DYNA
http://dyna.medellin.unal.edu.co/

UNIVERSIDAD NACIONAL DE COLOMBIA

Landsat 8-based estimation of total suspended solids (1.3-43 mg/L) using NDWI, MNDWI and AWEI indices at Tecocomulco Ramsar Wetland, Mexico

Ismael Ortega-García, René Cruz-Guerrero & German Cuaya-Simbro

Instituto Tecnológico del Oriente del Estado de Hidalgo, Tecnológico Nacional de México, Sahagún, Hidalgo, México. 23030537m@itesa.edu.mx, rcruz@itesa.edu.mx, gcuaya@itesa.edu.mx

Received: May 11th, 2025. Received in revised form: November 5th, 2025. Accepted: December 1st, 2025.

Abstract
This study integrates Landsat 8 multispectral imagery and three water segmentation indices (NDWI, MNDWI, AWEI) to estimate Total Suspended Solids (TSS) concentrations in Tecocomulco Lagoon, a Ramsar wetland in Hidalgo, Mexico. Using Google Earth Engine for image acquisition and Python-based processing with Rasterio and NumPy libraries, we applied the TSS algorithm: TSS = (RED/NIR) × F, where F=10 was calibrated against in-situ measurements. Eight spectral bands were processed from cloud-free images (<50% coverage) captured between January-December 2023. Results showed TSS concentrations ranging from 1.3 to 43 mg/L, with strong correlation ($r=0.89$, $p<0.05$) against CONAGUA reference data from three monitoring sites. Statistical validation yielded RMSE=4.2 mg/L, MAE=3.5 mg/L, and $R^2=0.79$. The combined use of three indices improved water body segmentation accuracy compared to single-index approaches. This methodology provides cost-effective monitoring for the 1,769-ha wetland, supporting conservation efforts for 15 duck species and the endangered Ambystoma mexicanum. The approach offers a replicable framework for TSS monitoring in similar high-sedimentation wetlands.

Keywords: remote sensing; Tecocomulco; Ramsar; satellite images.

Estimación basada en Landsat 8 de sólidos suspendidos totales (1.3-43 mg/L) usando los índices NDWI, MNDWI y AWEI en el Humedal Ramsar de Tecocomulco, México

Resumen
Este estudio integra imágenes multiespectrales de Landsat 8 y tres índices de segmentación de agua (NDWI, MNDWI, AWEI) para estimar las concentraciones de Sólidos Suspendidos Totales (SST) en la Laguna de Tecocomulco, un humedal Ramsar en Hidalgo, México. Utilizando Google Earth Engine para la adquisición de imágenes y procesamiento basado en Python con las bibliotecas Rasterio y NumPy, aplicamos el algoritmo de SST: SST = (ROJO/NIR) × F, donde F=10 fue calibrado con mediciones in situ. Se procesaron ocho bandas espectrales a partir de imágenes libres de nubes (<50% de cobertura) capturadas entre enero y diciembre de 2023. Los resultados mostraron concentraciones de SST que oscilaron entre 1.3 y 43 mg/L, con una fuerte correlación ($r=0.89$, $p<0.05$) respecto a los datos de referencia de CONAGUA provenientes de tres sitios de monitoreo. La validación estadística arrojó RMSE=4.2 mg/L, MAE=3.5 mg/L y $R^2=0.79$. El uso combinado de los tres índices mejoró la precisión de la segmentación del cuerpo de agua en comparación con los enfoques de índice único. Esta metodología proporciona un monitoreo eficaz para el humedal de 1,769 ha, apoyando los esfuerzos de conservación de 15 especies de patos y del Ambystoma mexicanum, especie en peligro de extinción. El enfoque ofrece un marco replicable para el monitoreo de SST en humedales similares con alta sedimentación.

Palabras clave: teledetección; Tecocomulco; Ramsar; imágenes satelitales.

How to cite: Ortega-García, I., Cruz-Guerrero, R. and Cuaya-Simbro, G., Landsat 8-Based estimation of total suspended solids (1.3-43 mg/L) using NDWI, MNDWI and AWEI indices at Tecocomulco Ramsar Wetland, Mexico DYNA, (93)240, pp. 36-45, January - March, 2026.

Revista DYNA, (93)240, pp. 36-45, January - March, 2026, ISSN 0012-7353
DOI: https://doi.org/10.15446/dyna.v93n240.120317

1. Introduction

There is currently a growing concern about water resources, due to the increase in temperature caused by climate change, which together with human activities, directly impact the disappearance of aquatic ecosystems, as in the case of Tecocomulco and other lakes in the basin of the Valley of Mexico [1], which has driven the application of computer vision technologies for the analysis of images captured by satellites, which orbit the earth performing environmental monitoring missions. Traditionally, water quality assessment is performed by in situ sampling that is subsequently analyzed in the laboratory, processes that are costly and limited in terms of spatial and temporal coverage. The Tecocomulco lagoon, recognized as a Ramsar site since 2003, faces problems derived from agricultural practices, indiscriminate logging, urbanization, and untreated waste, which have deteriorated water quality and consequently affected biodiversity and the health of riparian communities [2].

Nowadays, remote sensing is a promising alternative that allows obtaining data on a large spatial scale, through access to collections of satellite images and spatial data sets. Computational processing facilitates the interpretation of large volumes of data, extracting patterns and quantifying physicochemical parameters from the measurement of their reflectance on the Earth's surface, by measuring solar radiation, the images obtained are representations of wavelengths ranging from infrared to ultraviolet [3]. In addition, the use of satellites for earth surface observation has been constantly improved since the first space missions in the 1960s, capturing increasingly accurate images and data through sensors, such as Operational Land Imager (OLI) and Thermal Infrared Sensor (TIRS) of the Landsat missions led by the National Aeronautics Administration of the United States NASA.

Remote sensing as a technological discipline based on sciences such as physics, computing or geography, allows the application of techniques for the study of the earth's surface. Currently, the images obtained by the satellite missions of various space agencies are available for access and download in various digital platforms, such as: Google Earth Engine (GEE), Copernicus, USGS Browser Earth, etc. This article reports the research where the amount of Total Suspended Solids (TSS), considered as an optically active parameter, is calculated [4]. The calculation was made from images taken by the Landsat 8 satellite, which captures the study area every 16 days, allowing to regularly identify variations in water quality, making it possible to perform environmental monitoring of a wetland of great ecological importance such as Tecocomulco.

Previous studies have successfully applied TSS algorithms to various water bodies. Ritchie et al. (1987) [5] established the foundational correlation between red and NIR bands for sediment detection. Gómez & Dalence (2014) [6] validated similar approaches in the Bogotá River, achieving correlations of $r>0.85$. However, applications to high-altitude endorheic wetlands with significant agricultural impact remain limited. This study addresses this gap by adapting established algorithms to the unique conditions of Tecocomulco, an endorheic wetland at 2,514 m.a.s.l. characterized by high sedimentation and turbidity rates due to agricultural erosion [2], with fluctuating water levels influenced by seasonal precipitation patterns [1].

The Tecocomulco lagoon is located in the state of Hidalgo and comprises a regular area of 1,769 ha, with coordinates 19º52'N 098º23'W [7], this body of water arose during the Plio-Cuaternary (~3 million years ago) due to tectonic processes associated with the subduction of the Cocos plate under the North American plate. Volcanic activity in this area included the formation of dacite rock domes, basaltic lava flows and scoria cones, aligned with regional northeast-southeast oriented faults [8]. The volcanic activity created depressions and structures that facilitated the accumulation of water, contributing to the formation of lake bodies such as the Tecocomulco lagoon, whose location can be observed in Fig. 1, according to the visualization platform of the National Water Information System (SINA) of the National Water Commission, a Mexican state agency.

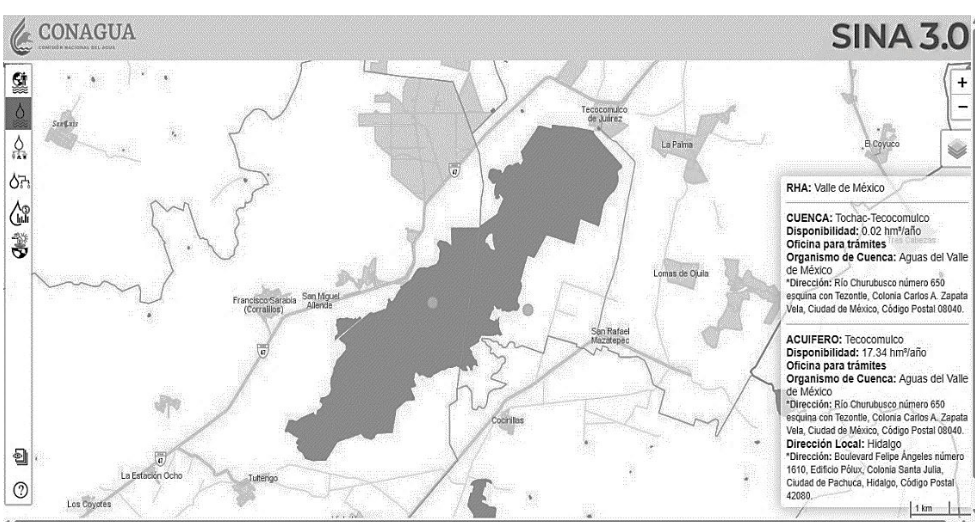

Figure 1. Screenshot of the SINA website where the country's main water bodies are monitored.
Source: Own elaboration. https://sinav30.conagua.gob.mx:8080/SINA/?opcion=acuiferos

Likewise, the Tecocomulco endorheic basin is located at 2514 (m.a.s.l.) and is considered part of the remnants of the ancient freshwater wetland system, which dominated the valley of Mexico basin in the center of the country 500 years ago. The underlying aquifers are recharged through the lake, whose water level is fluctuating due to the impact of meteorological phenomena causing extended periods of rain or drought [1]. The site is notable for the presence of 15 species of ducks, with congregations of ruddy ducks Oxyura jamaicensis of up to 5000 individuals, and coots Gallinula chloropus of up to 3000 [2].

The endangered salamander Ambystoma mexicanum also inhabits this body of water, threatened by invasive species such as the Asian carp (Cyprinus carpo), a non-endemic species that is commercially exploited by local fishermen. Given such a diverse biological environment, there are portions of fauna for waterfowl hunting, which is practiced during the winter and is a relevant economic activity for some people in the communities. Erosion of the basin caused by agricultural activities, logging, and overgrazing have caused a high rate of sedimentation and turbidity in the lake, which is identified in the Ramsar site catalog with the number 1322 [7].

The importance of the Tecocomulco lagoon has been internationally recognized since 2003 with its declaration as a Ramsar site, a category that emerged in 1971 in the convention that Mexico signed as a contracting party, as did 90% of the member states of the United Nations [7]. This importance lies in its ecological value due to its biodiversity and the environmental functions that characterize this lentic wetland. Its importance is key in the central migratory route of the North American region, which is evidenced by the birds such as the white pelican (Pelecanus erythrorhynchos), the Mexican duck (Anas diazi), the white heron (Ardea alba), the gray crane (Grus canadensis) and endangered species such as the common scoter (Emberiza schoeniclus) [9]. It also serves as a refuge during the nesting stages, which are critical for the development of these species.

With a significant reduction in recent decades, the natural reservoir located in the municipalities of Tepeapulco, Apan and Cuautepec in the state of Hidalgo is of great importance for international associations such as Ducks Unlimited de México, A.C. (DUMAC), who offer relevant information on Ramsar sites on their map server, as can be seen in Fig. 2. (DUMAC), who offer relevant information of the Ramsar sites in their cartographic server, as can be seen in Fig. 2, this project allows the visualization of wetlands of great importance for migratory species, where through various layers presents the characteristics of wetland coverage, as a sample of the development of useful computer applications for water management.

Research question: Can the Landsat 8 spectral band algebra (RED/NIR relationship) accurately estimate TSS concentrations in aquatic areas segmented by multiple indices in the Tecocomulco lagoon, as validated by measurements from government institutions?

Objectives:

- Demonstrate the opportunity in the use of satellite imagery for the analysis of physico-chemical parameters of water in the Tecocomulco Lagoon.
- To apply and validate computational algorithms for segmentation and pattern recognition and computational processing in different areas of aquatic coverage shown in satellite images of the Tecocomulco wetland.
- Compare the results obtained for SST parameters with historical measurements reported by official agencies.

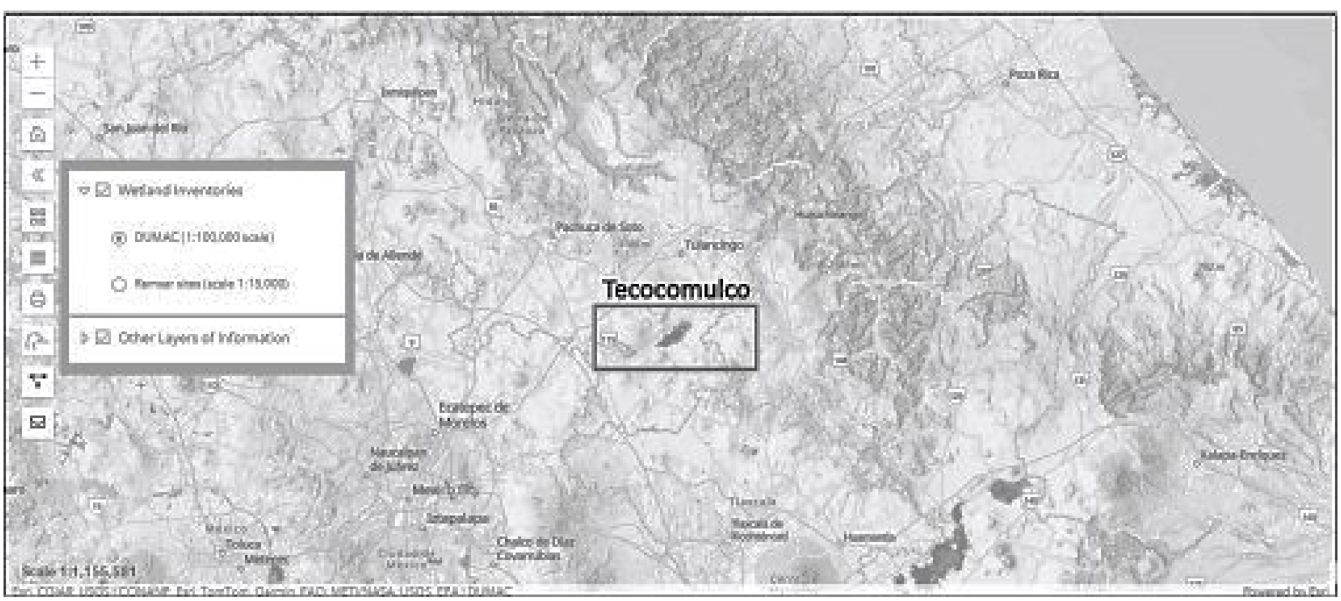

Figure 2. Geographic location of the Tecocomulco lagoon as seen on the DUMAC map server, which presents coverage characteristics of wetlands relevant to migratory species.
Source: Own elaboration

2. Methods and materials

2.1 Measurement of reflectance and spectral signature

Reflectance is an optical property of all surfaces and describes the portion of incident radiation that is capable of being reflected by that surface. It is commonly expressed as a percentage and is the physical principle that allows the inference of crustal characteristics in remote sensing. The measurement of reflectance is performed through sensors that are capable of capturing the reflected radiation at different wavelengths of the electromagnetic spectrum, in Fig. 3 can be seen the different wavelengths of which only a small portion is visible to the human eye.

For the capture of satellite images, according to Richards and Jia (2006) [3], a typical process can be considered as one that involves the capture of data through a sensor that captures the radiation from the earth's surface, followed by the correction of solar irradiance before reflection, then corrections that normalize light absorption and scattering to obtain more accurate values, finally reflecting in a reflectance value that can be used for computational analysis of multispectral images in remote sensing studies.

Each material on the earth's surface has a spectral signature, which can be considered as a graphic representation of the reflectance or emissivity as a function of the wavelength emitted by that material, so that vegetation, buildings or water have a unique spectral signature that is due to their physical-chemical properties and allows the identification of different materials on the earth's surface, through computational algorithms that operate directly with images and data captured by satellite sensors, which record the reflectance of lengths in specific ranges of the electromagnetic spectrum.

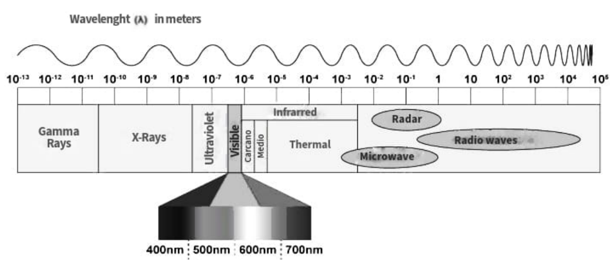

Figure 3. Graphical representation of the electromagnetic spectrum showing the wavelengths at which light is transmitted, indicating the spectrum visible to the human eye.
Source: Own elaboration

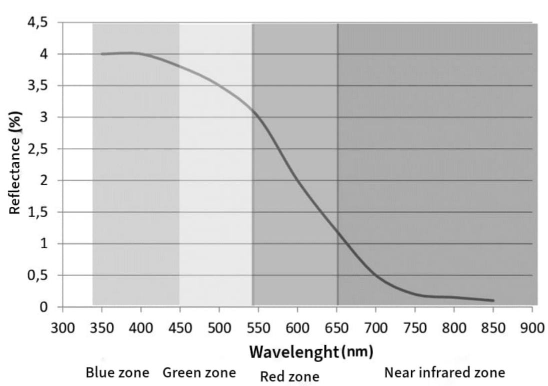

Figure 4. Graphical representation of the spectral signature of water in its pure form.
Source: Own elaboration

In Fig. 4 we can observe the spectral signature of water in its purest form, according to a study by Drozd & Fernandez (2016) [10] where different spectral signatures are analyzed for segments of the Uruguay River. It should be noted that the composition of surface water bodies becomes more complex as there are more variables, such as aquatic vegetation, climatic conditions or irregularities caused by the topographic circumstances of the terrain under study [10].

2.2 Water quality and total suspended solids

According to Richards & Jia (2016) [3], water quality is the set of physical, chemical and biological characteristics that determine its suitability for a specific use, among which human consumption, agriculture, industry, recreation and conservation of the aquatic ecosystem can be considered. These conditions can be affected mainly by the presence of chemical substances such as heavy metals, nutrients (nitrates and phosphates) and the presence of biological agents such as viruses or bacteria, affecting public health, the effectiveness of agriculture and the sustainability of the ecosystem in general.

The physical conditions of the water, such as turbidity, temperature, amount of suspended solids and color, can influence the life of aquatic organisms and the economic activities carried out by the population around the body of water, such as recreational and tourist activities, which have a direct impact on the economy of the communities.

The measurement of Total Suspended Solids (TSS) can be considered relevant in the determination of water quality as it is an indicator of pollution, since suspended particles can include sediments, organic matter, microorganisms and chemical contaminants. Therefore, high TSS levels can signal a deterioration of water quality due to increased human activities such as agricultural runoff or excessive erosion of

Figure 5. Photograph showing the transparency of the water, which at first sight reveals the aquatic vegetation of the natural reservoir. Taken on March 8, 2025.
Source: Own elaboration.

the watershed. The impact of high TSS levels in the water body darkens the aquatic environment, causing a decrease in light penetration, affecting the photosynthesis of algae and aquatic plants, thus altering the food chain and generating direct effects on biodiversity [3] The measurement of this parameter can be an indicator of water quality, because it reveals the dry weight of the particles present in the water column and is represented in units of mg/L [4]. Fig. 5 shows an image captured in the water mirror of the Tecocomulco lagoon, where the level of transparency of the reservoir can be observed in some areas.

2.3 Determination of the study area

The study of a water body requires the georeferential location to delimit the geographic area to be analyzed. This procedure is key in obtaining the results sought, known as ROI. The region of interest is defined by the coordinates of the longitude and latitude vertices of the area from which the satellite images are obtained. In our case, this delimitation was performed using USGS Explorer, as shown in Fig. 6, where the polygon was defined with the coordinates: [-98.442135, 19.833731, -98.335018, 19.906703], indicating the vertices that delimit the area and are subsequently used in the access platform for Landsat 8 collections.

2.4 Calculation of indices for water body segmentation and TSS measurement through band algebra

Once the study area was chosen, it was necessary to segment the aquatic coverage surface, since this is where the TSS calculation will be performed. For this task, there are several computer vision algorithms, such as edge detection and the application of filters. However, those that are best suited to work with georeferenced images, such as satellite collections, are the calculation of indices, which performs direct algebraic operations on the spectral bands, represented as matrices of reflectance of each of the pixels of the image, as shown in Fig. 7.

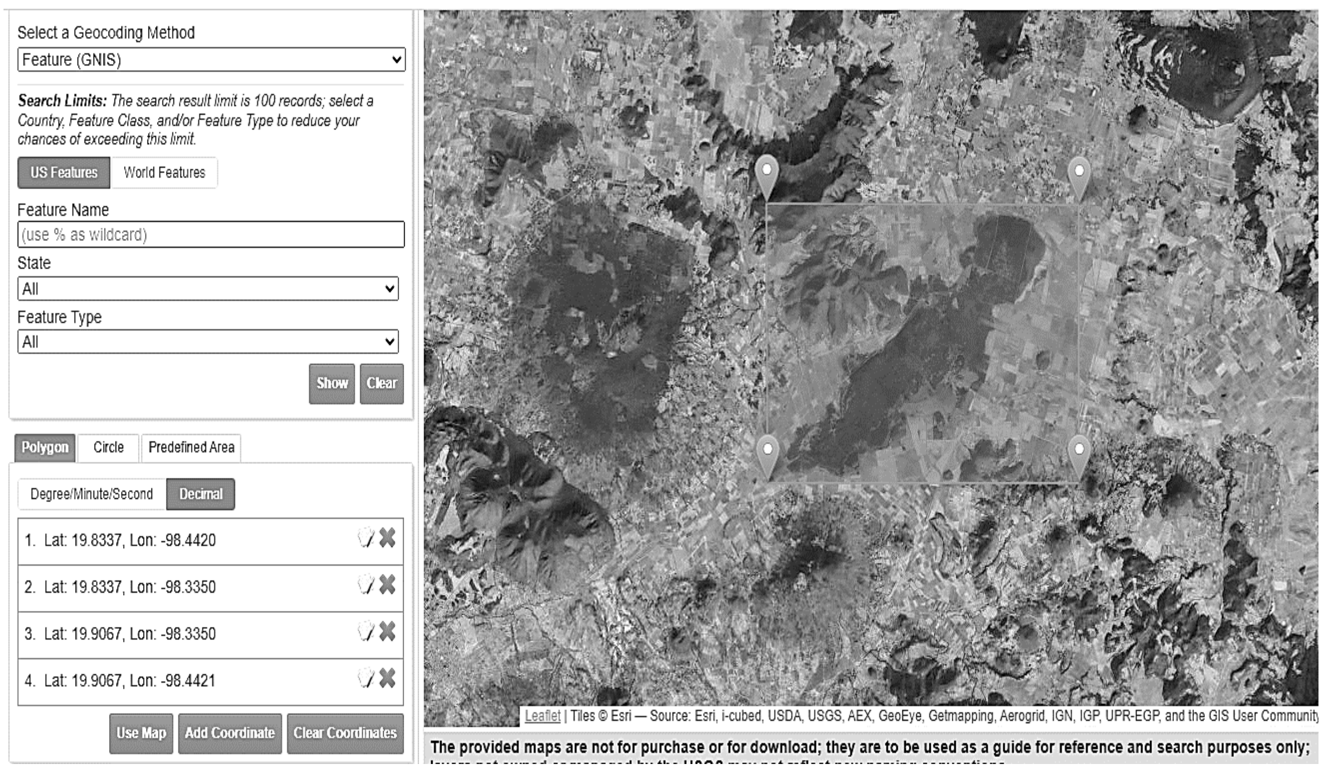

Figure 6. Definition of the region of interest: drawing a polygon in the USGS Earth Explorer platform to delimit the area comprising the Tecocomulco lagoon.
Source: Own elaboration

The use of indices to differentiate surface water from other elements of the landscape, due to their spectral properties, is one of the ways in which a water body can be segmented using the spectral bands of a satellite image. The NDWI index (Normalized Difference Water Index), proposed by McFeeters in 2013, uses the green and near-infrared (NIR) bands of satellite images, taking advantage of the high absorption of water in the wavelength range between 700 and 2500 nm of the NIR band and its moderate reflectance in the green, which allows for the highlighting of water masses. However, this index can overestimate urban areas and suffer interference from vegetation, thus reducing the area represented compared to the actual surface of the reservoir [11].

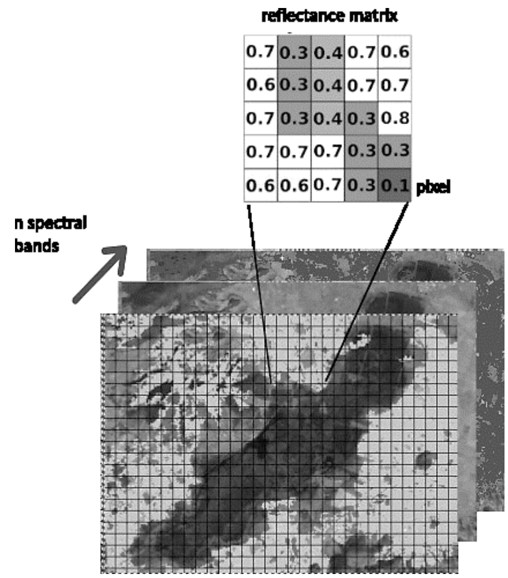

Figure 7. Visual interpretation of satellite images as arrays of pixel reflectance.
Source: Own elaboration

The use of indices to differentiate surface water from other elements of the landscape, due to their spectral properties, is one of the ways in which a water body can be segmented using the spectral bands of a satellite image. The NDWI index (Normalized Difference Water Index), proposed by McFeeters in 2013, uses the green and near-infrared (NIR) bands of satellite images, taking advantage of the high absorption of water in the wavelength range between 700 and 2500 nm of the NIR band and its moderate reflectance in the green, which allows for the highlighting of water masses. However, this index can overestimate urban areas and suffer interference from vegetation, thus reducing the area represented compared to the actual surface of the reservoir [11].

Normalized Differentiated Water Index (NDWI)

$$NDWI = \frac{(NIR - SWIR)}{(NIR + SWIR)} \tag{1}$$

Where:
NIR: Near Infrared (SR_B5)
SWIR: Show Wave Infrared (SR_B6)

To improve the representation of the aquatic surface, the MNDWI (Modified Normalized Difference Water Index) will also be used, which replaces the NIR band with the shortwave infrared (SWIR) band that captures wavelengths between 900 and 2500 nm, improving the separation between water and built-up areas, since the latter reflect the SWIR to a lesser extent [12]. This makes it possible to capture water surfaces more effectively in complex environments [13].

Modified Normalized Difference Water Index

$$MNDWI = \frac{(GREEN - SWIR)}{(GREEN + SWIR)} \tag{2}$$

Where:
GREEN = Green spectral band (SR_B3)

For its part, the AWEI (Automated Water Extraction Index) was developed by Feyisa et al. (2014) [14], combines multiple bands such as those of the visible spectrum, NIR, and SWIR, and can be implemented together with dynamic thresholds (such as Otsu's algorithm) to minimize false positives caused by shadows, clouds, or barren soils. Its design makes it a robust index in various environmental conditions [15].

Automated Water Withdrawal Index

$$AWEI = 4 * (GREEN - SWIR) - (0.25 * NIR) + (2.75 * SWIR) \tag{3}$$

The calculation of TSS sediment concentration by remote sensing is based on the relationship between water surface reflectance and the presence of suspended particles. Studies such as that of Ritchie et al. (1987) [5] have demonstrated the correlation between the red (600-700 nm) and near-infrared (700-800 nm) spectral bands, which are particularly sensitive to reflectance caused by suspended sediments in the water column. The authors, pioneers in the development of these studies, developed linear and nonlinear regression models that establish a relationship between atmospheric reflectance and TSS concentration. In addition, more recent studies have verified this correlation in measurements of this type in various surface aquifers, as reported in measurements made with spectral bands in the Bogotá River basin in Colombia [6].

The equation that determines the amount of TSS is as follows:

$$SST = \frac{RED}{NIR} * F \tag{4}$$

Where:
SST: Total Suspended Solids in mg/L.
RED: Red spectral band (SR_B4).
NIR: Near Infrared (SR_B5)
F: Scale factor according to the measurement.

2.5 Data collection and preprocessing

The satellite images were obtained through the development of a code created in the JavaScript language on the web development platform provided by GEE, which allows access to collections of satellite images ready for computational processing. Fig. 8 shows the interface for accessing USGS satellite images through GEE, which, once obtained, can be downloaded directly to the Google Drive file hosting service.

The images obtained from the spectral bands are downloaded to Drive with a TIF extension; these belong to the Landsat 8 Collection 2 Level 1, processed by the United States Geological Survey (USGS). These satellite image collections are available and have been prepared for computational processing.

Figure 8. Capture of the platform for access to satellite imagery and geospatial datasets. (GEE, 2025).
Source: Own elaboration

The USGS applies advanced algorithms to remove distortions caused by aerosol scattering, solar reflectance, and terrain topography, thus ensuring that the data represent realistic surface reflectance values. This standardized preprocessing allows users of the spectral bands to access information ready for computational analysis and calculation of spectral indices without requiring additional adjustments [16].

Data available at USGS Level 1 are radiometrically and geometrically calibrated, achieving better georeferencing, with an RMSE ≤ 12 m, using ground control points and digital elevation models. This collection includes multispectral (30 m), panchromatic (15 m), and thermal bands, in addition to key metadata such as cloud cover, date of capture, and solar/sensor angles. Tier 1 consists of Tier 1 (highest quality) and Tier 2 (which includes scenes with geometric or cloudiness limitations). Quality bands (QA_PIXEL/RADSAT) for automated filtering and optimized formats (COG) are incorporated into this collection, ensuring accuracy and compatibility with computational processing in the cloud [16].

2.6 Phyton and Rasterio

The use of Python together with the Rasterio library allows for the efficient processing of satellite images to calculate the amount of TSS in mg/L. Rasterio is an API that facilitates the reading of spectral bands (GeoTIFF) and their analysis using NumPy, making it possible to use mathematical operations to combine spectral bands and calculate physico-chemical indices of land surface areas with accurate georeferencing. Development with Python makes the analysis from pixels to the complete scene possible in a simple way. By means of matrix operations, the formulas described above are implemented.

This integration makes it possible to calculate TSS at the pixel level, generating visual representations of the TSS measurement distribution that serve to identify zones with different sediment concentrations, thus optimizing the water quality analysis in the Tecocomulco Lagoon [6].

The study is based on a multi-paradigm approach, integrating Remote Sensing, digital image processing, and software development. The primary data came from multispectral images from the Landsat 8 satellite (Collection 2 Level 1), acquired through the Google Earth Engine (GEE) platform and processed with JavaScript for the extraction of spectral bands. The calculation of Total Suspended Solids (TSS) was implemented in Python 3.1, using the libraries Rasterio (GeoTIFF data reading/manipulation), NumPy (for band algebra implementation), and OpenCV (JET color mapping for visualization). Statistical validation included correlation analysis (rr coefficient) between derived SST and CONAGUA in situ data, plus dynamic normalization using percentiles (2% and 98%) to adjust the range of values. Calibration of the scaling factor (F = 10) was based on previous studies in lentic wetlands [6].

The temporal scope covered January to December 2023, with 23 cloud-free scenes (<50% coverage) analyzed. Images were selected based on the 16-day Landsat 8 revisit cycle, prioritizing dry season captures (November-April) when TSS variability is highest. The three indices were integrated through Boolean algebra: Water_Mask = (NDWI > 0) AND (MNDWI > 0.1) OR (AWEI > -0.5), with thresholds determined via Otsu's algorithm.

The graphical interface (GUI), developed with Tkinter, allows for loading images in GeoTIFF format and visualizing measurement results in the area of aquatic coverage segmented by MNDWI, NDWI, and AWEI indices. In addition, a visualization of histograms and spectral signatures is generated using Matplotlib. The Rasterio library version used is 1.3.9, and cloudiness thresholds were defined (<50%). The validation data are available in the National Water Information System (SINA), ensuring methodological transparency [17]. Table 1 shows the values reported by CONAGUA for TSS measurements in three regions of the reservoir.

Table 1.
TSS measurement data in the Tecocomulco lagoon.

Site Key	Year	Zone	SST mg/L
DLHID1483	2023	THE ISLET LAGOON OF TECOCOMULCO	30
DLHID1484	2023	THALIMEDES LAGOON TECOCOMULCO	<25
DLHID1485	2023	TULTENGO LAGOON OF TECOCOMULCO	40

Carried out by CONAGUA and accessible in the National Water Information System.
Source:Own elaboration

3. Results and discussion

3.1 Results

The comparative evaluation of the collected data allowed for the observation of a solid correlation ($r > 0.89$) between the derived spectral indices and the laboratory analyses available on the SINA site of CONAGUA. The capacity of the segmentation algorithm using the NDWI, MNDWI, and AWEI indices allows for an improvement in segmentation compared to the use of only one index, as observed in the area detected as surface water; as shown on the left side of Fig. 9, this value is much lower than that observed on the right side. For this reason, in order to perform the segmentation more effectively, it was necessary to identify the area of aquatic coverage using more than one index.

Fig. 10 shows the interface developed in Python to load multispectral images from the Landsat 8 satellite, specifically designed to process the red (SR_B4) and near-infrared (SR_B5) spectral bands, which are essential for the calculation of Total Suspended Solids. The interface allows loading files in TIF format, guaranteeing the integrity of the geo-referenced data. The 'Calculate TSS' button activates the automatic processing of the images by means of matrix operations with the Rasterio and NumPy libraries, following the algorithm that implements the formula explained in the methodology section.

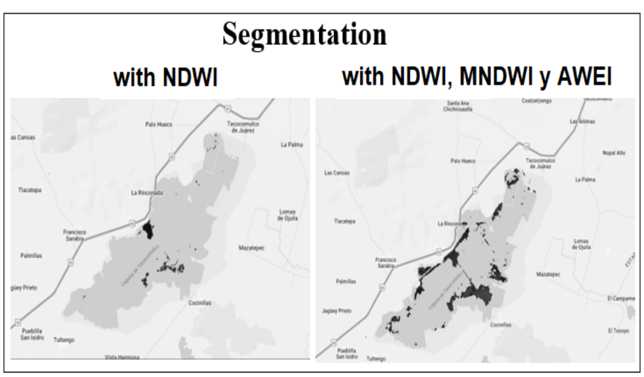

Figure 9. Segmentation masks captured from the calculation of the NDWI index in the first column and NDWI + NDWI + AWEI in the second column, from spectral images obtained from the Landsat 8 satellite through JavaScript code in the GEE platform.
Source:Own elaboration.

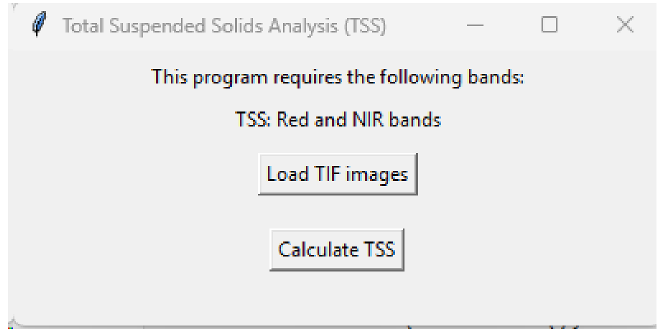

Figure 10. Interface of the module developed in Python for loading spectral bands and their computational processing.
Source: Own elaboration

The processing results are presented in Fig. 11. In the upper box, the area of surface water coverage can be seen on a color scale from blue to red, where the latter represents the highest TSS mg/L measurements, with a value not calculated (NaN) in non-aquatic areas.

To validate the accuracy of remote sensing algorithms in estimating Total Suspended Solids (TSS) in the Tecocomulco Lagoon, values derived from Landsat 8 images (processed in Python) were compared with in situ data reported by CONAGUA (Table 2).

The statistical metrics used for validation were the Pearson Correlation Coefficient (r), Root Mean Square Error (RMSE), Bias, Mean Absolute Error (MAE), and Coefficient of Determination (R^2), as shown in Table 2.

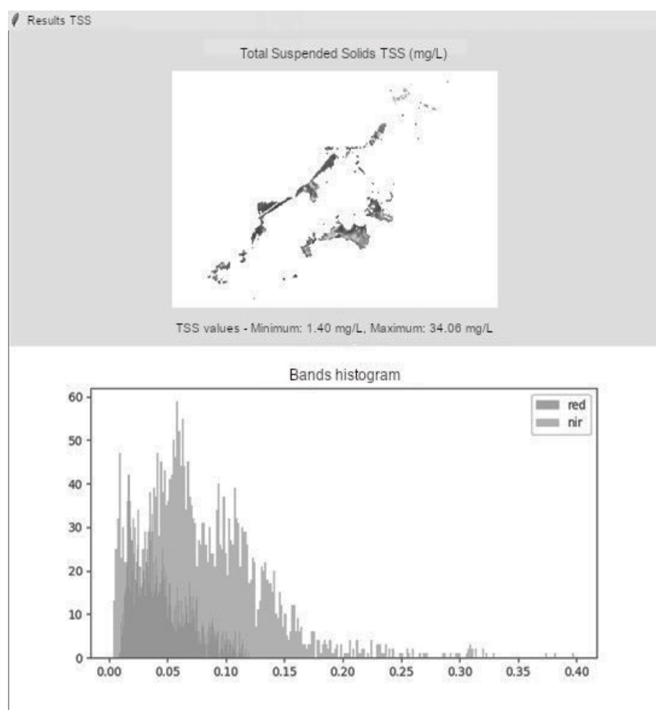

Figure 11. Presentation of results from the Python module for calculating TSS in mg/L at the water surface of the Tecocomulco Lagoon.
Source: Own elaboration.

Table 2
Use metrics to evaluate the correlation between results obtained from remote sensing and those from CONAGUA's SINA.

Metrics	Value	Equiation	Interpretation
Pearson's correlation coefficient (r)	$r = 0.89\ (p < 0.05)$	$r = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2 \sum(y_i - \bar{y})^2}}$	It indicates a strong and statistically significant correlation.
Mean Squared Error RMSE	4.2 mg/L	$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}$	It suggests a moderate discrepancy between the methods.
Bias	+1.8 mg/L	$\mathrm{Sesgo} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)$	This implies a slight underestimation by the satellite.
Mean Absolute Error MAE	3.5 mg/L	$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}\lvert y_i - \hat{y}_i\rvert$	It represents acceptable accuracy for monitoring applications.
Coefficient of Determination (R^2)	0.79	$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}$	It indicates that 79% of the variability in the laboratory data is explained by the model.

Source: Own elaboration

3.2 Discussion

The linear relationship between satellite data and data obtained through laboratory studies of in situ samples validates the use of RED and NIR bands to estimate the SST parameter in turbid waters, consistent with previous studies [6], such as that of Gómez & Dalence (2014). The accuracy shown by the RMSE metric is comparable to other studies in lentic wetlands [6], although calibrating the scale factor (F) with a greater number of iterations is suggested. Meanwhile, the positive bias can be interpreted as an underestimation that could be due to atmospheric interference or the spatial resolution of Landsat 8 (30 m/pixel), which averages values in heterogeneous areas. Finally, the scarcity of data restricts statistical robustness. Similar studies recommend expanding the validation with more sampling at different times of the year.

The observed RMSE of 4.2 mg/L and positive bias of +1.8 mg/L are within acceptable ranges for operational monitoring, comparable to Gómez & Dalence's (2014) [6] findings in Colombian wetlands (RMSE = 5.1 mg/L). The 30m spatial resolution limitation could be addressed in future studies using Sentinel-2 imagery (10m resolution). The underestimation bias likely results from sub-pixel mixing in shallow areas (<1m depth) where bottom reflectance influences the signal. Ecologically, the TSS range of 1.3-43 mg/L indicates moderate to high turbidity levels that may impact submerged vegetation photosynthesis and benthic invertebrate communities, critical food sources for migratory waterfowl.

4. Conclusions

This study has shown that the integration of remote sensing, spectral indices, and segmentation models constitutes an effective methodology for the delineation and analysis of water bodies in complex environments, thus answering our research question. The application of NDWI, MNDWI, and AWEI allowed for the accurate identification of water surfaces, reducing overestimation in urban areas and minimizing vegetation interference. Thanks to the use of satellite images from the Landsat 8 mission and their preprocessing in Google Earth Engine (GEE), the extraction of relevant data was facilitated, thus optimizing the spectral analysis of water quality.

The correlation found between spectral values and in situ measurements supports the validity of this methodology for environmental monitoring aimed at water management. On the other hand, the integration of computational tools, such as Python and Rasterio, into data processing has been crucial for satellite image processing, ensuring the consistency and reliability of the results. However, certain limitations have been identified that should be considered in future research. The spatial resolution of the sensors used may be insufficient to detect microvariations in smaller bodies of water, suggesting that incorporating higher-resolution images or combining them with data from airborne sensors could improve the accuracy of the analysis. Furthermore, atmospheric conditions during image capture can influence data quality, highlighting the need to implement improvements in radiometric correction algorithms and consider interpolation techniques to minimize these effects.

Finally, the findings of this study underscore the relevance of remote sensing and machine learning in water resources management and monitoring. Automating segmentation and spectral analysis reduces the dependence on physical sampling, thus optimizing water quality assessment processes and facilitating decision-making in environmental policies. This methodology demonstrates the practical application of established remote sensing techniques to a previously unstudied high-sedimentation wetland, providing baseline TSS data for future monitoring programs.

References

[1] Lanza, G., and Gomez, G., Analysis of the short term geo-environmental evolution of an endorreic basic using satelital images: the caso of Tecocomulco Lagoon, México: Investigaciones

Geográficas, Boletín del Instituto de Geografía, UNAM, (58), pp. 66-79, 2005. DOI: https://doi.org/10.14350/rig.30048
[2] Delgadillo, A.E., Determinación de parámetros fisicoquímicos, estado eutrófico y metales pesados de la laguna de Tecocomulco, Hidalgo; identificación de compuestos quelantes de Hydrocotyle ranunculoides L.f., Instituto de Ciencias Básicas e Ingeniería, UAEH, [online]. (58), pp. 1-4, 2012. Available at: http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S0188-46112005000300005&lng=es&nrm=iso
[3] Richards, J.A., and Jia, X., Remote sensing digital image analysis, Berlin: Springer, pp. 265-278, 2005. DOI: https://doi.org/10.1007/3-540-29711-1
[4] Sabry, F., Teledetección del agua: avances en técnicas de visión por computadora para la teledetección del agua, Madrid: mil millones de conocimientos, 2024.
[5] Ritchie, J.C., Cooper, C.M., and Yongqing, J., Using landsat multispectral scanner data to estimate suspended sediments in moon lake, mississippi., Remote Sensing of Environment, 23(1), pp. 65-81, 1987. DOI: https://doi.org/10.1016/0034-4257(87)90071-X
[6] Gómez-Diaz, J.L., and Delance-Martinic, J.S., Determinación del parámetro sólidos suspendidos totales (SST) mediante imágenes de sensores ópticos en un tramo de la cuenca media del río Bogotá (Colombia)., Revista UD y la Geomática, 9, pp. 19-27, 2014. DOI: https://doi.org/10.14483/23448407.7943
[7] Secretariat-Ramsar-Convention, The list of wetlands of international importance, ramsar convention, [online]. pp. 34, 2005, Available at: https://www.ramsar.org/sites/default/files/documents/library/sitelist.pdf
[8] García-Tovar, G.P., Martínez-Serrano, R.G., Sole, J., Correa-Tello, J.C., Núñez-Castillo, E.Y., Guillou, H., and Monroy-Rodríguez, E., Geology, geochronology and geochemestry of the Plio-Quaternary volcanism of the Apán-Tecocomulco Volcanic Field, Transmexican Volcanic Belt, Revista Mexicana de Ciencias Geologicas, [online]. 32(1), pp. 100-122, 2015. Available at: https://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S1026-87742015000100009&lng=es&nrm=iso
[9] Apolonio-Ortega, L.A., La evaluación de la avifauna de la Laguna de Tecocomulco y alrededores, estado de Hidalgo: prioridades de conservación., Ciudad de México: Facultad de Ciencias UNAM, [online]. 2017. Available at: https://repositorio.unam.mx/contenidos/400714
[10] Drozd, A., and Fernandez, V., Capítulo VII: Caracterización de firmas espectrales en distintos tramos del río Uruguay y su relación con variables biológicas y físicas del agua. Informe de actividades de investigación de la subcomisión de medio ambiente y uso sostenible del agua, estudio de la calidad del agua en el río Uruguay vigilancia de playas y estado trófico., Montevideo, Universidad de la República de Uruguay, [online]. 2016, pp. 298-308. Available at: https://www.caru.org.uy/web/sub_medio_ambiente/Informe%20Bienal%202013-14.pdf
[11] McFeeters, S.K., Using the Normalized Difference Water Index (NDWI) within a geographic information system to detect swimming pools for Mosquito Abatement: a practical approach, Department of Geography, California State University Fresno, 2013, pp. 3544-3561. DOI: https://doi.org/10.3390/rs5073544
[12] Xu, H., Modification of Normalized Difference Water Index (NDWI) to enhance open water features in remotely sensed imagery. International Journal of Remote Sensing, 27(14), pp. 3026-3033, 2006. DOI: https://doi.org/10.1080/01431160600589179
[13] Kulkarni, R., Khare, K., and Khanum, H., Detecting, extracting, and mapping of inland surface water using Landsat 8 operational land imager: a case study of Pune district, India., F1000Research, 11, pp. 774, 2022. DOI: https://doi.org/10.12688/f1000research.121740.1
[14] Feyisa, G.L., Meilby, H., Fensholt, R., and Proud, S.R., Automated water extraction index: a new technique for surface water mapping using Landsat imagery, Remote Sensing of Environment, 140, pp. 23-35, 2014. DOI: https://doi.org/10.1016/j.rse.2013.08.029
[15] Laonamsai, J., Julphunthong, P., Saprathet, T., Kimmany, B., Ganchanasuragit, T., Chomcheawchan, P., and Tomun, N., Utilizing NDWI, MNDWI, SAVI, WRI, and AWEI for Estimating Erosion and Deposition in Ping River in Thailand, Hydrology, 10(3), art. 70, 2023. DOI: https://doi.org/10.3390/hydrology10030070
[16] Eros, C., Landsat 8-9 Operational Land Imager / Thermal Infrared Sensor Level-1, Collection 2 [dataset], U.S. Geological Survey, 2022. DOI: https://doi.org/10.5066/P975CC9B
[17] CONAGUA-SINA, SINA 3.0. Sistema Nacional de Información del Agua: base de datos de calidad del agua en cuerpos superficiales. Comisión Nacional del Agua, México, [en linea]. Disponible en: https://sinav30.conagua.gob.mx:8080/SINA/?opcion=acuiferos.

Note:
Data Availability: The Python code for TSS calculation and GEE scripts for image acquisition are available at: https://github.com/Ismaelog9/SST_Tecocomulco

I. Ortega-García, received his BSc. Eng. in Computer Systems in 2017. He has worked as a consultant in application development in C#, front-end and back-end development, and is currently a systems analyst in the auto parts manufacturing sector. He has participated as an exhibitor at the Multidisciplinary International Congress organized by the National Institute Technology of Mexico and the Technological Institute of Eastern Hidalgo State. He is currently pursuing a Master's degree in Computer Systems at ITESA with an interest in the areas of artificial intelligence and deep learning, industry 4.0, and AI applications in environmental monitoring.
ORCID: 0009-0008-1525-7569.

G. Cuaya-Simbro, received the Dr. in Computer Science doing research in the area of machine learning. He has 8 years of professional experience in the industry, working as a senior consultant in Information Technology developed by IBM, such as mainfrain solutions, high availability, virtualization and cognitive computing, among others. He is currently a member of the SNI of CONAHCYT level C and is a research professor in the technological system of Mexico (TecNM), where he has taught undergraduate and graduate courses, directed graduate work at both levels, and has collaborated with national and international research groups to develop research projects that directly impact national problems, such as Education, Health, Food Sovereignty, among others, and with special emphasis on issues that contribute to achieving the UN SDGs. The lines of research currently working are those related to Industry 4.0 and Digital Transformation, such as Artificial Intelligence, Machine learning, robotics, IoT, Cloud Computing, BlockChain, Bigdata, cyber-physical systems, among others.
ORCID: 0000-0001-6303-154X

R. Cruz-Guerrero, received the PhD. in Computer Science from the Autonomous University of the State of Hidalgo, Mexico. He is a member of the National System of Researchers (SNI) at the CONAHCYT and research professor at the Technological System of Mexico (TecNM). His main research interests include Ontological Models, Industry 4.0, Artificial Intelligence, Bigdata, Deep learning and Semantic systems.
ORCID: 0000-0003-1276-2419