/m^2 and an altimetric accuracy of each point better than 20 cm RMSE. The data were acquired during October 2010 at a time when the lake showed a very low water level, but it was not desiccated and caused the appearance of patches without information. Nevertheless, as the lakebed bottom is extremely flat, the data recorded offered sufficient spatial resolution to interpolate the areas without signal.

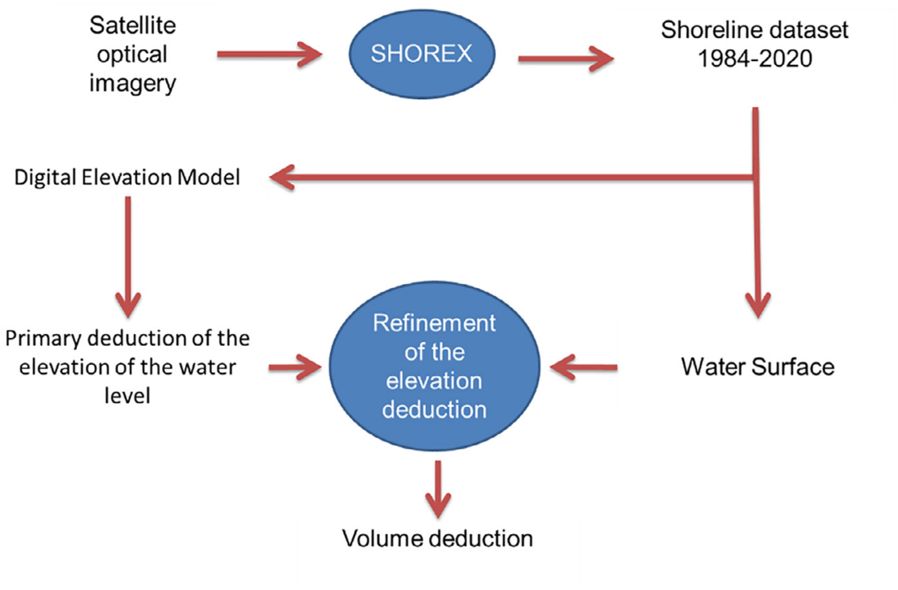

Fig. 1. Methodology for defining the surface, elevation and volume of the lake.

2.2. Methods

The volume deduction is based on the obtention of the shoreline contour of the lake and its combination with the DEM (Fig. 1) for each date.

2.2.1. Shoreline Extraction and Water Surface Definition

The downloading of the satellite imagery, pre-processing, and extraction of SDS were carried out using the system for shoreline extraction SHOREX [6,7, http://cgat.webs.upv.es/]. The images were downloaded after filtering those showing more than 30% of cloud coverage for the whole scene, while a manual cloud-checking process was carried out to discard those with clouds covering the lake and the immediate area. The shoreline position was automatically defined for the resulting 1043 images as the water/land intersection during their acquisition time. Initially, the shore was approximately defined at pixel level by a mask created according to the AWEINSH index [8] setting a constant threshold value of 0. Using this approximate line as the reference, the final subpixel shoreline was extracted by applying the algorithm described by Pardo-Pascual et al. [9] by operating over the Short-Wave Infrared bands (SWIR1) and using a third-degree polynomial and 3 × 3 analysis kernels. According to previous assessments of this extraction methodology (see [6]) the resulting shorelines are expected to offer a subpixel accuracy of about 3 m RMSE.

In accordance with these shoreline positions, the contour of the water present in the lake was defined as a polygon. Subsequently, the quantification of the water surface was carried out using the QGIS software (www.qgis.org). The data gaps linked to L7 images acquired during the failure of the were filled using morphological filters in the approximate mask defined by the index AWEINSH.

2.2.2. Mean elevation of the water level and water depth

Assuming that the water surface is flat over a certain area [10] the elevation data available in the DEM was assigned to the points defining the shoreline, subsequently allowing to calculate the mean elevation of the water surface. Only the western half of the lake was considered in

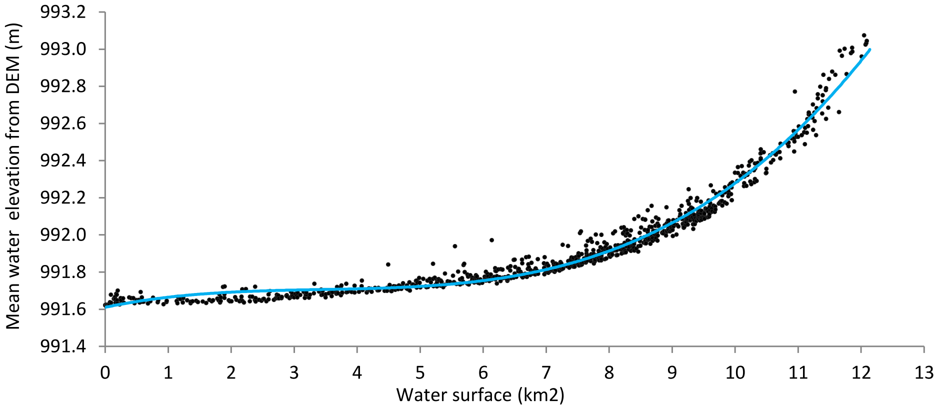

Fig. 2. Relationship between the lake water surface and its mean elevation according to the DEM (dots), and the numerical model relating both variables (solid line).

this step as its gentler slopes may minimize the positioning errors of the shorelines on the resulting elevation values. Furthermore, a refinement of the elevation deduction was proposed based on modeling the relationship between the water surface and the elevation. Thus, all the pairs of surface and elevation values recorded over time were numerically related by a third-degree polynomial function (Eq. 1, Fig. 2) that describes the relationship with a goodness of fit of $R^2 = 0.978$.

$$y = 0.0019x^3 - 0.0191x^2 + 0.0716x + 991.61; \tag{1}$$

being x the water surface (in km^2) and y the mean elevation according with the DEM (in m).

The water depth was obtained by subtracting to all the values of the series the minimum elevation value registered along time (991.6 m), that was assimilated to the complete desiccation of the lake.

2.2.3. Volume Estimation

The volumetric changes along time were defined by comparing the elevation and the surface area of each date with those registered during the first date of the study period. Surface and elevation differences were multiplied to obtain the relative volumetric changes. A volume value of zero was assigned to the date with the largest negative volumetric change, as it was assumed to be linked to the complete desiccation. The volumetric series was recalculated by adding the relative volumetric change value of that reference date, therefore enabling the transformation of all the deduced values into absolute volumes.

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.

Data Availability

Multi-decadal shoreline position, water level elevation and volumetric changes in the Gallocanta Lake (NE Spain) (Original data) (Mendeley Data).

CRediT Author Statement

Jesús Palomar-Vázquez: Conceptualization, Methodology, Data curation, Software, Formal analysis, Writing – review & editing; **Carlos Cabezas-Rabadán:** Conceptualization, Methodology, Data curation, Visualization, Writing – original draft, Writing – review & editing; **Alfonso Fernández-Sarría:** Data curation, Resources; **Enrique Priego-de-los-Santos:** Data curation, Resources; **Ramón Pons-Crespo:** Data curation, Resources; **Josep E. Pardo-Pascual:** Conceptualization, Methodology, Writing – review & editing, Supervision, Funding acquisition, Project administration.

Acknowledgments

This research is supported by the projects 'Monitorización de precisión de las fluctuaciones de agua de humedales salinos intermitentes RAMSAR en la cuenca media del Ebro mediante teledetección espacial' with the financial support of Fundación Biodiversidad from the Spanish Ministry for Ecological Transition and the Demographic Challenge (MITECO),and MONOBESAT (PID2019-111435RB-I00) funded by the Spanish Ministry of Science, Innovation and Universities, and the Margarita Salas contract within the Re-qualification programme by the Spanish Ministry of Universities financed by the European Union - NextGenerationEU.

References

[1] E. Luna, C. Castañeda, F.J. Gracia, R. Rodríguez-Ochoa, Late Quaternary pedogenesis of lacustrine terraces in Gallocanta Lake, NE Spain, Catena 147 (2016) 372–385, doi:10.1016/j.catena.2016.07.046.

[2] C. Castañeda, J. Herrero, The water regime of the Monegros playa-lakes established from ground and satellite data, J. Hydrol. 310 (2005) 95–110, doi:10.1016/j.jhydrol.2004.12.007.

[3] C. Castañeda, F.J. Gracia, J.A. Conesa, B. Latorre, Geomorphological control of habitat distribution in an intermittent shallow saline lake, Gallocanta Lake, NE Spain, Sci. Total Environ. 726 (2020) 138601, doi:10.1016/j.scitotenv.2020.138601.

[4] J.F. Crétaux, W. Jelinski, S. Calmant, A. Kouraev, V. Vuglinski, M. Bergé-Nguyen, M.C. Gennero, F. Nino, R.A.D. Rio, A. Cazenave, SOLS: a lake database to monitor in the Near Real Time water level and storage variations from remote sensing data, Adv. Space Res. 47 (2011) 1497–1507, doi:10.1016/j.asr.2011.01.004.

[5] J. Li, D.P. Roy, A global analysis of Sentinel-2a, Sentinel-2b and Landsat 8 data revisit intervals and implications for terrestrial monitoring, Remote Sens. 9 (2017) 902, doi:10.3390/rs9090902.

[6] E. Sánchez-García, J.M. Palomar-Vázquez, J.E. Pardo-Pascual, J. Almonacid-Caballer, C. Cabezas-Rabadán, L. Gómez-Pujol, An efficient protocol for accurate and massive shoreline definition from mid-resolution satellite imagery, Coast. Eng. 160 (2020) 103732, doi:10.1016/j.coastaleng.2020.103732.

[7] C. Cabezas-Rabadán, J.E. Pardo-Pascual, J. Palomar-Vázquez, Characterizing the relationship between the sediment grain size and the shoreline variability defined from sentinel-2 derived shorelines, Remote Sens. 13 (14) (2021) 2829, doi:10.3390/rs13142829.

[8] G.L. Feyisa, H. Meilby, R. Fensholt, S.R. Proud, Automated water extraction index: a new technique for surface water mapping using Landsat imagery, Remote Sens. Environ. 140 (2014) 23–35, doi:10.1016/j.rse.2013.08.029.

[9] J.E. Pardo-Pascual, J. Almonacid-Caballer, L.A. Ruiz, J. Palomar-Vázquez, Automatic extraction of shorelines from landsat TM and ETM+ multi-temporal images with subpixel precision, Remote Sens. Environ. 123 (2012) 1–11, doi:10.1016/j.rse.2012.02.024.

[10] D.J. Penton, I.C. Overton, Spatial modelling of floodplain inundation combining satellite imagery and elevation models, in: Proceedings of the MODSIM 2007 International Congress on Modelling and Simulation, Clayton south, Vic, Australia, Modelling and Simulation Society of Australia and New Zealand CSIRO, 2007.