

Carolina Doña Monzó

Monitoring water quality and hydrological patterns of wetlands using recent techniques in remote sensing



Doctorado en Teledetección

Departament de Física de la Terra i Termodinàmica

Facultat de Física

PhD Thesis 2016



VNIVERSITAT
DE VALÈNCIA

PhD Thesis
November 2016
Carolina Doña Monzó

PhD advisors: Prof. Vicente Caselles Miralles
Prof. Juan Manuel Sánchez Tomás
Prof. Antonio Camacho González

VNIVERSITAT
DE VALÈNCIA (ò) Facultat de Física

UNIVERSITAT DE VALÈNCIA

Facultat de Física

Departament de Física de Terra i Termodinàmica



VNIVERSITAT
DE VALÈNCIA

Monitoring water quality and hydrological patterns of wetlands using recent techniques in remote sensing

Doctorado en Teledetección

PhD Thesis

November 2016

CAROLINA DOÑA MONZÓ

PhD advisors: Prof. Vicente Caselles Miralles

Prof. Juan Manuel Sánchez Tomás

Prof. Antonio Camacho González

D. VICENTE CASELLES MIRALLES, Catedrático de Física de la Tierra en el departamento de Física de la Tierra y Termodinámica de la Facultad de Física, Universidad de València,

D. JUAN MANUEL SÁNCHEZ TOMÁS, Profesor Contratado Doctor en el departamento de Física Aplicada de la Escuela Politécnica de Cuenca, Universidad de Castilla-La Mancha Y

D. ANTONIO CAMACHO GONZÁLEZ, Profesor Titular en el departamento de Microbiología y Ecología de la Facultad de Biología, Universidad de València,

CERTIFICAN: Que la presente memoria "*Monitoring water quality and hydrological patterns of wetlands using recent techniques in remote sensing*", presentada por Carolina Doña Monzó ha sido realizada bajo su dirección constituyendo su Tesis Doctoral para optar a grado de Doctora en Teledetección por la Universidad de València.

Y para que así conste, en cumplimiento de la legislación vigente, firman el presente certificado en Burjassot, a 30 de noviembre de 2015



Fdo.: Vicente Caselles Miralles



Fdo.: Juan Manuel Sánchez Tomás



Fdo.: Antonio Camacho González

Este trabajo ha sido posible gracias a la financiación económica recibida por parte de:

El Proyecto EL USO DE LA TELEDETECCIÓN PARA MEJORAR LA OBSERVACIÓN DE LA ATMÓSFERA Y EL CLIMA (CLIMATEL) (CGL2007-64666/CLI), financiado por el Ministerio de Educación y Ciencia (MEC), así como su acción complementaria CGL2007-29819-E/CLI.

El Proyecto EL CAMBIO CLIMÁTICO A TRAVÉS DE LA TELEDETECCIÓN (CGL2010-17577), financiado por el Ministerio de Ciencia e Innovación (MICINN).

El Proyecto EL USO DE LA TELEDETECCIÓN PARA UNA MEJOR COMPRESIÓN DEL CAMBIO CLIMÁTICO (PROMETEU/2009/086), financiado por la Generalitat Valenciana, Conselleria d'Educació i Ciència, en el marco del Programa Prometeo (I+D) para grupos de investigación de excelencia, y a su continuidad PROMETEU/2014/086

Los Proyectos PATRONES ECOLÓGICOS EN LAGUNAS MESETARIAS: LAS CLAVES PARA SU CONSERVACIÓN (ECOLAKE) (CGL2012-38909), MITIGACIÓN Y ADAPTACIÓN AL CAMBIO CLIMÁTICO EN LOS PRINCIPALES TIPOS DE HUMEDALES MEDITERRÁNEOS IBÉRICOS: BALANCES DE CARBONO Y MODELOS DE RESPUESTA DE ESPECIES Y HÁBITATS (CLIMAWET) (CGL2015-69557-R) y ESTIMACIÓN Y SEGUIMIENTO DE LA TEMPERATURA Y EMISIVIDAD DE LA SUPERFICIE TERRESTRE MEDIANTE LOS ÚLTIMOS AVANCES DE TELEDETECCIÓN (ESCLITEL) (CGL2013-46862-C2-1-P) y su subproyecto ESTIMACIÓN Y SEGUIMIENTO DEL BALANCE DE ENERGÍA EN SUPERFICIE Y DE LA CALIDAD DE AGUAS INTERIORES MEDIANTE LOS ÚLTIMOS AVANCES DE TELEDETECCIÓN (CGL2013-46862-C2-2-P), financiados por el Ministerio de Economía y Competitividad del Gobierno de España y por la Unión Europea a través del Fondo Europeo de Desarrollo Regional - FEDER "Una manera de hacer Europa".

La ayuda para la Formación del Personal Investigador (FPI) de carácter predoctoral, englobada en el marco del Plan Nacional de Investigación Científica, Desarrollo e Innovación Tecnológica 2008-2011 por el Ministerio de Ciencia e Innovación (MICINN), con efectos económicos y administrativos del 1 de agosto de 2011 al 31 de julio de 2015.

A la piñita

Deseo expresar aquí mi agradecimiento a las personas e instituciones que han hecho posible con su ayuda y apoyo la realización de esta Tesis Doctoral.

En primer lugar a mis directores, el Dr. Juan Manuel Sánchez y los Profesores Antonio Camacho y Vicente Caselles, quienes, además de ofrecerme la oportunidad de iniciarme en el mundo de la investigación, me han apoyado siempre para seguir adelante. Tengo que agradecerles también el tiempo que han dedicado en dirigir mi trabajo y la confianza que han depositado en mí.

Al Profesor José Antonio Martínez Lozano, director del departament de Física de la Terra i Termodinàmica, agradecerle su disposición a resolver siempre cualquier inconveniente.

A los Profesores Enric Valor y César Coll por la acogida en el grupo y la amabilidad mostrada en todo momento... *y cómo no, por hacer tan agradable el coffee time.*

A José Antonio Domínguez por haber dado la opción de poder iniciarme en este tema de investigación que tanto me gusta y disfruto.

A los doctores Carlos Rochera y Antonio Picazo por su amabilidad y toda la ayuda concedida tanto en temas de laboratorio como en haber resuelto las dudas que han ido apareciendo en este tema del que tanto tengo que aprender.

A la Dra. Raquel Nicolòs por su apoyo, sus consejos y también por haberme hecho más agradable la estancia por el departamento.

A mis compañeros de despacho: Rubén, Flor, M^a Jesús, Xumi, Jesús y Lluís- *Peplita del Turia*, gracias por ser tan buenos compañeros y por llenar de alegría y hacerme más llevadero el día a día. En especial quería agradecer al trio pokecalavera la paciencia que han tenido conmigo en este último periodo. A los que ya no están o no son tan asiduos, también agradecerles en todo en lo que me han ayudado y los buenos momentos que me han hecho pasar. También agradecer a los compañeros, presentes o pasados, de departamento y del campus que también han hecho más agradable mi estancia por estos mundillos: Emilio, Alex, Felipe, Pepelu, Álvaro, Sara, Carlos, Guillem, Facundo, Roberto, Níobe, Amparo, Drazen, Sergio...y muchos más. Quería dedicarle un espacio a Vicente, muchas gracias por todo xiquet, se te echa de menos por aquí...

Y a toda mi Gentucilla, no es necesario que entre en detalles, ¡ellos saben quiénes son! los haya nombrado ya o no. Querría hacer mención especial a mi Amooooool, a mi novio postizo, gracias, gracias y gracias...y a ti, sí sí, a ti también, gràcies.

Y como no, a la piñita: el tete, la mare y el papa. *GRACIAS*, y no hay palabras que describan lo agradecida que estoy con vosotros por todo, gracias por estar apoyándome, animándome en todo momento, hasta en los más críticos. Gracias a vosotros hoy puedo escribir estas palabras y me llena de felicidad que las podamos leer todos...*os quiero muchoOooO, como la trucha al truchoOooO.*

A todas las personas que han pasado por mi caminito, las que están y las que no, a todos vosotros:

;;;GRACIAS!!!

Picanya, noviembre 2016

Índice

NOTA INTRODUCTORIA	v
ABSTRACT	vii
MEMORIA: RESUMEN DE TRABAJOS	1
1. Introducción	1
1.1. Humedales. Proceso de eutrofización y comportamiento hidrológico.....	3
1.2. La teledetección en el control de masas de agua.....	5
1.3. Respuesta espectral del agua.....	8
1.3.1. Clorofila- <i>a</i>	9
1.3.2. Sólidos en suspensión.....	10
1.4. Antecedentes.....	11
1.5. Objetivos de la Tesis.....	14
2. Zona de estudio y datos empleados	17
2.1. Zonas de estudio.....	19
2.1.1. La Albufera de Valencia.....	19
2.1.2. La Mancha Húmeda.....	22
2.2. Medidas de campo y fuentes de datos de campo empleadas.....	25
2.2.1. Estudio de la calidad del agua.....	25
2.2.2. Estudio para estimación de área inundada.....	26
2.3. Imágenes de satélite.....	27

3. Metodología	33
3.1. Datos de campo.....	35
3.1.1. Concentración de clorofila- <i>a</i>	36
3.1.2. Concentración de sólidos en suspensión.....	37
3.1.3. Transparencia.....	38
3.2. Procesado de las imágenes de satélite.....	38
3.3. Fusión de datos.....	43
3.4. Métodos tradicionales. Modelos linealizables e índices.....	46
3.4.1. Estimación de parámetros de calidad de aguas.....	46
3.4.2. Estimación de área inundada	47
3.5. Aprendizaje automático. Algoritmos genéticos.....	49
4. Resultados y discusión	55
4.1. Datos de campo.....	57
4.2. Algoritmos para la estimación de parámetros de calidad de agua.....	59
4.2.1. Modelado y optimización.....	59
4.2.2. Aplicación y validación.....	62
4.3. Algoritmos para la estimación de área inundada.....	67
4.3.1. Modelado y optimización.....	67
4.3.2. Aplicación y validación.....	70
4.4. Elaboración de mapas temáticos. Seguimiento operativo.....	72
4.4.1. Seguimiento de la eutrofización en La Albufera de Valencia.....	72
4.4.2. Patrones hidrológicos en base al área inundada.....	76
4.5. Discusión.....	78

5. Conclusiones	83
Bibliografía	87
Anexo A: Artículo “Empirical relationships for monitoring water quality of lakes and reservoirs through multispectral images”	97
Anexo B: Artículo “Integrated satellite data fusion and mining for monitoring lake water quality status of the Albufera de Valencia in Spain”	107
Anexo C: Ponencia “Monitoring inland water quality status using images from the SPOT-5 take-5 experiment”	121
Anexo D: Artículo “Monitoring hydrological patterns of temporary lakes using remote sensing and machine learning models. Case study of La Mancha Húmeda Biosphere Reserve in Central Spain”	129

NOTA INTRODUCTORIA

La presente memoria resume el trabajo realizado por la doctoranda durante los últimos años, dedicados principalmente al estudio de calidad de aguas y área de inundación de lagunas.

La normativa de la Universidad de Valencia, según el reglamento sobre depósito, evaluación y defensa de la Tesis doctoral aprobado en el consejo de gobierno del 28 de junio de 2016, recoge dos posibilidades para la presentación de la Tesis doctoral: 1) la escritura de una Tesis tradicional y 2) la escritura de una Tesis por compendio de publicaciones. En este caso, la segunda opción ha sido la seleccionada. Para poder realizar la Tesis de acuerdo con esta opción, los requisitos a cumplir son:

- 1) La doctoranda debe presentar un mínimo de tres artículos, ya publicados o aceptados en revistas indexadas en algún índice internacional como JCR (WoS) y/o SJR (Scopus) y ser el primer firmante de cada uno de ellos.*
- 2) La Tesis debe incluir un resumen global de la temática, de los principales resultados y de las conclusiones.*
- 3) Ha de incluir una copia completa de los trabajos publicados o admitidos para publicación.*

De acuerdo con la normativa vigente en la Universidad de Valencia, la realización de esta Tesis ha considerado como base 4 publicaciones seleccionadas a tal efecto. Estos trabajos son:

1. “Empirical relationships for monitoring water quality of lakes and reservoirs through multispectral images”.

C. Doña; J. M. Sánchez; V. Caselles; J. A. Domínguez; A. Camacho.

IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol.7 (5):1632-1641. Mayo 2014.

2. “Integrated satellite data fusion and mining for monitoring lake water quality status of the Albufera de Valencia in Spain”.

C. Doña; N. B. Chang; B; V. Caselles; J. M. Sánchez; A. Camacho; J. Delegido; W. Vannah.

Journal of Environmental Management, vol.151:416-426. Enero 2015.

3. “Monitoring inland water quality status using images from the SPOT-5 take-5 experiment”.

C. Doña; J. M. Sánchez; V. Caselles; A. Camacho; A. Picazo; C. Rochera; J. M. Galve.

Proceedings ESA Living Planet Symposium 2016. Praga, Republica Checa. 9-13 mayo 2016 (ESA SP-740, agosto 2016).

4. "Monitoring hydrological patterns of temporary lakes using remote sensing and machine learning models. Case study of La Mancha Húmeda Biosphere Reserve in Central Spain".

C. Doña; N. B. Chang; B; V. Caselles; J. M. Sánchez; A. Camacho; L. Planells; M. Bisquert; V. García-Santos; S. Imen.

Remote Sensing, vol.8:618. Julio 2016.

La memoria que se presenta a continuación es un resumen de los trabajos citados anteriormente. Con el objetivo de ordenar y compactar las ideas, así como de orientar mejor al lector, se ha optado por dividir el contenido en diferentes capítulos, siguiendo los apartados de una Tesis tradicional.

Este trabajo está estructurado en cinco capítulos. En el primer capítulo se realiza una breve introducción sobre el tema tratado y las herramientas a utilizar, así como los objetivos que se pretenden conseguir. En el siguiente capítulo se habla de las diferentes zonas de estudio y de las bases de datos verdad-terreno empleadas, y también se citan algunas de las características de los sensores utilizados en el estudio. En el capítulo tercero se describe la metodología seguida para la obtención de variables de calidad de agua in situ y en el laboratorio, así como el estudio del área de inundación, y se explica cómo se ha realizado el tratamiento de las imágenes y el desarrollo de los diferentes algoritmos. El capítulo cuarto muestra los principales resultados obtenidos del análisis de los datos medidos, y de los algoritmos obtenidos, así como la discusión que se deriva de ellos. Por último, el capítulo quinto resume las conclusiones más relevantes de los resultados logrados en el presente trabajo.

Finalmente se incluyen los 4 anexos donde se recogen las diferentes publicaciones completas, las cuales se irán referenciando a lo largo de toda la Memoria.

Por otra parte, como se pretende optar a mención internacional del título de doctor, según el Artículo 9, del Reglamento sobre depósito, evaluación y defensa de la tesis doctoral, parte de la Tesis doctoral, al menos el resumen y las conclusiones deben estar redactados en inglés. Por este motivo, a continuación se presentará un abstract y el capítulo final de conclusiones estará escrito en inglés.

ABSTRACT

Wetlands contribute to a wealth of ecosystem services, including the regulation of the hydrological cycle for flood and drought control, and provide water supply, wildlife refuges, aesthetic enjoyment and recreational opportunities, among others. According to the Ramsar Convention, wetlands, in a broad sense, include all lakes and rivers, underground aquifers, swamps and marshes, wet grasslands, peatlands, oases, estuaries, deltas and tidal flats, mangroves and other coastal areas, coral reefs and all constructed sites, such as fish ponds, rice paddies, reservoirs and salt pans. Although these areas have a critical value to sustainable development, they are detrimentally impacted by urban growth, agricultural land reclamation and derived pollution.

The water quality and the ecological status of these aquatic ecosystems can deteriorate due to, among others, eutrophication. Lake eutrophication is a critical issue in the interplay of water supply, environmental management, and ecosystem conservation. Integrated sensing, monitoring, and modeling for a holistic lake water quality assessment and ecological status with respect to multiple constituents is in acute need.

On the other hand the water balance and hydrological variations are intimately tied to potential changes in a lentic ecosystem. Understanding the dynamics of water in lakes helps the goal of conservation and recovery of these valuable ecosystems. This is especially relevant given several environmental initiatives, such as the European Water Framework Directive (WFD), which came into force in 2000, and the Habitats Directive, delivered in 1992. These directives require each member state in the European Union to achieve a good ecological/conservation status for their water bodies and associated habitats and species, forcing the establishment of conservation actions.

Remote sensing techniques can be used to estimate water quality variables such as the concentration of chlorophyll-*a*, of total suspended particles, and water transparency. The first part of this Thesis describes empirical algorithms for the estimation of these variables using Landsat Thematic Mapper (TM) data. In this case, the ground data were taken from several Spanish lakes covering a variety of trophic statuses, ranging from oligotrophic to hypereutrophic. The studied lakes were la Albufera de Valencia and lakes and ponds of the Southeast Regional Park in Madrid. Empirical equations were obtained to estimate chlorophyll-*a* from the ratio in reflectance values between bands 2 and 4 of TM, transparency (Secchi disk) from reflectance in band 2, and total suspended particles from reflectance in band 4. The spectral equivalence between TM and Deimos-1 was also tested. By applying the proposed algorithms to this sensor, the temporal resolution is improved by up to 3 days, and this also increases spatial resolution to 22 m. The algorithms were validated using 3 Deimos-1 scenes of la Albufera de Valencia together with ground measurements. Results of this validation showed root mean square errors of 40 mg/m³ for chlorophyll-*a* concentration (Mean absolute difference percentage MADP = 22%), 10 mg/l for total suspended particles concentration (MADP=15%) and 0.10 m for SD (MADP=40%). Then, results were acceptable in terms of chlorophyll-*a* and total suspended particles concentration estimation with MADP values of $\pm 22\%$ and $\pm 15\%$, respectively. In any case, these results show the potential of Deimos-1 as a substitute of TM in water quality monitoring in small/medium water bodies, providing continuity to 3 decades of TM imagery.

Following this work, we developed an integrated algorithm for data fusion and mining of satellite remote sensing images to generate daily estimates of some water quality parameters of interest, such as chlorophyll-*a* concentrations and water transparency, to be applied for the assessment of the water quality of la Albufera de Valencia. In this case, remote sensing data from Moderate

Resolution Imaging Spectroradiometer (MODIS) and Landsat Thematic Mapper (TM) and Enhance Thematic Mapper (ETM+) images were fused to carry out an integrative near-real time water quality assessment on a daily basis. Landsat images are useful to study the spatial variability of the water quality parameters, due to its spatial resolution of 30 m, in comparison to the low spatial resolution (250/500 m) of MODIS. While Landsat offers a high spatial resolution, the low temporal resolution of 16 days is a significant drawback to achieve a near real-time monitoring system. This gap may be bridged by using MODIS images that have a high temporal resolution of 1 day, in spite of its low spatial resolution. Synthetic Landsat images were fused for dates with no Landsat overpass over the study area. Finally, with a suite of ground truth data, a few genetic programming (GP) models were derived to estimate the water quality using the fused surface reflectance data as inputs. The GP model for chlorophyll-*a* estimation showed a Root Mean Square Error (RMSE) = 8 mg/m³ (MADP=6%), and the GP model for water transparency estimation using Secchi disk showed a RMSE = 4 cm (MADP=9%). This methodology improves the traditional effort by means of linear regression techniques, as shown above. The spatio-temporal variations of water transparency and chlorophyll-*a* concentration can be then monitored simultaneously on a daily basis throughout the lake for environmental management.

Sentinel-2 offers the opportunity to continue with this inland water quality monitoring task, thanks to its 5-day revisit cycle and 10-30 m spatial resolution. In the framework of the Sentinel-2 preparatory activities, the ESA developed the SPOT5 take-5 experiment. From early April to the end of August 2015, SPOT-5 satellite was relocated in a 5-day orbit, before being decommissioned. Based on the spectral matching between both VNIR sensors, SPOT-5 was used to simulate Sentinel-2 products and show the benefits of its high spatial resolution to monitor small water bodies. Then, we developed algorithms to estimate the

water quality parameters from SPOT-5 images in three water bodies belonging to the Júcar river basin. Several experimental campaigns were carried out concurrent with SPOT-5 overpasses or close in date to accomplish this aim. Chlorophyll-*a* concentration, transparency and total suspended particles concentration were measured in la Albufera de Valencia. Again, genetic programming models were used to generate nonlinear regression equations between ground measurements and reflectance values from the SPOT-5 spectral bands. Results showed MADP values of $\pm 8\%$, $\pm 5\%$ and $\pm 10\%$ in the estimation of chlorophyll-*a* concentration, transparency and total suspended particles concentration, respectively. Focusing on la Albufera de Valencia Lake, results are similar to those already reported in the previous work. These results show the potential of Sentinel-2 to monitor and study the spatio-temporal trend of these water quality parameters.

On the other hand, remote sensing technologies also facilitate the content-based mapping over space and time, leading to multitemporal change detection of the hydrological variations in wetlands. Mapping surface water bodies allows for the investigation of water balance dynamics by providing information on the temporal and spatial variations of surface water coverage, this being especially relevant under the current climate change scenario.

The Biosphere Reserve of La Mancha Húmeda is currently the main wetland area in the Iberian Peninsula. This Lake District is one of the wetland complexes most threatened by anthropogenic activity, mainly by groundwater overexploitation due to excessive use for irrigated agriculture. This area is an important refuge for endangered waterfowl species, following the protection criteria for birds in Europe, and also holds endangered habitats. Because of its natural and ethnographic values, it was designated a Biosphere Reserve. This reserve comprises a set of temporary lakes, often saline, where water level fluctuates seasonally. Water inflows come mainly from direct precipitation and runoff of

small lake watersheds. Most of these lakes lack surface outlets and behave as endorheic systems, where water withdrawal is mainly due to evaporation, causing salt accumulation in the lake beds. Remote sensing was also used to estimate the temporal variation of the flooded area in these lakes and their associated hydrological patterns related to the seasonality of precipitation and evapotranspiration. Landsat 7 ETM+ images for the reference period 2013–2015 were jointly used with ground-truth datasets. Several inverse modeling methods, such as two-band and multispectral indices, single-band threshold, classification methods, artificial neural network, support vector machine and genetic programming, were applied to retrieve information on the variation of the flooded areas. Results were compared to ground-truth data, and the classification errors were evaluated by means of the kappa coefficient. Comparative analyses demonstrated that the genetic programming approach yielded the best results, with a kappa value of 0.98 and a total error of omission-commission of 2%. The dependence of the variations in the water-covered area on precipitation and evaporation was also investigated. The results show the potential of the tested techniques to monitor the hydrological patterns of temporary lakes in semiarid areas, which might be useful for management strategy-linked lake conservation and specifically to accomplish the goals of both the European Water Framework Directive and the Habitats Directive.

CAPÍTULO 1

Introducción

En este capítulo se va a introducir el concepto de humedal, cuáles son sus principales características, su importancia en el medio y por qué son esenciales para el bienestar de la humanidad, su problemática en cuanto a la calidad de sus aguas, y la necesidad de tener conocimiento de su comportamiento hidrológico para su conservación y recuperación. Seguidamente se explicará cómo se pueden estudiar estos sistemas mediante técnicas de teledetección y se mostrará la revisión bibliográfica correspondiente. Finalmente se expondrán los objetivos de la Tesis.

1.1. Humedales. Proceso de eutrofización y comportamiento hidrológico

Los humedales son áreas terrestres que están saturadas o inundadas de agua de manera estacional o permanente. Estas áreas aportan una gran cantidad de servicios ecosistémicos, incluyendo la regulación del ciclo hidrológico que ayuda al control de inundaciones y de la sequía, el suministro de agua, ser refugios de vida silvestre, y oportunidades de recreo, entre otros. De acuerdo con la Convención Relativa a los Humedales de Importancia Internacional especialmente como Hábitat de Aves Acuáticas, conocida como el Convenio de Ramsar, los humedales continentales incluyen ríos, lagos, acuíferos subterráneos, marismas, charcas, pantanos y llanuras de inundación. En cuanto a los humedales costeros, éstos incluyen marismas de agua salada, estuarios, manglares, albuferas o lagunas litorales, incluso arrecifes de coral. También hay humedales artificiales, como estanques piscícolas, arrozales y salinas. A pesar de que estas áreas tienen un valor fundamental para el bienestar de la humanidad, se están viendo perjudicadas por el crecimiento urbano, por la extensión de la agricultura y la contaminación que todo ello conlleva.

La calidad de las aguas y el estado ecológico de los ecosistemas acuáticos se pueden ver deteriorados por el proceso de eutrofización. Este proceso produce cambios biológicos, químicos y físicos en la masa de agua. Etimológicamente el término eutrófico significa “bien nutrido” y fue definido por primera vez por Naumann en 1919. Lo definió como “un incremento en el nivel de nutrientes, especialmente el nitrógeno y el fósforo”. Posteriormente se han ido añadiendo más definiciones, la mayoría de ellas con un factor en común, el enriquecimiento o aumento de nutrientes. Aunque es un fenómeno natural que se produce de forma muy lenta e irreversible, la presión antrópica ha causado que este proceso se acelere, recibiendo en este caso el nombre de eutrofización cultural. Este tipo de eutrofización se produce debido a aguas de desecho o residuales que los humanos vierten en los ecosistemas acuáticos.

La entrada de materia orgánica y de nutrientes, principalmente fósforo y nitrógeno en la zona eufótica de los ecosistemas leníticos, es decir, la zona en la que la luz penetra lo

suficiente para que dominen los procesos de fijación sobre los de consumo de carbono, altera temporalmente las condiciones de equilibrio de la masa de agua. Se produce un aumento de la turbidez y un cambio del color del agua hacia el verde debido al incremento de la producción primaria, es decir, la producción de las microalgas planctónicas. La fotosíntesis se activa con la proliferación del fitoplancton produciéndose una liberación de oxígeno tan intensa que las capas superficiales de la masa de agua quedan sobresaturadas y, en consecuencia parte de ese oxígeno es liberado a la atmósfera. La respiración de la materia orgánica, a su vez, consume oxígeno, con lo que se pueden producir episodios de hipoxia o anoxia cuando el consumo excede a los aportes (Domínguez, 2002).

Desde el punto de vista trófico la calidad de las masas de agua se puede clasificar desde un estado oligotrófico (bajo nivel de nutrientes) a un estado hipereutrófico (muy alto contenido de nutrientes). Se han desarrollado diferentes índices para poder realizar la ordenación de las masas de agua en función de su estado trófico, tales como el índice de Carlson (1974). En 1982 la OCDE (Organización para la Cooperación y el Desarrollo Económico) estableció una clasificación de acuerdo a los valores que alcanzan la concentración de fósforo total, la concentración de clorofila-*a* y la profundidad de visión del disco de Secchi (transparencia). En la tabla 1.1 se muestran los valores límites anuales establecidos por la OCDE (1982):

Categoría Trófica	TP mg/l	[Chl- <i>a</i>] media µg/l	[Chl- <i>a</i>] máxima µg/l	Secchi media m	Secchi mínimo m
Ultraoligotrófico	< 4,0	< 1,0	< 2,5	>12,0	> 6,0
Oligotrófico	< 10,0	< 2,5	< 8,0	>6,0	> 3,0
Mesotrófico	10 - 35	2,5 - 8	8 - 25	6 - 3	3 - 1,5
Eutrófico	35 - 100	8 - 25	25 - 75	3 - 1,5	1,5 - 0,7
Hipereutrófico	>100	>25	>75	<1,5	<0,7

Tabla 1.1. Valores límite (anuales) para la clasificación trófica de la OCDE considerando el contenido en fósforo [TP], la concentración de clorofila-*a* [Chl-*a*] y la profundidad de visión del disco de Secchi.

Además del fósforo total, en la tabla 1.1 se puede observar que otro de los parámetros utilizados como indicador de la calidad del agua es la concentración de clorofila-*a*. Esta

se utiliza como indicador de la biomasa del fitoplancton, el cual muestra una relación positiva con el incremento de nutrientes. Otro de los parámetros utilizados es la profundidad del disco de Secchi, con el que se puede estimar la profundidad de la zona fótica de la masa de agua, es decir, hasta donde penetra la luz solar en cantidad suficiente para permitir la dominancia de los procesos productivos, y por lo tanto donde hay actividad fotosintética significativa. Además, la medida del disco de Secchi nos proporciona información sobre el estado de eutrofia del sistema ya que en los procesos de eutrofización la falta de transparencia va ligada al crecimiento algal que provoca turbidez en el agua.

En el año 2000 entró en vigor la Directiva Marco del Agua (Directiva 2000/60/CE) y en 1992 se adoptó la Directiva de Hábitats. Ambas directivas establecen que cada estado miembro de la Unión Europea debe lograr un buen estado ecológico/estado de conservación de sus recursos hídricos, hábitats y especies asociadas, obligando a establecer ciertas acciones de conservación. Además, establecen que se deben preservar las zonas protegidas, teniendo en cuenta los datos científicos y técnicos disponibles.

Para poder aplicar las herramientas necesarias para la conservación y recuperación de estos valiosos sistemas también se necesita de la comprensión de la dinámica del agua (Castañeda *et al.*, 2009), ya que el balance de agua y las variaciones hidrológicas en los humedales están íntimamente ligados a los cambios potenciales que se producen en estos sistemas.

La implantación de la Directiva Marco del Agua y Hábitats ha implicado que se realicen numerosos estudios de calidad y detección de agua sobre estos ecosistemas, con el objetivo de su cumplimiento en los plazos acordados.

1.2. La teledetección en el control de masas de agua

La teledetección es una herramienta útil para poder estimar diferentes parámetros de interés en numerosos campos de investigación, entre ellos el estudio de la variabilidad de la cobertura u otras características del agua y el estudio de la calidad de masas de agua. Con este tipo de tecnologías se puede estudiar la superficie de interés sin ningún

tipo de contacto con ella. Una de las formas de realizar detección remota es mediante sensores que están a bordo de satélites en órbita alrededor de la Tierra. Estos sensores registran la radiancia reflejada y emitida por los elementos de la superficie terrestre y de la atmósfera en una amplia zona del espectro electromagnético (fig. 1.1).

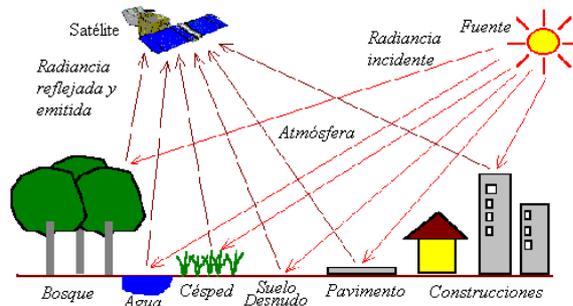


Figura 1.1. Esquema de radiancias medidas desde satélite.

(Fuente: Centre for Remote Imaging, Sensing & Processing (CRISP) (modificado))

Una de las utilidades de este tipo de técnicas es la de poder realizar un seguimiento de las masas de agua sin necesidad de muestreos *in-situ*. La teledetección facilita el monitoreo en espacio y tiempo, dando lugar a una detección del cambio multitemporal de las variaciones de los parámetros de interés, en nuestro caso las variaciones hidrológicas de los humedales y de los parámetros de calidad del agua, permitiendo el estudio de la dinámica del balance hídrico y las variaciones temporales y espaciales de la cobertura de agua y de los parámetros de calidad en superficie.

En la actualidad existe un amplio abanico de satélites para poder realizar este tipo de estudios. Dependiendo de lo que se quiera abordar se utiliza un sensor u otro atendiendo a las características que estos ofrecen, en combinación con nuestros intereses. A modo de resumen, las características que distinguen a los diferentes sensores son: la frecuencia entre la adquisición de una imagen y la siguiente en una misma zona (resolución temporal), el tamaño del pixel (resolución espacial), el número y anchura de bandas o regiones del espectro en las que el sensor puede medir (resolución espectral), y finalmente la sensibilidad del sensor, a mayor precisión radiométrica mayor detalle se podrá captar en la imagen (resolución radiométrica).

En los últimos años las técnicas de teledetección se han utilizado para estudiar diferentes parámetros de calidad de agua (Serrano *et al.*, 1997; Härmä *et al.*, 2001; Peña *et al.*, 2004; Duan *et al.*, 2006; Allan *et al.*, 2011; McCullough *et al.*, 2012). Las imágenes del sensor Thematic Mapper (TM) a bordo del satélite Landsat-5 han sido ampliamente utilizadas en este tipo de estudios debido a su resolución espacial de 30 m en las bandas del visible e infrarrojo cercano y al amplio archivo histórico de imágenes disponible que permite tener un conocimiento de la masa de agua desde los años 80. Con todo esto se pueden desarrollar estudios multitemporales de detección de cambio de los diferentes parámetros (por ejemplo, Mayo *et al.*, 1995; Alle y Johnson, 1999; Olmanson *et al.*, 2008). Sin embargo, Thematic Mapper tiene una resolución espectral limitada y esto, junto con su baja resolución temporal de 16 días, hace que sea difícil estimar alguno de los parámetros de calidad de agua de forma precisa. Sensores como MODIS (Moderate Resolution Imaging Spectroradiometer) o MERIS (MEdium Resolution Imaging Spectrometer), que tienen una mejor resolución temporal-espectral, son más adecuados para monitorear los diferentes parámetros de calidad (Peña *et al.*, 2004, McCulloch *et al.*, 2012); pero en este caso, la resolución espacial de sus imágenes es demasiado baja (250-500 m) para analizar los cuerpos de agua de pequeño-medio tamaño (Mancino *et al.*, 2009). Otros estudios más recientes han fusionado imágenes MODIS y Landsat, aprovechando su equivalencia espectral, para poder resolver este problema y obtener así imágenes con la resolución espacial de Landsat a una mejor resolución temporal (Chang *et al.*, 2014).

En los estudios de detección de agua también se han utilizado imágenes de la familia Landsat, QuickBird, o SPOT-4 (Satélites Pour l'Observación de la Terre o Satélites Para la Observación de la Tierra), cuya resolución espacial es apta para estimar la superficie del agua de lagos (Bai *et al.*, 2011; McFeeters, 2013), deltas o cuencas hidrográficas (Sakamoto *et al.*, 2013; Li *et al.*, 2013; El-Asmar *et al.*, 2011), la extensión de lagunas temporales y permanentes en zonas áridas (Campos *et al.*, 2012), y líneas de costa (Maglione *et al.*, 2014).

1.3. Respuesta espectral del agua

Las diferentes coberturas de la superficie terrestre se comportan de forma distinta a la hora de reflejar la energía que les incide (fig. 1.2). Además, en un mismo cuerpo la distribución de la energía reflejada varía con las diferentes longitudes de onda (signatura espectral), y presentará máximos y mínimos de reflectividad característicos de cada material. La medida de estas curvas desde satélite permite extraer gran cantidad de información sobre la cobertura de la superficie (fig. 1.2).

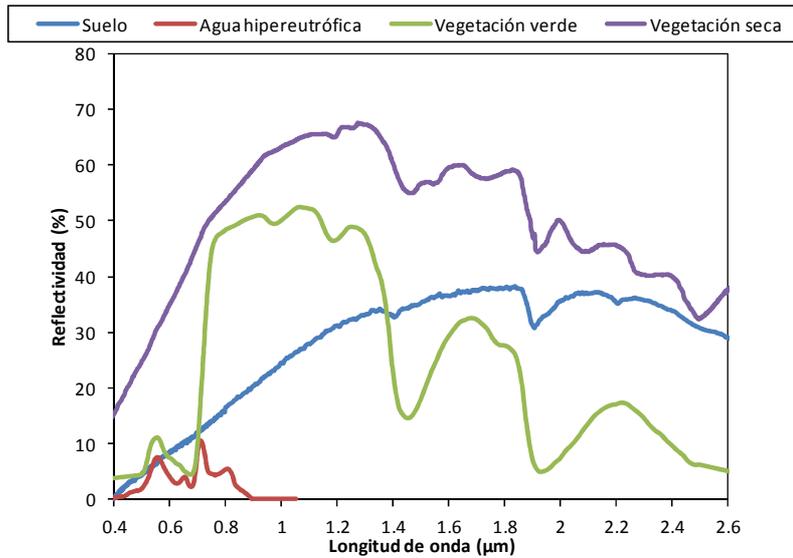


Figura 1.2. Signatura espectral de diferentes superficies.

En la figura 1.2 se puede observar que el agua presenta los valores de reflectividad más bajos respecto al resto de superficies. Esto se debe a que el agua presenta una elevada transmisión en el espectro visible y absorción en el infrarrojo. El agua presenta un espectro característico con valores máximos de reflectividad en la región del espectro del azul, pudiéndose explicar así su color, reduciéndose poco a poco hasta prácticamente hacerse nula en el infrarrojo cercano y medio. Esta última característica espectral permite discriminar las masas de agua del resto de superficies.

Las masas de agua presentan un carácter tridimensional, es decir, tienen una contribución proveniente del fondo, de la columna de agua y de la superficie, que está

reflejado en su respuesta espectral, por lo que las variaciones en cada uno de estos componentes alteran también su respuesta espectral.

En el estudio de calidad de agua nos vamos a centrar en los componentes que alteran la superficie y la columna de la masa de agua, como pueden ser la concentración de partículas en suspensión y clorofila-*a*. A estos componentes se les conoce como Constituyentes Ópticamente Activos (OAC) y pueden ser discriminados de la signatura espectral del agua debido a su interacción espectral selectiva.

1.3.1. Clorofila-*a*

El espectro de la clorofila-*a* presenta dos picos de absorción en el intervalo espectral del visible, uno localizado entre 400-500 nm y el otro entre 600-700 nm, zonas correspondientes al azul y rojo del espectro. Además, presenta un máximo de reflectividad en el verde, motivo de su color.

Si se analiza el espectro de diferentes concentraciones de clorofila-*a* en agua pura medida en un tanque de laboratorio (fig. 1.3), se puede observar que la curva espectral presenta el máximo de reflectividad característico de la clorofila-*a* en la zona del verde del espectro (550 nm) y un máximo relativo aproximadamente en 700 nm que podría corresponder a un mínimo de absorción de las algas en agua (Domínguez, 2002).

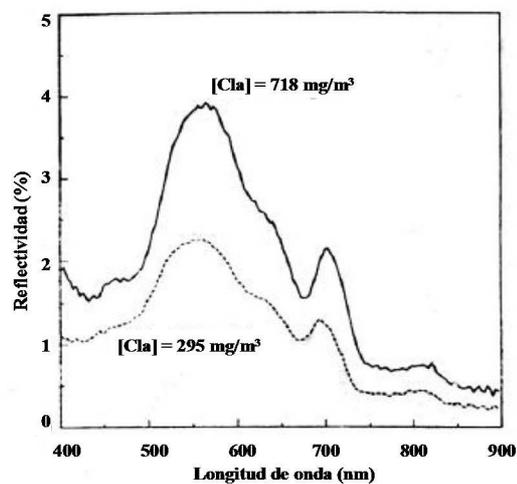


Figura 1.3. Curvas de reflectividad de concentración de clorofila-*a* disuelta en agua pura (Han *et al.*, 1994).

Igualmente cuando se estudia el espectro de la clorofila-*a* en aguas naturales, éste varía de tal forma que la reflectividad del agua disminuye en el azul y aumenta en el verde (fig. 1.4). Esto facilita la medida de la abundancia algal mediante técnicas de teledetección.

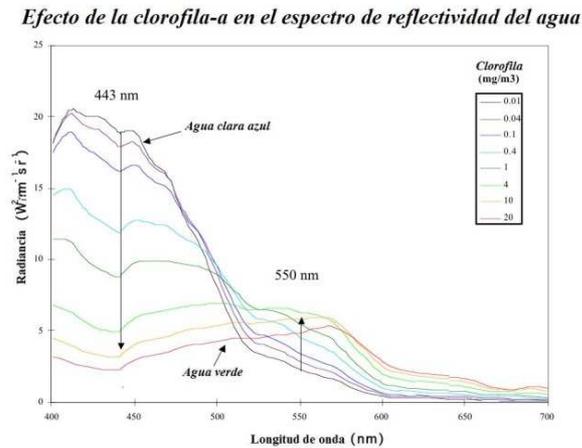


Figura 1.4. Efectos de la clorofila en el espectro de reflectividad del agua. (Fuente: Coastal Ocean Observing Centre)

1.3.2. Sólidos en suspensión

La reflectividad del agua aumenta proporcionalmente con el aumento de la concentración de cualquier tipo de sólidos y partículas en la parte del espectro comprendida entre 400 y 1000 nm. El incremento de la reflectividad y la variación de la signatura espectral dependerán del tamaño de la partícula y de las propiedades del sólido en cuestión. Por ejemplo, con la disminución del tamaño de partícula la reflectividad aumentará y si las partículas son de origen orgánico ésta dependerá también de sus características cromáticas. Así, en el caso del fitoplancton, la reflectividad aumentará en la zona del verde y disminuirá en el azul (ver 1.3.1). Si se trata de partículas inorgánicas de origen arcilloso, a más concentración de estas partículas la reflectividad aumentará, especialmente en la zona del rojo (fig. 1.5).

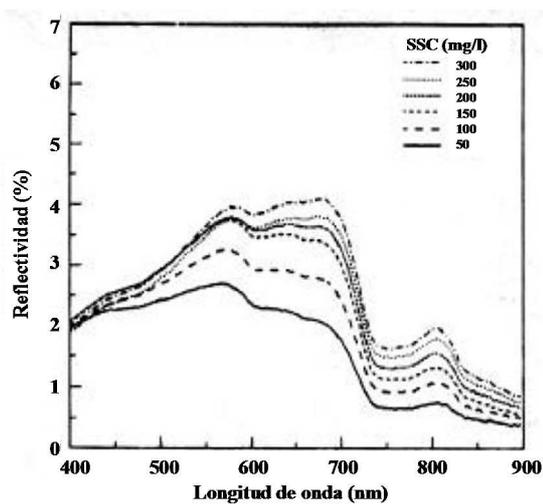


Figura 1.5. Variación de la reflectividad en función de la concentración de sólidos en suspensión (SSC) (mezcla de barro y arcilla) (Hand-Rundsquist, 1994).

1.4. Antecedentes

A lo largo de los últimos años se han realizado numerosos trabajos de aplicación de la teledetección para el estudio de calidad de masas de agua y para la estimación del área inundada. En este apartado nos vamos a centrar, en primer lugar, en algunos de los trabajos que han establecido algoritmos entre los parámetros de calidad de agua en superficie, y la información de reflectividad de las bandas del sensor Thematic Mapper (TM) de Landsat, y posteriormente presentaremos diferentes técnicas que se han utilizado para la estimación de área inundada.

Caselles *et al.* (1986) llevaron a cabo un estudio de calidad de agua en la Albufera de Valencia en el que concluyeron que las reflectividades corregidas de las bandas TM1 y TM3 eran las que mejor permitían determinar los parámetros indicativos de la biomasa: clorofila-*a* y partículas en suspensión, pero no obtuvieron resultados satisfactorios para la estimación de la transparencia.

Mayo *et al.* (1995) realizaron un estudio de la distribución de clorofila-*a* en el lago Kinneret en Israel. En su trabajo proponen un índice compuesto por las reflectividades corregidas: $(TM1-TM3)/TM2$ para la estimación de la concentración de clorofila-*a*.

También se han utilizado algoritmos polinomiales para la estimación de la concentración de clorofila-*a* y transparencia, como en el estudio publicado por Allee *et al.* (1999), los cuales propusieron un algoritmo que relacionaba las radiancias de las bandas TM1, 2, 3 y 5 para la estimación del logaritmo natural de la concentración de clorofila-*a*, y las radiancias de las bandas TM2, 3 y 5 para la estimación de la transparencia. Más adelante Giardino *et al.* (2001) llevaron a cabo un estudio de determinación de variables de calidad del agua en lagos sub-alpinos de Italia. Estos autores encontraron buenas correlaciones con las reflectividades corregidas de las bandas TM1 y TM2 para la obtención de valores de transparencia y de concentración de clorofila-*a*. Además, algunos autores también han realizado estudios temporales. Por ejemplo, Kloiber *et al.* (2002) obtuvieron un algoritmo de estimación de transparencia con las bandas TM1 y TM3 con datos de 500 lagos en Minnesota, EEUU, y utilizando 13 imágenes de satélite llevaron a cabo un estudio de transparencia de los lagos del área en un periodo de 25 años. Más adelante Bilge *et al.* (2003) realizaron un estudio en el embalse Porsuk de Turquía donde desarrollaron modelos de regresión múltiple entre las cuentas digitales de las bandas TM1, 2, 3 y 4 y parámetros de calidad de agua como la concentración de partículas en suspensión y de clorofila-*a*. Estos autores emplearon datos de dos décadas de 13 estaciones situadas en el embalse.

Algunos investigadores también han utilizado medidas espectrales realizadas *in situ* para la construcción de algoritmos de determinación de la concentración de clorofila-*a*, como puede ser el estudio publicado por Duan *et al.* (2007) donde, a partir del análisis de los espectros obtenidos en el lago Chagan (China), encontraron un algoritmo utilizando el cociente entre las cuentas digitales TM4/TM3. Como en el estudio anterior, Domínguez *et al.* (2009) también analizaron los espectros, en este caso los obtenidos en los lagos del Parque Regional del Sureste de Madrid. Tras analizar los resultados, optaron por utilizar la reflectividad de la banda TM2 para la determinación de la transparencia y de la concentración de clorofila-*a*, y la reflectividad corregida de la banda TM4 para la estimación de las partículas en suspensión. En ese trabajo no se obtuvieron resultados satisfactorios para la determinación de clorofila-*a*. También se han realizado estudios de evaluación de calidad de agua en lagos de pequeño tamaño

como es el caso del trabajo realizado por Mancino *et al.* (2009) en los lagos de Monticchio, Italia. Estos autores encontraron algoritmos para la estimación de la concentración de clorofila-*a* y de la transparencia con técnicas de regresión múltiple entre las reflectividades corregidas de las bandas TM1, 2 y 3. Allan *et al.* (2011) realizaron un estudio de determinación de concentración de clorofila-*a* en los lagos de Rotoura, en Nueva Zelanda, y concluyeron que a partir del sensor TM se puede realizar un seguimiento de clorofila-*a* de las masas de agua utilizando el logaritmo natural de TM3 y disponer así de una herramienta para poder controlar la calidad del agua en zonas de difícil acceso humano. Para el estudio de calidad, aparte de la utilización de regresiones simples y múltiples, también se ha recurrido a otro tipo de técnicas, como la utilizada por Chang *et al.* (2014) con los algoritmos genéticos.

En cuanto a la estimación de área inundada, uno de los primeros estudios fue publicado por Work y Wilmer en 1976, en el cuál identificaron estanques y lagos utilizando una sola banda o combinaciones de bandas de Landsat 1. A este primer estudio le siguieron unos cuantos más. McFeeters (1996) propuso el índice de diferencia normalizada de agua (NDWI) para delinear las características de los cuerpos de agua abiertos. Este índice utiliza las bandas del infrarrojo cercano (NIR) y del verde del sensor TM a bordo de Landsat 4 y 5. Los valores positivos de este índice corresponden a los píxeles de agua en la imagen. Más adelante Gao (1996) propuso un NDWI modificado para estudiar el agua en vegetación. Este índice se compone de las bandas del NIR y del infrarrojo de onda corta (SWIR). Más tarde, Xu (2006), propuso una modificación adicional del NDWI incluyendo las bandas del verde y del SWIR para el estudio de las características de los cuerpos de agua abiertos.

También se han utilizado bandas individuales y transformaciones de imágenes como la transformación “tasseled cap” (Jain *et al.*, 2005, Ouma *et al.*, 2006). Esta transformación supone pasar de las 6 bandas ópticas de Landsat a 3 bandas con un claro sentido físico que representan el brillo de la imagen (albedo), el verdor y la humedad. En el estudio de Ouma *et al.* (2006) se utilizó la componente humedad. Lira (2006) utilizó variantes del Análisis de Componentes Principales (PCA), otra técnica aplicada en la transformación de la imagen, para estimar la extensión de aguas abiertas con

imágenes de MODIS, ASTER y Landsat. Más recientemente Fisher *et al.* (2013) publicaron el índice de agua obtenido por análisis discriminante lineal (LDAWI) compuesto por las bandas del verde, rojo, NIR y SWIR de las imágenes SPOT-5 para estimar la superficie de un lago. También se han utilizado métodos de clasificación temática clásicos, tales como el método de máxima probabilidad utilizado para estimar la superficie cubierta por agua en humedales (Gardelle *et al.*, 2009 y Soliman *et al.*, 2011), métodos basados en los clasificadores por píxeles o en los basados en objetos (Sun *et al.*, 2012). Feyisa *et al.* (2014) también proponen el Índice de Extracción de Agua Automatizado (AWEI) el cuál funcionó mejor en zonas de montaña, donde la sombra causada por el terreno es una importante fuente de error a la hora de realizar la clasificación. Algunos métodos aplicados a la vigilancia humedales combinan la banda NIR o la del SWIR para estimar los píxeles cuya superficie está cubierta por agua y el NDVI para estudiar la vegetación de los humedales (Moser *et al.*, 2014, Bustamante *et al.*, 2006).

1.5. Objetivos de la Tesis

El propósito principal de esta Tesis es el de obtener las herramientas necesarias para poder llevar a cabo el seguimiento de calidad y dinámica del agua en ecosistemas leníticos, usando como modelos un conjunto de humedales situados en la Península Ibérica.

En nuestro caso los humedales estudiados son la Albufera de Valencia y un conjunto de lagunas situadas en la Reserva de la Biosfera de La Mancha Húmeda. En cada una de nuestras zonas se ha realizado un tipo diferente de estudio dependiendo de sus principales necesidades de gestión. Por tanto, los principales objetivos de esta Tesis son dos, por un lado desarrollar algoritmos empíricos para la estimación de parámetros de calidad del agua que sean aplicables a sistemas acuáticos como la Albufera de Valencia, mientras que el segundo objetivo es el de desarrollar una metodología para poder estimar el área inundada en el conjunto de lagunas objeto de estudio en La Reserva de la Biosfera de La Mancha Húmeda. Los pasos seguidos para cumplir estos objetivos son:

Objetivo 1

- i) Realizar campañas de campo en la Albufera de Valencia para ampliar la base de datos y procesar las respectivas imágenes TM. (*Anexo A*)
- ii) Procesar imágenes Landsat y Deimos-1 para el entrenamiento y validación de los modelos obtenidos. (*Anexo A*)
- iii) Obtener imágenes sintéticas Landsat a partir de la fusión de imágenes Landsat y MODIS mediante el algoritmo STAR-FM, y así resolver la escasa resolución temporal de 16 días de Landsat. (*Anexo B*)
- iv) Desarrollar algoritmos con métodos clásicos, regresiones simples, regresiones múltiples, métodos de aprendizaje automático, y algoritmos genéticos, para la estimación de parámetros de calidad de aguas con datos de las imágenes utilizadas. (*Anexos A, B y C*)

Objetivo 2

- i) Realizar campañas de campo en el conjunto de lagunas de la Biosfera de la Reserva de la Mancha Húmeda. (*Anexo D*)
- ii) Comparar diferentes metodologías para la obtención de área inundada a partir de imágenes Landsat. (*Anexo D*)
- iii) Validar la metodología propuesta. (*Anexo D*)

CAPÍTULO 2

Zonas de estudio y datos utilizados

En este capítulo se describe, en primer lugar, la localización y principales características de las diferentes masas de agua objeto de estudio. Seguidamente se presentan las diferentes bases de datos de campo empleadas y, finalmente, se detallan algunas de las características de los sensores utilizados.

2.1. Zonas de estudio

Esta Tesis se ha realizado a partir de datos de 14 masas de agua localizadas en dos áreas diferentes. Una de ellas es la Albufera de Valencia, dónde se ha realizado el estudio de calidad del agua, mientras que las demás masas de agua se encuentran situadas en la Reserva de la Biosfera de La Mancha Húmeda, dónde se ha realizado el estudio del área inundada.

2.1.1. La Albufera de Valencia

El Parque Natural de la Albufera de Valencia constituye uno de los espacios naturales más representativos e interesantes de la Comunidad Valenciana, y se ha reconocido su valor en el ámbito nacional e internacional. La Albufera de Valencia es un lago somero situado en la costa mediterránea de Valencia, España (39° 20 'N, 0° 20'W) (fig. 2.1). Es el mayor lago de la Península Ibérica, siendo su superficie de 23,2 km² con una profundidad media de 1,2 m (Romo *et al.*, 2008). Después de, junto con su entorno, ser reconocido como Parque Natural en 1985 y ser incluido en la lista Ramsar de humedales protegidos en 1989, el lago ha sido objeto de numerosos estudios limnológicos (Vicente y Miracle, 1992).

El lago es un antiguo golfo marino, cerrado por un cordón litoral. Pertenece en su totalidad al término municipal de Valencia, por el este, la separa del Mediterráneo una estrecha franja arenosa de unos 1200 metros cubierta de pinares de pino blanco y monte bajo conocida como la Dehesa del Saler, que constituye el dique natural del lago. El lago está conectado con el mar Mediterráneo mediante tres canales denominados "golas" (El Pujol, El Perellonet y El Perelló), que regulan mediante compuertas el nivel de las aguas del lago y cuya gestión está vinculada con las necesidades del ciclo del cultivo de arroz, ya que éste es el cultivo dominante de la zona. El ciclo hidrológico de la Albufera está íntimamente ligado con el cultivo del arroz, por lo tanto la gestión del agua depende en gran medida de estas operaciones. El manejo del cierre y apertura de las compuertas de las "golas" está a cargo de la Junta de Desagüe de la Albufera, constituida desde el 20 de julio de 1862.

Las compuertas se cierran como norma general el día 1 de noviembre, inundando los arrozales, y se vuelven a abrir el 1 de enero. Este día comienza el desagüe natural de campos y acequias que se complementa a finales de enero y principios de febrero, con la ayuda de motores y bombas. Coincide entonces el nivel más bajo de las aguas y la desecación total de los campos de arroz para llevar a cabo laboreos imprescindibles para el cultivo. A finales de abril o primeros de mayo vuelven nuevamente a cerrarse las compuertas, los campos se inundan de nuevo, esta vez con menos agua, procediéndose a una labor previa a la siembra, que se realiza a principios de mayo. En junio, julio y agosto se produce la fase de pleno cultivo y a mediados de agosto se secan ya los campos para la recolección de arroz a principios de septiembre. El nivel de la Albufera continúa bajo hasta principios de noviembre, en que comienza de nuevo el ciclo anual.

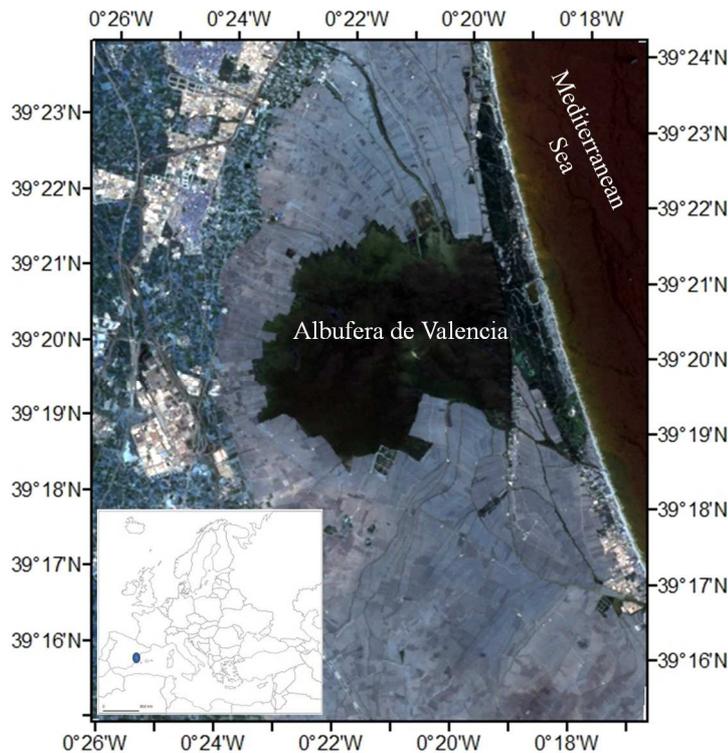


Figura 2.1. Imagen en color verdadero del sensor TM (RGB=321) de la Albufera de Valencia.

Las aportaciones a la Albufera provienen principalmente de los canales de riego que traen las aguas dulces de los ríos Turia y Júcar, varios barrancos y surgencias denominadas ullals (Soria *et al.*, 2002). También recibe agua residual tratada de las zonas urbanas e industriales cercanas. Incluso hoy en día, a pesar de las mejoras en los sistemas de alcantarillado, pequeñas cantidades de aguas residuales no tratadas son vertidas en este cuerpo de agua. A partir del reconocimiento del parque en régimen jurídico como Parque Natural por decreto de la Generalitat Valenciana de 8 de julio de 1986 y su incorporación cuatro años más tarde a las lista de humedales de importancia internacional designada por el Gobierno español, como hábitat de aves acuáticas (Soria *et al.*, 2002), la Albufera ha sido motivo de muchos estudios, que en menor o mayor medida intentan contribuir a la mejora de su estado en todos los ámbitos. A nivel estatal, mediante el Plan Hidrológico de la Cuenca del Júcar, se ha intentado mejorar la calidad del agua del lago, aportando caudales externos, aunque no llegando a los objetivos propuestos por la Directiva Marco del Agua (2000/60/CE).

Los núcleos de población que rodean al lago han ido experimentado un fuerte crecimiento (consecuencia del desarrollo agrario, turístico e industrial), aumentando la presión urbanística, humana e incrementando la contaminación sobre el espacio natural, principalmente sobre el lago y la calidad de sus aguas. La intensificación de la agricultura y la transformación de cultivos tradicionales han generado conflictos en relación con el agua, y el aterramiento de zonas húmedas. Debido a todos los impactos asociados principalmente a los cambios del uso de suelo, la Albufera de Valencia se ha convertido en un sistema hipereutrófico, mostrando valores de concentración de clorofila-*a* superiores a 100 mg/m³, y valores de transparencia de agua por debajo de 20 cm (Soria y Vicente, 2002). El alto nivel trófico se mantiene hoy en día no sólo por los aportes de nutrientes externos, sino también por la fuerte carga interna de nutrientes acumulados en los sedimentos del lago. Las diferentes fuentes de entrada de nutrientes y las diferencias en los depósitos sedimentarios de nutrientes a lo largo del lago, junto con el efecto del viento y las corrientes, determinan una cierta heterogeneidad espacial en la distribución de las variables de calidad del agua tales como la concentración de clorofila-*a* y la transparencia del agua. A diferencia de otros sistemas hipereutróficos, la

Albufera tiene la particularidad de disponer de las compuertas o “golas” que posibilitan la renovación periódica de sus aguas. Estos flujos dentro del lago determinan la evolución anual del sistema (Caselles *et al.*, 1986). Esta variabilidad no se puede controlar con suficiente resolución espacial-temporal a través de técnicas de control limnológicas, lo que justifica el uso de técnicas de teledetección.

2.1.2. La Mancha Húmeda

La otra zona de estudio se centra en una serie de lagunas situadas en la Reserva de la Biosfera de la Mancha Húmeda, una región rica en humedales, la más extensa de la Península Ibérica, que comprende hasta 30.000 hectáreas albergando humedales y lagunas (Laguna *et al.*, 2013) distribuidos dentro de las provincias de Albacete, Ciudad Real, Cuenca y Toledo, en la región de Castilla-La Mancha (fig. 2.3).

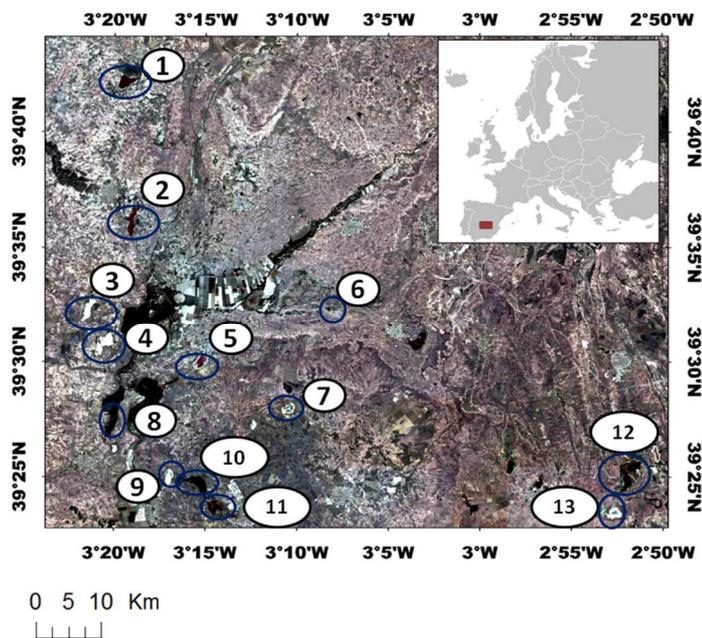


Figura 2.2. Imagen en color verdadero del sensor TM (RGB=321) de la zona de estudio de La Mancha Húmeda. Las lagunas estudiadas son: 1) El Longar, 2) Larga de Villacañas, 3) Tírez, 4) Peñahueca, 5) Grande de Quero , 6) Mermejuela, 7) Salicor, 8) Grande de Villafranca , 9) Las Yeguas, 10) Camino de Villafranca, 11) La Veguilla, 12) Manjavacas, y 13) Alcahozo.

En 1981, la UNESCO designó el área como Reserva de la Biosfera de la Mancha Húmeda dentro del Programa El Hombre y la Biosfera (MAB), un programa científico para promover la mejora de las interacciones entre las personas y sus entornos. En 2014 esta Reserva de la Biosfera fue reconocida como una de las zonas de humedales más grandes de Europa, y muchos de sus humedales están incluidos en la Red Natura 2000 (Directiva Europea de Hábitats y Aves) y la Convención de Ramsar. Situada en una zona muy llana (La Mancha), la Reserva de la Biosfera incluye llanuras de inundación y, sobre todo, una gran variedad de lagunas de agua salada, en su mayoría endorreicas, es decir que no evacuan cantidades significativas de agua ni por desagüe superficial ni por infiltración, sino que lo realizan fundamentalmente por evaporación. Esta zona también representa una de las principales áreas de lagos salinos de Europa.

Para este estudio se ha seleccionado un conjunto de 13 lagunas salinas poco profundas (tabla 2.1), todas ellas situadas entre 638 y 690 m sobre el nivel del mar.

Tabla 2.1. Principales características de las lagunas objeto de estudio.

Laguna* ¹	Coordenadas °N	Coordenadas °W	Altitud (masl)	Área Max ^{*2} (km ²)	Estacionalidad
<i>Alcahozo (13)</i>	39,38	2,85	660	0,19	Alta
<i>Camino de Villafranca (10)</i>	39,41	3,26	638	1,4	Moderada
<i>Grande de Quero (5)</i>	39,50	3,24	664	0,9	Alta
<i>Grande de Villafranca (8)</i>	39,46	3,33	645	0,9	Baja
<i>La Veguilla (11)</i>	39,39	3,24	638	0,19	Baja+
<i>Larga de Villacañas (2)</i>	39,62	3,34	660	1,4	Baja+
<i>Las Yeguas (9)</i>	39,41	3,28	638	0,6	Alta
<i>El Longar (1)</i>	39,70	3,32	690	0,3	Alta
<i>Manjavacas (12)</i>	39,41	2,86	670	1,4	Moderada+
<i>Mermejuela (6)</i>	39,54	3,13	660	0,09	Alta
<i>Peñahueca (4)</i>	39,51	3,34	650	1,6	Alta
<i>Salicor (7)</i>	39,47	3,17	668	0,6	Alta
<i>Tírez (3)</i>	39,54	3,36	650	1,3	Alta

*¹ Entre paréntesis la codificación de las lagunas (indicadas en figura 2.2)

*² Fuente: Confederación hidrográfica del Guadiana (<http://www.chguadiana.es/>)

+ Estas lagunas reciben aguas residuales tratadas que aumentan artificialmente el suministro de agua.

La mayoría son lagunas temporales que se encuentran dentro de las cuencas de los ríos Záncara y Cigüela. Estas lagunas se enclavan en paisajes agrícolas, y tienen cuencas muy pequeñas que normalmente albergan cultivos de vid o cereal. Estas lagunas pueden presentar un cinturón marginal de plantas halófilas y, algunas de ellas, también presentan vegetación helofítica (p. ej. Manjavacas y La Veguilla) en zonas influenciadas por los vertidos de aguas residuales depuradas en las que la salinidad del agua es mucho más baja. Los sedimentos del área son de naturaleza continental, el relieve es llano y en su litología dominan calizas (Gosálvez *et al.*, 2012). La mayoría de las lagunas carecen de salidas superficiales y se comportan como sistemas endorreicos, con la evaporación como el principal proceso de extracción de agua, causando la acumulación de sales en los lechos de las lagunas. El nivel de salinidad de estas lagunas abarca el intervalo mesosalino a hipersalino.

Los hidroperiodos de las lagunas fluctúan, aunque éstas son en su mayoría temporales, como resultado de su mezcla de patrones climáticos mediterráneo-continental semiáridos (Laguna *et al.*, 2013), los cuales pueden, por ejemplo, ser ilustrados por el patrón de inundación de la Laguna de Alcahozo en un intervalo de tiempo específico (fig. 2.3). El clima en esta zona muestra una estación seca pronunciada en verano, con una precipitación promedio anual de 400-500 mm (Vidal *et al.*, 2013).

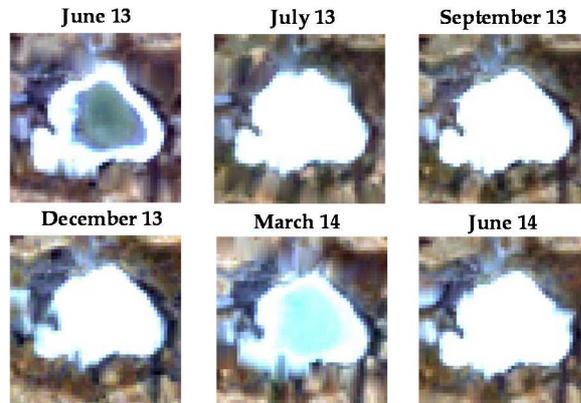


Figura 2.3. Imágenes en color verdadero del sensor ETM+ (RGB=321) de la evolución temporal de la laguna de Alcahozo para el período junio 2013-junio 2014.

Las lagunas de La Mancha Húmeda han sufrido un progresivo deterioro debido a la alteración de sus patrones hidrológicos y la contaminación, el primero vinculado principalmente a la continua sobreexplotación de los acuíferos de los últimos 50 años debido al aumento de la demanda agrícola (Florín *et al.*, 1999) o al ser destino de aguas residuales que han reducido la concentración salina de sus aguas y las han eutrofizado. Muchas de las lagunas de la Reserva de la Biosfera se drenaron durante el siglo XX. Mientras que otras, debido a sus niveles de salinidad y la imposibilidad de uso agrícola, se utilizaron como vertederos o para eliminar las aguas residuales de los pueblos cercanos.

2.2. Medidas de campo y fuentes de datos de campo empleadas

En este apartado se explican las diferentes medidas de campo realizadas y las fuentes de datos utilizadas para las dos partes del estudio, calidad de agua y área de inundación de las lagunas.

2.2.1. Estudio de la calidad del agua

En cuanto al estudio de calidad de agua los datos de campo utilizados son los de las concentraciones de clorofila-*a* (mg/m^3), y de partículas en suspensión (mg/l) y la transparencia del agua (m).

La concentración de clorofila-*a*, [Chl-*a*], se emplea para estimar la biomasa algal. La medida de la concentración de partículas en suspensión, [TSS], en aquellos medios en los que el fitoplancton es mayoritariamente su constituyente, es otra forma de estimar la biomasa, aunque incluye, además de células vivas del plancton, restos celulares y productos de degradación, así como partículas inorgánicas en suspensión. Por último, la transparencia, medida con el disco de Secchi (SD), mide la penetración de la luz en la masa de agua.

Los datos de campo proceden de las siguientes fuentes de datos:

- Datos de campañas experimentales de campo realizadas en el 2006, 2010/2011, y en 2015 en la Albufera de Valencia (*Anexos A, B y C*).

El estudio se completa con datos de campo procedentes de:

- Datos de campañas de campo tomados del trabajo de Caselles *et al.* (1986), en la Albufera de Valencia (*Anexo A*).
- Datos de campañas de campo cedidos por el Centro de Estudios Hidrográficos (CEH) del CEDEX de las lagunas de las Madres, el Campillo y el Porcal localizadas en el Parque Regional del Sureste, Madrid (*Anexo A*).
- Datos de campañas de campo realizadas en los embalses de Tous y Escalona en 2015 (*Anexo C*).

2.2.2. Estudio para estimación de área inundada

Los datos utilizados para el estudio de la estimación de área inundada consisten en (tabla 2.2):

Tabla 2.2. Fuentes de datos empleadas para cada laguna, A= datos verdad-terreno de la columna de agua, B= datos GPS y C= imágenes Earth Explorer™ ©Digital globe.

(Las abreviaturas de las lagunas son: Alcahozo-AL, Camino de Villafranca-CA, Grande de Quero-QUE, Grande de Villafranca-GRAN, La Veguilla-VE, Larga de Villacañas-LAR, Las Yeguas-YE, Longar-LO, Manjavacas-MAN, Mermejuela-MER Peñahueca-PE, Salicor-SAL, Tírez-TI).

Fecha	Fuente	Lagunas
09/02/11	C	LO/YE/CA/VE/LAR/PE/QUE/SAL
29/06/12	C	VE/LAR/PE
15/07/12	C	CA/SAL
28/02/13	A	AL/CA/MAN
29/05/13	A	AL/CA/LO/VE/MAN/GRAN/LAR/MER/PE/QUE/TI
23/07/13	A	CA/VE/YE/QUE/TI
28/10/13	A	PE/TI
25/11/13	A	AL/CA/LO/LAR/SAL
16/12/13	A	CA/LO/GRAN/LAR/PE/TI
27/01/14	A	CA/VE/MAN/YE/PE/TI
29/04/14	A	CA/MER
24/06/14	A	LO/VE/YE/GRAN/LAR/MER/QUE/SAL
23/07/14*	A	AL/CA/LO/VE/MAN/YE/MER/PE/TI
22/02/15	B	AL/MAN
21/05/15	B	AL/MAN

*Inspección de campo en AL/CA/LO/VE/MAN/YE

- Datos de profundidad de la columna de agua de las 13 lagunas estudiadas. Estos datos se midieron en los limnómetros situados en cada una de las lagunas durante el periodo de muestreos 2013/2014 del proyecto ECOLAKE. Una campaña de campo más extensiva se llevó a cabo en julio de 2014 en las lagunas de Alcahozo, Camino de Villafranca, El Longar, La Veguilla, Las Yeguas, y Manjavacas, para tener un mejor conocimiento del estado de estas lagunas.
- Datos del perímetro medido in situ por GPS de las lagunas de Alcahozo y Manjavacas en el año 2015.
- Perímetros de las lagunas digitalizados manualmente a partir de imágenes de alta resolución espacial de Google Earth TM.

2.3. Imágenes de satélite

Las imágenes empleadas en esta Tesis pertenecen a los satélites Landsat 5 y 7, Terra, Deimos-1 y SPOT-5. Todos estos satélites tienen órbita heliosíncrona, es decir, que se desplazan en órbitas generalmente circulares y polares (el plano de la órbita es paralelo al eje de rotación de la Tierra) de modo que, aprovechando el movimiento de rotación terrestre, puede captar imágenes de diferentes puntos cada vez que pase por el mismo punto de la órbita. Estas órbitas son posibles entre 300 y 1500 km de altura, en el caso de Landsat 5, 7 y MODIS está situado a 705 km de altura, Deimos-1 a 686 km y SPOT-5 a 822 km.

Los satélites Landsat-5 y 7 pertenecen a la misión espacial LANDSAT, la cual se inició en marzo de 1972 con el lanzamiento de Landsat-1. El sensor Thematic Mapper (TM) a bordo de Landsat-5 fue lanzado el 1 de marzo de 1984. El 30 de noviembre de 2011 se detuvo la adquisición global de imágenes Landsat-5 debido a problemas en la transmisión de datos a tierra, causados por la degradación de un componente electrónico. Este sensor tenía un periodo de revisita de 16 días. Operaba en 6 bandas en óptico, de las cuales 3 son en el visible y 3 en el infrarrojo con una resolución espacial de 30 m, además de una banda en el infrarrojo térmico con 120 m de resolución espacial (tabla 2.3). Características muy similares presenta el sensor Enhanced Thematic Mapper + (ETM+) a bordo de Landsat-7, operativo desde abril de 1999 hasta

el día de hoy, a excepción de la mejora de la banda del infrarrojo térmico, que pasa a 60 m de resolución espacial.

Otro de los sensores empleados es el Moderate Resolution Imaging Spectroradiometer (MODIS) a bordo de la plataforma Terra. Este satélite fue lanzado en 1999 y contiene 36 bandas espectrales de diferentes resoluciones espaciales, 250 m, 500 m y 1000 m y un periodo de revisita diario. Aunque fue diseñado para tener una vida operativa de 6 años, todavía sigue operativo.

También se han utilizado en esta Tesis algunas imágenes de Deimos-1, también conocido como Spain-DMC 1. Se trata del primer satélite español, y está operado por Deimos Imaging S.L. El satélite forma parte de la contribución española a la constelación de satélites para el seguimiento de catástrofes cuyas siglas en inglés son DMC. Deimos-1 fue lanzado el 29 de julio de 2009 y opera en 3 bandas espectrales situadas en el óptico, 2 de ellas en el visible y otra en el infrarrojo próximo cuya resolución espacial es de 22m. En cuanto a la resolución temporal, tiene un periodo de revisita de 3 días (tabla 2.3).

Por último, también se han utilizado imágenes SPOT-5. En el marco de las actividades preparatorias del nuevo Sentinel-2, la Agencia Espacial Europea (ESA) desarrolló el experimento SPOT-5 take 5. Desde principios de abril hasta finales de agosto de 2015, la órbita de SPOT-5 fue reubicada, antes de su desconexión, a una órbita cuyo periodo de revisita era de 5 días. Aprovechando la equivalencia espectral de ambos sensores, se utilizó SPOT-5 para simular los productos del reciente Sentinel-2. Las imágenes simuladas consisten en 2 bandas pertenecientes al visible y otras 2 al infrarrojo con una resolución temporal de 5 días y una resolución espacial de 10 m.

Son varios los motivos que nos llevaron a centrar el estudio en todos estos sensores. En primer lugar, y más importante, la media-alta resolución espacial de Landsat, Deimos-1 y SPOT-5 permite estudiar masas de agua de pequeño-medio tamaño, lo que no sería posible con otros sensores de menor resolución espacial. Además, las imágenes Landsat están disponibles gratuitamente vía online en la web del United States Geological Survey (USGS) (<http://glovis.usgs.gov/>), y a través del Plan Nacional de Teledetección

(PNT). Un aspecto clave también es que el extenso archivo histórico disponible de estas imágenes, permitiría realizar estudios de evolución de las masas de agua desde 1984.

A continuación se muestran las principales características de los sensores utilizados en cada una de las partes de nuestro estudio de calidad.

En la primera parte de la Tesis, correspondiente a la publicación mostrada en el *Anexo A*, se trabajó con imágenes Landsat-5-TM e imágenes Deimos-1, ya que Deimos-1 guarda una cierta equivalencia espectral con Landsat-5-TM (tabla 2.3), aunque mejorando su resolución espacial y temporal. Esto posibilita aprovechar los modelos ya existentes elaborados para TM y extender su aplicación.

Para la realización del trabajo presentado en el *Anexo A* se dispuso de 11 imágenes, 3 TM y 8 Deimos-1, todas ellas pertenecientes a la zona de la Albufera de Valencia. Las imágenes TM se descargaron de la web de la USGS, a excepción de la imagen de la Albufera del año 1985 que fue facilitada por el Grupo de Teledetección Térmica de la Universidad de Valencia (GTT). Las fechas de las imágenes TM utilizadas en este trabajo son de las siguientes fechas: 22/07/1985, 20/07/2010 y 05/08/2010.

Tabla 2.3. Características de Landsat5-TM y Deimos-1.

Satélite	Banda	Resolución espectral (μm)	Resolución espacial (m)	Resolución temporal (días)	Resolución radiométrica (bits)	Área de cobertura (km^2)
Landsat-5	1	0,452-0,518 (visible-azul)	30	16	8	185x170
	2	0,528-0,609 (visible-verde)	30			
	3	0,626-0,693 (visible-rojo)	30			
	4	0,776-0,904 (IR cercano)	30			
	5	1,567-1,784 (IR medio)	30			
	6	10,45-12,42 (IR térmico)	120			
	7	2,097-2,349 (IR medio)	30			
Deimos-1	2	0,520-0,600 (visible-verde)	22	3	8	300x300
	1	0,630-0,690 (visible-rojo)	22			
	0	0,770-0,900 (IR cercano)	22			

En el caso de las imágenes Deimos-1, éstas fueron proporcionadas por Deimos Imaging S. L., en el marco del proyecto “Monitoring the Trophic State of the Albufera Lake using Deimos-1 Images” en el que el GTT se encontraba involucrado (tabla 2.4).

Tabla 2.4. Fechas de las imágenes Deimos-1 disponibles para este trabajo.

Deimos-1	2010					2011		
	19/07	07/08	19/10	4/11	9/12	23/01	12/04	12/07

En el caso de la zona de estudio complementaria, correspondiente a las masas de agua del Parque Regional del Sureste (Comunidad de Madrid), los datos de las imágenes fueron proporcionados directamente por el Centro de Estudios Hidrográficos; en este caso información de reflectividad en las bandas 2 y 4 del sensor TM, con fechas coincidentes con los datos de campo, también facilitados para este estudio.

En la segunda parte de la Tesis, correspondiente a la publicación presentada en el *Anexo B*, se trabajó con imágenes Landsat-5-TM, Landsat-7-ETM+, e imágenes Terra-MODIS.

Tabla 2.5. Características de Landsat y MODIS.

SENSOR	Bandas fusionadas	Resolución espectral (μm)	Resolución temporal (días)	Resolución espacial (m)
Landsat TM/ETM+	1	0,452 – 0,518	16	30
	2	0,528 – 0,609		
	3	0,626 – 0,693		
	4	0,776 – 0,904		
	5	1,567 – 1,784		
	7	2,097 – 2,349		
MODIS TERRA	3	0,459 – 0,479	1	500
	4	0,545 – 0,565		500
	1	0,620 – 0,670		250
	2	0,841 – 0,876		250
	6	1,628 – 1,652		500
	7	2,105 – 2,155		500

Mientras que las imágenes Landsat ofrecen una resolución espacial media, la baja resolución temporal de 16 días es un gran inconveniente para poder realizar el control de calidad de las masas de agua. Este problema se puede resolver mediante el uso de imágenes MODIS que tienen una alta resolución temporal de 1 día, a pesar de su baja resolución espacial. La solución pasa por realizar una fusión de los datos de las imágenes MODIS y Landsat aprovechando la equivalencia espectral entre los sensores y así poder obtener imágenes sintéticas de Landsat a una mejor resolución temporal. En

la tabla 2.5 se muestran las principales características de las bandas fusionadas de estos sensores.

Como se ha mencionado anteriormente, las imágenes Landsat se descargaron del servidor Earth Explorer de la USGS, y las imágenes MODIS se descargaron de la web MODIS reprojection tool (<https://mrtweb.cr.usgs.gov/>). En total se han utilizado un conjunto de 28 imágenes, (19 MODIS y 9 Landsat). Las imágenes descargadas corresponden a las coincidentes con las campañas de campo más las imágenes adicionales necesarias para llevar a cabo la fusión de las imágenes (tabla 2.6).

Tabla 2.6. Fechas de imágenes Landsat y MODIS utilizadas en el estudio.

DOY	22/06	24/06*	13/07*	17/07	09/08	11/08*	25/08	27/08	28/08	30/08	08/09	16/09	17/09	19/09	16/09*	5/10
2006																
MODIS	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
LANDSAT	X			X	X		X							X	X	X

DOY	09/04	12/04*	26/04
2011			
MODIS	X	X	X
LANDSAT	X		X

*Días con datos *in situ*

En la última parte del estudio de calidad, correspondiente al trabajo presentado en el *Anexo C*, se trabajó con las imágenes simuladas SPOT-5. Los productos se descargaron de <https://spot-take5.org/> dentro del proyecto SPOT-5 take 5 de la ESA. Todos los datos descargados están procesados por el centro de datos terrestres THEIA y son distribuidos con una política libre y abierta a través de la articulación de la web de la ESA-CNES, portal de spot-take5.org. Las imágenes utilizadas en este trabajo corresponden a las fechas 9, 14 y 24 de julio de 2015.

En la parte de estimación de área inundada, correspondiente a la publicación presentada en el *Anexo D*, también se recurrió a imágenes del sensor Enhanced Thematic Mapper (ETM+) a bordo de Landsat-7. Se utilizaron un total de 15 imágenes (tabla 2.2): una imagen coincidente con una determinación precisa mediante trabajo de campo para realizar el estudio comparativo de todos los métodos testeados, y las catorce restantes para evaluar la precisión del método que mejores resultados proporcionó comparando

los resultados obtenidos con los datos de referencia. Los productos utilizados en este caso se descargaron también del servidor Earth Explorer de la USGS.

CAPÍTULO 3

Metodología

En este capítulo se describe la metodología empleada en los muestreos realizados en el lago de la Albufera de Valencia y los embalses de Tous y Escalona, y en el posterior tratamiento para la obtención de la concentración de clorofila-*a* y de partículas en suspensión, en el laboratorio. También se detalla el procesado y tratamiento de las imágenes de satélite, así como las técnicas utilizadas para el desarrollo de los diferentes algoritmos de estimación de parámetros de calidad del agua y de área inundada.

3.1. Datos de campo

Para el desarrollo de esta Tesis se han empleado medidas de campo tomadas en el marco de varias campañas de muestreo llevadas a cabo en la Albufera de Valencia, en los años 2006, 2010, 2011 y 2015, y en los embalses de Tous y Escalona en 2015. Todos estos muestreos se realizaron de forma coincidente, o muy próxima en el tiempo, con el paso de los sensores utilizados. En las campañas del 2010, 2011 y 2015 se seleccionaron diferentes puntos de muestreo alejados de la orilla y abarcando la totalidad de las masas de agua (fig. 3.1). Los datos de la Albufera de 2006 fueron cedidos por el CEH-CEDEX.

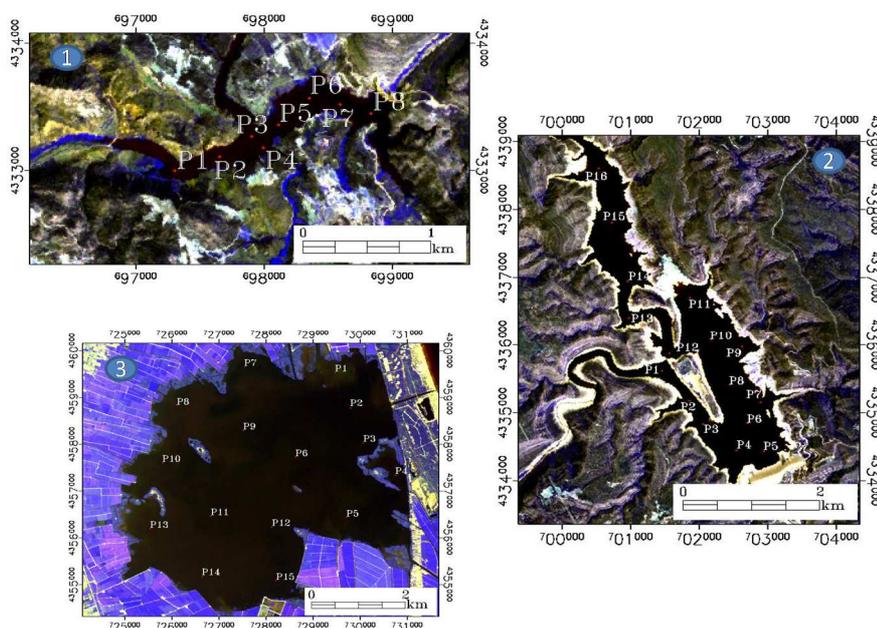


Figura 3.1. Localización de los puntos de muestreo en los embalses de: 1) Escalona y 2) Tous, para la campaña de 2015, y 3) la Albufera de Valencia para las campañas del 2010, 2011 y 2015.

Los días de medida para las campañas del 2010, 2011 y 2015 (tabla 3.1) se seleccionaron en base al calendario de paso del satélite por la zona y también a las condiciones meteorológicas.

El protocolo de muestreo en cada punto consistió en recoger entre 0,5-1 l de agua en superficie, al tiempo que se anotaban sus coordenadas GPS. Una vez rotuladas, las botellas con las muestras se introducían rápidamente en una nevera, de forma que las muestras permaneciesen refrigeradas hasta el momento de su posterior análisis en el laboratorio. Finalmente también se midió la transparencia. Esto fue así a excepción de las dos primeras fechas de 2010 en que se realizó un proceso de filtrado “*in situ*” de agua para la posterior obtención de concentración de clorofila-*a* en el laboratorio.

3.1.1. Concentración de clorofila-*a*

En el caso de las campañas de 2010, 2011 y 2015 la concentración de clorofila-*a* se determinó en el laboratorio filtrando las muestras de agua hasta la saturación del filtro, utilizando para ello filtros de fibra de vidrio Whatman GF/F (fig.3.2-a). Después, cada uno de los filtros se introdujo en un tubo de ensayo, se añadieron 5 ml de acetona al 90% en cada uno, de manera que el filtro quedara totalmente sumergido, y se dejó durante unas 24 horas a una temperatura de -20°C.

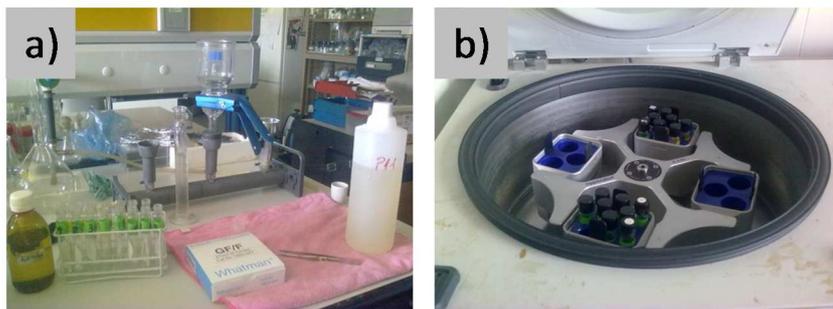


Figura 3.2. a) Montaje del proceso de filtrado y b) de centrifugado de las muestras de agua.

Transcurridas las 24 horas, el extracto obtenido se centrifugó (fig.3.2-b) para depositar todas las partículas presentes en el fondo del tubo. El espectrofotómetro se calibró con acetona al 90%, tras lo cual se midieron las absorbancias a diferentes longitudes de onda. La clorofila-*a* se determinó a partir de estos valores de absorbancia mediante el método tricromático de Strickland y Parsons (1968), adaptado de la forma descrita en Picazo et al. (2013):

$$[\text{Chl-a}] = (11,6 A_{665} - 1,31 A_{645} - 0,14 A_{630}) (\nu \nu Z) \quad (3.1)$$

donde A_{665} , A_{645} y A_{630} son los valores de absorción a las longitudes de onda de 665, 645 y 630 nm, respectivamente, v es el volumen del extracto en ml (5 ml de acetona al 90% en nuestro caso), V es el volumen de la muestra de agua filtrada en litros, y z es el paso óptico de la cubeta espectrofotométrica en cm (1 cm en nuestro caso).

En el caso de la campaña del 2005 las concentraciones de clorofila-*a* se midieron mediante fluorometría *in situ* con una posterior re-calibración con datos de laboratorio (Domínguez et al., 2011).

3.1.2. Concentración de sólidos en suspensión

Los sólidos en suspensión se determinaron en el laboratorio a través del método gravimétrico (diferencia de pesada). Para cada muestra se utilizó un crisol y un filtro de fibra de vidrio AH-934, que se pesaron con una balanza de precisión. Al igual que en el caso anterior, las muestras de agua se filtraron hasta la saturación del filtro, y se tomó nota del volumen filtrado. El conjunto de crisoles + filtros se introdujeron durante 5 horas en una estufa a una temperatura de 103°C hasta que el filtro perdió el agua y solamente quedó el material en suspensión en él. A continuación se volvieron a pesar y se calculó la diferencia de pesos (fig.3.3). La concentración de las partículas en suspensión viene determinada por esa diferencia de pesos dividida por el volumen de agua filtrado en l (APHA-AWWA-WEF 1992).



Figura 3.3. Determinación de partículas en suspensión.

3.1.3. Transparencia

La medida de la transparencia del agua se realizó *in situ* usando un disco de Secchi (SD). Se trata de un disco de aproximadamente 20 cm de diámetro dividido en cuadrantes alternantes negros y blancos (fig.3.4). Se considera la medida a la profundidad a la que el disco deja de ser visible desde la superficie cuando se sumerge en la masa de agua. Se trata de la profundidad de visión, y por lo tanto es una medida bastante subjetiva.



Figura 3.4. Imagen del Disco de Secchi.

3.2. Procesado de las imágenes de satélite

Las imágenes de los diferentes sensores se procesaron atendiendo a las necesidades del estudio en cuestión. Las imágenes precisan de un nivel de procesado u otro en función del producto descargado. Todo el procesado de las imágenes se realizó con un software de tratamiento digital de imágenes. A continuación, se detalla el procesado aplicado a las imágenes de satélite en cada uno de los trabajos que conforman esta Tesis.

En la primera parte de la Tesis, correspondiente al trabajo presentado en el *Anexo A*, se trabajó con productos de Landsat-TM (L1T) y Deimos-1. Ambos productos vienen corregidos geométricamente y contienen información de cuentas digitales (DN), por lo que se necesita llevar a cabo ciertas correcciones hasta llegar al producto deseado. Lo primero que se hizo fue convertir las DN en una magnitud con significado físico como es la radiancia espectral del sensor, $L_{\lambda, \text{Sat}}$ ($\text{W} \cdot \text{m}^{-2} \cdot \text{sr}^{-1} \cdot \mu\text{m}^{-1}$). Las ecuaciones de calibración para esta conversión son las siguientes:

$$L_{\lambda Sat} = G_{rescale} \cdot DN + B_{rescale} \quad (3.2)$$

$$L_{\lambda Sat} = (DN/G_{rescale}) + B_{rescale} \quad (3.3)$$

donde $G_{rescale}$ y $B_{rescale}$ son los coeficientes de calibrado radiométrico del sensor para cada banda. En el caso de TM se aplicó la ecuación 3.2, cuyos coeficientes se tomaron del estudio de Chander *et al.* (2009). En el caso de las imágenes Deimos-1, la ecuación que se aplicó es la 3.3, y los coeficientes de calibrado, diferentes para cada imagen, venían indicados en la cabecera de éstas.

Tras la calibración se dispone de valores de radiancia en el techo de la atmósfera (TOA) para cada banda espectral. Estas radiancias se transformaron a continuación a reflectividades TOA mediante la expresión:

$$\rho_{\lambda TOA} = L_{\lambda Sat}/L_{\lambda Sol} \quad (3.4)$$

donde

$$L_{\lambda Sol} = (E_{\lambda 0} \cdot \cos\theta_z)/k \cdot \pi \quad (3.5)$$

$L_{\lambda Sol}$ es la radiancia espectral solar ($W \cdot m^{-2} \cdot sr^{-1} \cdot \mu m^{-1}$), $E_{\lambda 0}$ es la irradiancia solar espectral extraterrestre ($W \cdot m^{-2} \cdot \mu m^{-1}$), θ_z es el ángulo cenital solar y k es el factor corrector de la distancia Tierra-Sol:

$$k = [1 + 0,0167 \cdot (\text{sen}(2\pi(\text{día} - 93,5))/365)]^2 \quad (3.6)$$

Los valores de $E_{\lambda 0}$ son diferentes para cada banda espectral y cada sensor, en el caso de TM se tomaron los valores del estudio de Chander *et al.* (2009), y para el caso de Deimos-1 se tomaron los datos el documento técnico disponible en el servidor del PNT (DEIMOS-1: Payload Description and Data Characteristics).

Como bien se indica con el subíndice TOA (Techo de la Atmósfera), la radiancia recibida por el sensor no se corresponde con la que proviene de la superficie, ya que la señal recibida por el sensor está afectada por la contribución de la capa atmosférica interpuesta entre la superficie y éste. Para obtener la magnitud física de la superficie objeto de estudio es necesario corregir esta perturbación atmosférica.

La reflectividad de la superficie medida desde satélite ($\rho_{\lambda sup}$), que asumiendo una superficie Lambertiana, es decir, que refleja por igual en todas las direcciones, se puede aproximar a (Gilabert *et al.*, 1994):

$$\rho_{\lambda sup} = \frac{k \cdot \pi \cdot (L_{\lambda Sat} - L_{\lambda p})}{\tau_{\lambda v} \cdot (E_{\lambda 0} \cdot \cos \theta_z \cdot \tau_{\lambda z} + E_{\lambda b})} \quad (3.7)$$

donde $L_{\lambda p}$ es la radiancia de camino, resultado de la interacción de la radiación electromagnética con los diferentes componentes atmosféricos ($W \cdot m^{-2} \cdot sr^{-1} \cdot \mu m^{-1}$), $E_{\lambda b}$ es la irradiancia solar difusa ($W \cdot m^{-2} \cdot sr^{-1}$), $\tau_{\lambda v}$ es el coeficiente de transmisividad atmosférica en el camino Tierra-sensor y $\tau_{\lambda z}$ en el camino Sol-Tierra. En la figura 3.5 se muestra el esquema de las interacciones que intervienen en la señal captada por el sensor.

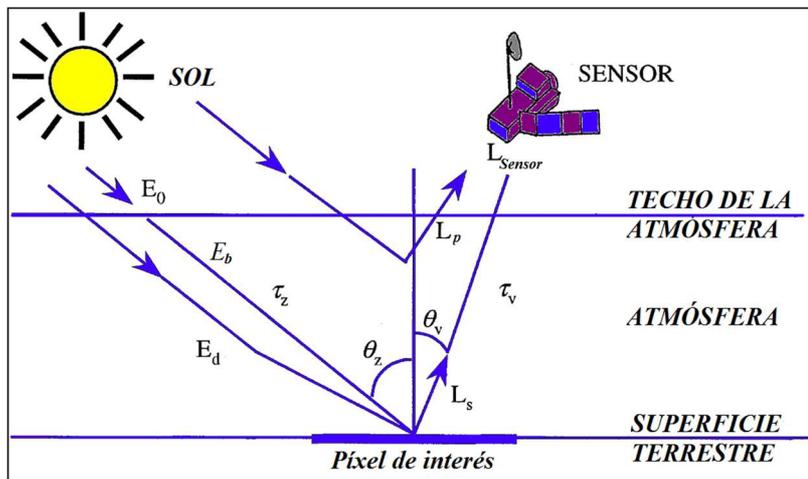


Figura.3.5. Radiancia recibida por el sensor e interacciones atmosfera-radiación electromagnética.

En esta parte de la Tesis se han aplicado diferentes métodos de corrección atmosférica para cada uno de los sensores empleados. Los métodos utilizados se basan en la metodología del objeto oscuro. Esta técnica es aplicable sobre imágenes que contengan zonas de reflectividad baja, como puede ser el agua o sombras. De acuerdo con esta metodología los valores mínimos de la imagen corresponden a píxeles cuyos valores deben ser cero y el valor que tiene se debe a la contribución de la atmósfera.

Para las imágenes TM se ha utilizado el método propuesto por Gilabert *et al.* (1994). Estos autores utilizan funciones analíticas mediante las cuales, a partir de los valores del pixel oscuro de las bandas TM1 y TM3 se pueden calcular los parámetros atmosféricos que intervienen en la ecuación 3.7, y finalmente calcular la reflectividad de la superficie. Tanto el agua como la vegetación presentan bajos valores de reflectividad en las bandas TM1 y TM3, siendo estas suficientes para poder estimar la cantidad y el tipo de aerosoles que hay presentes en la atmósfera.

Las imágenes Deimos-1 no se corrigieron por este método, ya que este sensor no posee banda en intervalo azul del espectro. En este caso se empleó el método de Chávez (1996), tal y como propone el Plan Nacional de Teledetección (PNT) (Chuvieco, E. y Hantson, S., 2010). Según este método los términos de $E_{\lambda b}$ y $\tau_{\lambda v}$ de la ecuación 3.7 se ignoran y la radiancia de camino se calcula a partir del valor del pixel oscuro de cada banda espectral del sensor.

Para la estimación del pixel oscuro se utilizó como zona oscura un área de mar alejada de la costa, que se encuentra dentro de los límites de las imágenes procesadas. El valor del pixel oscuro se determinó como el primer valor del histograma del área que sobrepasara el 1% de los píxeles.

En la segunda parte del estudio, correspondiente a la publicación mostrada en el *Anexo B*, se utilizaron imágenes de Landsat y MODIS. En este caso el producto MODIS descargado fue el de reflectividad de superficie MOD-09GA / GQ con 500/250 m de resolución espacial, mientras que para Landsat, al igual que en el caso anterior, se trató del producto L1T, con 30 m de resolución espacial. Con el fin de homogeneizar criterios, en esta parte de la Tesis la corrección atmosférica de las imágenes Landsat se realizó siguiendo el mismo método de corrección del producto MODIS descargado. Para ello se empleó el código de transferencia radiativa 6S (Vermonte *et al.*, 1997, Masek *et al.*, 2006).

Las imágenes MODIS se remuestrearon al tamaño de pixel de Landsat, es decir, a 30 m, empleando la técnica del vecino próximo, que es la que menos cambia el valor del pixel

original. Además, las imágenes de ambos sensores se reproyectaron a la zona UTM 31N. Por otra parte, se aplicó un método de interpolación bilineal (ESRI, 2013) con el fin de resolver el problema técnico en el corrector de barrido lineal (SLC) presente en el sensor ETM+ desde 2003. En la figura 3.6 se muestra un esquema de los pasos seguidos en el procesado de las imágenes de ambos sensores:

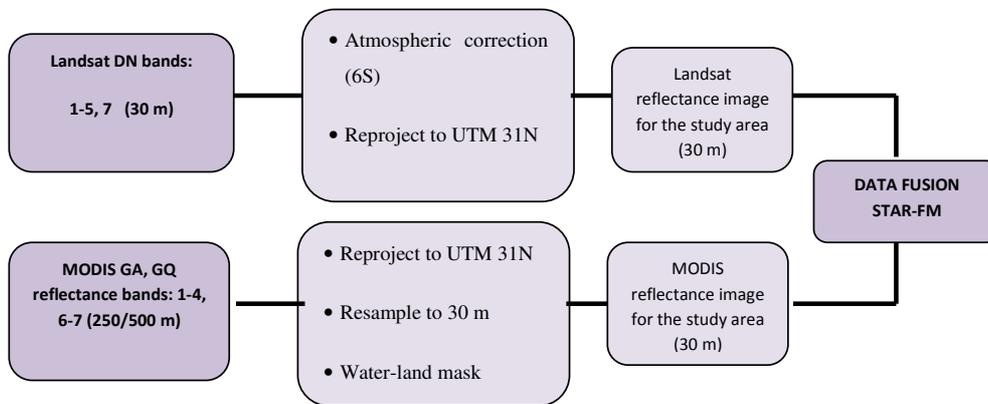


Figura 3.6.- Esquema del procesado de las imágenes Landsat y MODIS.

En la última parte del trabajo realizado para calidad del agua, correspondiente al trabajo presentado en el *Anexo C*, se trabajó con el producto orto-rectificado, y ya corregido atmosféricamente, L2A de las imágenes SPOT-5 take 5. Este producto está procesado por el centro de datos terrestres THEIA y se distribuye a través de la web de la ESA-CNES mediante el portal de spot-take5.org.

En la parte del estudio del estudio de los patrones hidrológicos, correspondiente al artículo presentado en el *Anexo D*, se utilizaron imágenes del sensor ETM+. En concreto se utilizaron los productos de reflectividad de superficie corregidos de la contribución atmosférica mediante el código de transferencia radiativa 6S (CDR_sr). En este caso también se llevó a cabo la normalización de las imágenes (Scroeder *et al.*, 2011, Yang *et al.*, 2012) usando una única escena (imagen del mes de julio 2014) como imagen de referencia. Para la normalización radiométrica se utilizó el método iterativo de normalización de detección multivariada de alteraciones (IRMAD) (Canty y Nielsen, 2008) para reducir al mínimo la variabilidad espectral causada por efectos estacionales

debidos al Sol-superficie-sensor (Yang *et al.*, 2012). La técnica IRMAD fue desarrollada para realizar la normalización radiométrica automática de imágenes multiespectrales e hiperespectrales. Esta técnica permite encontrar combinaciones lineales entre las bandas de la imagen de referencia y la imagen a normalizar. Utilizando un análisis de correlación canónica se obtienen nuevas imágenes multiespectrales. Las componentes de las nuevas imágenes se conocen como variables canónicas. Esta técnica considera que los valores de reflectividad de algunas zonas de una escena adquirida en diferentes periodos de tiempo tendrán zonas invariantes, es decir, que el valor de reflectividad de estos píxeles no variará con el tiempo. Con este supuesto, los píxeles que presentan las menores diferencias en estas variables canónicas se consideran píxeles pseudo-invariantes. Estos píxeles, a posteriori, son los utilizados para normalizar cada una de las bandas, banda por banda, a partir de la imagen de referencia. En nuestro caso se utilizaron las ecuaciones de regresión lineal que mostraban en las seis bandas de la imagen valores del coeficiente de regresión R^2 mayores a 0,90 y errores cuadráticos medios menores del 10%. Además, como en el trabajo presentado en el *Anexo B*, se tuvo que realizar una corrección del bandeo producido por el error del SLC en las imágenes ETM+. Finalmente, para separar el área de estudio (lagunas) del resto de superficies se realizó una máscara de agua digitalizando las masas de agua objeto de estudio. Esta máscara fue realizada con la imagen de un año húmedo (2010) ya que con esto se delimitó la zona de máxima inundación.

3.3. Fusión de datos

La fusión de datos de imágenes de satélite implica la combinación de las características espectrales, temporales o espaciales de dos o más imágenes para formar una nueva imagen con la información de cada una de las imágenes integradas (Genderen y Pohl, 1994). El resultado es una imagen fusionada o sintética que posee las mejores características de las imágenes individuales. Puesto que la fusión de datos se puede basar en una o más propiedades de las imágenes originales (espaciales, temporales, espectrales...) una gran ventaja de la fusión radica en que las imágenes sintéticas

pueden proporcionar más información que las imágenes originales (Pohl y Genderen, 1998).

Existen diferentes técnicas disponibles de fusión de datos (fig. 3.7), y la selección del algoritmo a aplicar depende del tipo de producto que queramos obtener. Estas técnicas de fusión se dividen en 3 grupos, dependiendo del nivel al cual se realiza el procesamiento (Pohl y Genderen, 1998): 1) nivel de píxel, 2) nivel de objeto y 3) nivel de decisión.

A nivel de píxel (nivel bajo), como su nombre indica, la fusión se realiza píxel a píxel. La información del píxel de la imagen fusionada es el resultado de la información asociada a los píxeles de las imágenes de referencia. En cuanto al nivel de objeto (nivel medio), la fusión se realiza extrayendo objetos reconocidos de las imágenes de referencia y a continuación se realiza la fusión. Esta extracción se puede realizar extrayendo información de extensión, textura, vecindad, etc. En el caso del nivel de decisión (nivel alto) como su nombre indica se fusionan decisiones. Las imágenes de referencia se procesan por separado, se extrae la información y en la última etapa es cuando se realiza la fusión; por ejemplo se podrían fusionar decisiones de tipo objeto.

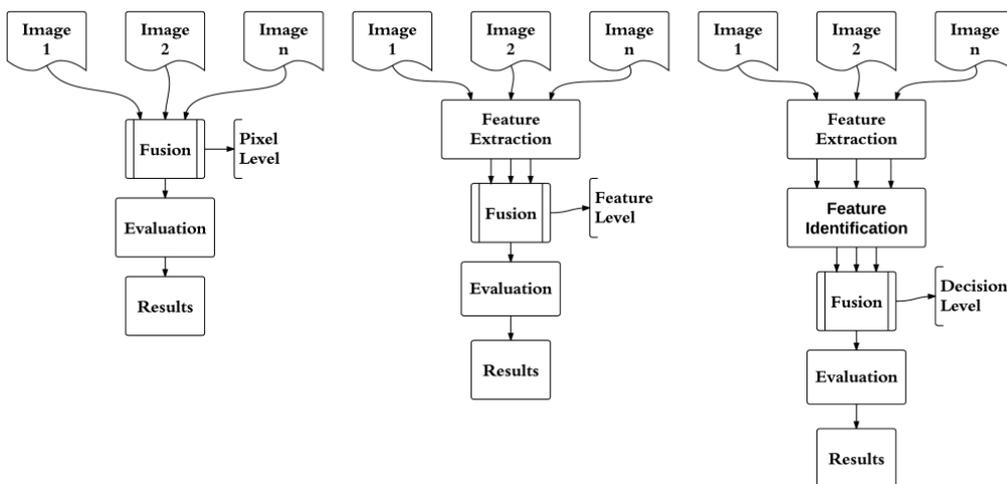


Figura 3.7. Niveles de procesamiento de la fusión de imágenes (Pohl y Genderen, 1998).

En esta Tesis se fusionaron imágenes Landsat 5,7 y MODIS debido a que la resolución espacial de Landsat (30 m) es idónea para poder estudiar la variabilidad espacial de los parámetros de calidad de agua, en comparación con los 250/500m de resolución

espacial de MODIS. El problema de Landsat radica en su escasa resolución temporal de 16 días en comparación con la visita diaria de MODIS. Este largo intervalo de tiempo sin imágenes del sensor se puede solventar aplicando técnicas de fusión de imágenes aprovechando la equivalencia espectral de los diferentes sensores. De esta forma se obtienen imágenes sintéticas de Landsat en los días en que el sensor no pase por nuestra zona de estudio.

Para este estudio se seleccionó el algoritmo de fusión STAR-FM (Temporal Adaptive Reflectance Fusion Model). Se trata de un software de fusión a nivel de pixel producido por la National Aeronautics and Space Administration of the United States (NASA) (Gao *et al.*, 2006). La fusión a nivel de pixel se realiza principalmente por la combinación de los atributos físicos que se caracterizan por el pixel, como la información espacial o espectral. Las técnicas utilizadas se pueden aplicar a cualquier sensor que tengan bandas espectrales similares, como por ejemplo Landsat y MODIS, dado que se requieren similitudes espectrales. Para la fusión entre las imágenes Landsat y MODIS, se requiere al menos un par de imágenes, una Landsat y otra MODIS del mismo día, seguido por otra imagen MODIS correspondiente a la fecha en la que se quiera generar la imagen Landsat sintética. Aunque el algoritmo STAR-FM puede trabajar en base a un solo par de entrada Landsat-MODIS, también se pueden utilizar dos pares de imágenes Landsat-MODIS, anteriores y posteriores a la fecha deseada (Gao *et al.*, 2006).

El algoritmo de fusión STAR-FM funciona de forma que a través de la matriz de pixeles de las imágenes Landsat y MODIS se selecciona un pixel central y se reasigna el valor del nuevo pixel basado en los pixeles candidatos que están más cercanos al pixel central y que son a su vez espectralmente similares. Los pixeles que presentan mayores diferencias espectrales son filtrados. Por lo tanto, la reflectividad del pixel en cuestión no es tan solo el valor del pixel más cercano, sino que es una ponderación de los pixeles candidatos. Esta ponderación se aplica en base a la probabilidad de cada uno de los pixeles candidatos seleccionados. Los factores de ponderación más altos son asignados si el pixel candidato es espectralmente y temporalmente similar al pixel central, además de su distancia geométrica desde el pixel central (Gao *et al.*, 2006). A

partir de todo este proceso se genera la imagen Landsat sintética basada en los píxeles candidatos de entrada de la imagen MODIS cuya fecha corresponde a la que se quiere generar la imagen sintética, es decir, la fecha de predicción.

Cabe señalar que una limitación principal de STAR-FM es la presencia de nubes en las escenas. Si hay nubes presentes en el área de estudio no se puede estimar la reflectividad de superficie. Por otra parte, la aplicación de esta metodología en zonas de agua altamente dinámicas no es viable, como por ejemplo zonas de corrientes rápidas, o en el caso de formación de blooms de algas o zonas con condiciones muy ventosas.

3.4. Métodos tradicionales. Modelos linealizables e índices.

3.4.1. Estimación de parámetros de calidad de aguas (Doña et al., 2014- Anexo A)

Como ya se ha comentado en la introducción de la Tesis, existen diversos métodos para la obtención de parámetros de calidad de agua utilizando técnicas de teledetección. Una de las técnicas más comunes aprovecha la combinación de bandas de un sensor estableciendo regresiones lineales simples o múltiples.

En la primera parte de la tesis se utilizaron este tipo de técnicas para estimar parámetros de calidad como son la concentración de clorofila-*a*, transparencia y partículas en suspensión. En este caso el estudio se realizó utilizando las bandas del sensor Landsat TM que son comunes con Deimos-1. Estas bandas pertenecen al intervalo del espectro correspondiente al verde, rojo e infrarrojo cercano.

Para determinar la concentración de clorofila-*a* se empleó la banda TM2, perteneciente al intervalo del verde, ya que el espectro de una masa de agua con presencia de clorofila-*a* presenta un máximo absoluto de reflectividad en 566 nm. En el caso de la estimación de concentración de partículas en suspensión se empleó la banda TM4, correspondiente al infrarrojo cercano, ya que en el espectro del agua las partículas en suspensión presentan un máximo relativo de reflectividad en 830 nm en este caso, dentro de dicha banda. Finalmente, para la estimación de la transparencia se recurrió a la banda perteneciente al verde, TM2, ya que la máxima penetración de la luz en estas masas de agua tiene lugar en este intervalo espectral (Domínguez, 2002).

En esta parte de la Tesis se emplearon los datos de campo del estudio de Caselles *et al.* (1985), los datos del CEH del Parque Regional del Sureste del 3-4 de julio del 2000 y los datos de las campañas de la Albufera del 20/07/10 y del 05/08/10, así como los respectivos valores de reflectividad pertenecientes a las bandas TM2 y TM4.

Las bases de datos se dividieron en dos grupos, 2/3 de los datos se usa para la construcción de los algoritmos, y el 1/3 restante para realizar la validación de los mismos. Se evaluaron regresiones lineales, potenciales, logarítmicas, exponenciales y polinomios de segundo y tercer orden. Además, también se realizó una intercomparación con los resultados obtenidos empleando otras bandas y algoritmos propuestos por otros autores. Seguidamente el algoritmo que mostró los mejores resultados estadísticos se aplicó a un set de imágenes Deimos-1, realizando una validación de estos algoritmos con datos de campo pertenecientes a fechas de campañas coincidentes con la pasada del sensor. Finalmente se aplicó el mejor algoritmo para cada uno de los parámetros a una serie de imágenes Landsat-5 y Deimos-1, y se obtuvieron mapas de la distribución espacial de cada uno de ellos en la Albufera de Valencia para diferentes fechas.

3.4.2. Estimación de área inundada (Doña *et al.*, 2016-Anexo D)

En lo referente al estudio de la estimación de área inundada las técnicas tradicionales más extendidas se basan en el uso de índices de agua. Estos índices se basan en las características del espectro de agua para poder identificar zonas inundadas utilizando las bandas del infrarrojo cercano y medio, que es dónde el agua muestra valores muy bajos de reflectividad. En esta Tesis se han estudiado los siguientes índices (ecs. 3.8-3.11): a) el índice de agua de diferencia normalizada (NDWI_G) propuesto por Gao (1996) y que está formado por las bandas del infrarrojo cercano y del medio, en nuestro caso, las bandas 4 y 5 del sensor ETM+; b) el propuesto por McFeeters (1996) (NDWI_M) con las bandas del verde y del infrarrojo cercano; y c) el índice de agua de diferencia normalizada modificado (MNDWI) propuesto por Xu (2006) con la utilización de las bandas del verde y del infrarrojo medio, valores positivos de estos índices corresponden a píxeles de agua.; d) el propuesto por Ángel-Martínez (1994) y

utilizado por el CEDEX para cartografiar aguas continentales (I_CEDEx) (Bustamante et al., 2006), este índice emplea las bandas del rojo, infrarrojo cercano e infrarrojo medio, si este índice presenta valores inferiores a 0,4, se considera como área inundada.

$$NDWI_G = \frac{ETM+4 - ETM+5}{ETM+4 + ETM+5} \quad (3.8)$$

$$NDWI_M = \frac{ETM+2 - ETM+4}{ETM+2 + ETM+4} \quad (3.9)$$

$$MNDWI = \frac{ETM+2 - ETM+5}{ETM+2 + ETM+5} \quad (3.10)$$

$$I_{CEDEx} = \frac{ETM+4}{ETM+3} - \frac{ETM+4}{ETM+5} \quad (3.11)$$

Además de los cuatro anteriores, también se abordaron algunos índices de vegetación como e) el índice de vegetación normalizado (NDVI) propuesto por Rouse *et al.* (1974), el cual utiliza las bandas del espectro del rojo y del infrarrojo cercano; y f) el índice de vegetación ajustado al suelo (SAVI) propuesto por Huete *et al.* (1988), que utiliza estas mismas bandas. En este caso, un valor negativo en estos índices implica que el pixel está inundado (ecs. 3.12, 3.13). Además también se estudió el umbral propuesto por Bustamante *et al.* (2005) con la banda del infrarrojo medio del sensor ETM+ y, por otro lado, se estudió la banda del infrarrojo cercano para la discriminación de los pixeles de agua, en este último caso con el histograma de la imagen.

$$NDVI = \frac{ETM+4 - ETM+3}{ETM+3 + ETM+4} \quad (3.12)$$

$$SAVI = \frac{ETM+4 - ETM+3}{ETM+3 + ETM+4 + L} \cdot (1 + L) \quad (3.13)$$

donde L es la constante para ajustar la línea de vegetación-suelo al origen.

Para la intercomparación de los modelos se utilizaron los datos del 23 de julio del 2014, coincidente con una determinación precisa mediante trabajo de campo en 6 de las lagunas (Alcahozo, Camino de Villafranca, El Longar, La Veguilla, Las Yeguas, y Manjavacas). Los datos del resto de campañas se utilizaron para validar el método que proporcionó mejores resultados.

Para evaluar la bondad de los mapas de clasificación obtenidos se calculó la matriz de confusión para cada uno de los métodos aplicados. Como es sabido, una matriz de confusión ordena todos los casos del modelo en categorías, determinando si el valor de predicción coincide con el valor real. Esta matriz es una herramienta estándar de evaluación de modelos estadísticos. Para evaluar la consistencia de los resultados se utilizó esta matriz para calcular el coeficiente kappa (κ). Este índice presenta valores comprendidos entre $[-1 +1]$, siendo aceptables valores mayores de 0,4 (Landis y Koch, 1977). La matriz de confusión también se puede utilizar para calcular el error de comisión y la precisión del usuario. El error de comisión es el porcentaje de píxeles asignados erróneamente a una determinada clase por el clasificador, mientras que la precisión del usuario es la probabilidad de que un píxel asignado a una clase por el clasificador corresponda correctamente a esa clase. El error de omisión es el porcentaje de píxeles que pertenecen a la clase real pero fueron clasificadas incorrectamente, y la exactitud del productor es la probabilidad de que el clasificador haya asignado correctamente una clase dada por los datos de referencia.

Finalmente, se realizó una validación del mejor método con datos independientes al entrenamiento-testeo del método. Estos datos pertenecen a diferentes fechas, lo que además permite evaluar el funcionamiento del método para otras épocas del año.

3.5. Aprendizaje automático. Algoritmos genéticos

Se pueden utilizar diferentes técnicas de aprendizaje automático para identificar patrones y realizar tanto la clasificación como regresión dentro de un conjunto de datos de interés. Los algoritmos de aprendizaje automático se pueden diferenciar entre algoritmos de aprendizaje supervisados, los no supervisados y los semi-supervisados. En el primer tipo de algoritmo, el aprendizaje supervisado tendremos un conjunto de entrenamiento o de aprendizaje, que servirá para el diseño del clasificador y habrá otro conjunto de validación para estudiar cómo funciona el clasificador. En el caso del aprendizaje no supervisado, también denominado clustering, los datos no han sido pre-procesados, es decir, que no tenemos ningún conjunto de datos para entrenar al clasificador. En este tipo de clasificación tenemos muestras que tienen un conjunto de

características y que no se sabe a qué clase pertenecen. La finalidad de este tipo de técnicas es de poder obtener la mejor agrupación con los objetos que tengan características similares y así poder obtener las diferentes clases. Por último, el aprendizaje semi-supervisado utiliza tanto los datos de entrada sin procesar y procesados para generar una función que explique la relación entre los valores de entrada y de salida. Estas técnicas de aprendizaje que también se pueden utilizar en problemas de regresión incluyen a las redes neuronales artificiales (ANN), las máquinas de soporte vectorial (SVM) y la programación genética (GP) (Chang et al., 2014).

La programación genética es un método de aprendizaje automático que está incluido en un tipo de técnicas denominadas “algoritmos evolutivos”, basadas en los principios del Darwinismo. Este tipo de algoritmos resuelven los problemas imitando procesos evolutivos naturales pudiendo decodificar comportamientos basados en datos empíricos mediante un modo de aprendizaje no supervisado, examinando el conjunto de datos mediante asociación, clasificación, y agrupación, es decir, uniendo todas las características del aprendizaje automático (Seifert, 2004). Todo esto es muy útil ya que el usuario no tiene que especificar ningún tipo de solución o tener un conocimiento previo de la relación existente entre las entradas del modelo y el objetivo deseado.

El primer paso de la programación genética es el de inicializar la población (datos) mediante la creación de una serie de programas al azar. A mayor número de datos, mayor será la capacidad de modelar con mayor precisión el problema, sin embargo, esto requiere más tiempo de procesado y de memoria del ordenador (Francone, 1998). Si un programa cumple los criterios de error mínimos establecidos por el usuario, el proceso del programa genético se completa. En cambio, si no sucede esto, se comienza otra iteración para crear una generación mejorada de programas (Francone, 1998). Esta nueva generación se forma aplicando 3 operadores genéticos para reemplazar los programas que no han obtenido buenos resultados. Los principales operadores genéticos son:

- Reproducción: programas que presentan un buen ajuste y se copian sin cambios en la nueva generación.

- Crossover (recombinación): incluyen instrucciones de intercambios cruzados o nodos dentro de programas con mejor resultado y aptos para desarrollar un nuevo programa. Hay tres tipos de cruces que se pueden aplicar cuando se genera un nuevo programa (Engelbercht, 2007):
 - Asexual: se utiliza un solo programa para crear el nuevo.
 - Sexual: se utilizan dos programas para crear el nuevo.
 - Multi-recombinación: se utilizan nodos de más de dos programas para crear el nuevo.
- Mutación: Cambios al azar realizados en los programas que presentan los mejores ajustes.

Una de las principales ventajas de la programación genética es que la longitud del programa está limitada solamente por las capacidades de memoria del equipo utilizado. Esto hace que los programas puedan crecer y evolucionar sin ningún tipo de restricción hasta que se haya cumplido el criterio de parada. Otra de las grandes ventajas es que la solución de estos programas es una caja blanca, por lo que podemos ver las relaciones existentes entre los parámetros de estudio. Los algoritmos genéticos desarrollados en esta Tesis se han realizado con el software Discipulus®, diseñado por Francone (1998). Discipulus® está diseñado para determinar los mejores 30 modelos en base a los resultados obtenidos con los datos reservados para el entrenamiento y la validación.

En el segundo estudio presentado en la presente Tesis (*Anexo B*) se hizo uso de los algoritmos genéticos para la obtención de los algoritmos de estimación de los diferentes parámetros de calidad de agua estudiados. Una vez realizada la fusión de las imágenes, se juntaron las bases de datos de la Albufera de Valencia de las campañas de 2006 y los datos correspondientes al 12/04/2011, así como los respectivos valores de reflectividad pertenecientes a las bandas fusionadas.

En la última parte perteneciente al estudio de calidad de aguas (*Anexo C*), los datos utilizados para estudiar los algoritmos genéticos fueron los correspondientes a las campañas en la Albufera de Valencia del 24 de junio y 9 de julio de 2015 y los datos del 24 de julio de los embalses de Escalona y Tous, así como los respectivos valores de reflectividad pertenecientes a las bandas de Sentinel-2 simuladas.

Respecto a la parte de estimación de área inundada (*Anexo D*) también se utilizó esta técnica. En este caso los datos de reflectividad con los que se desarrollaron los algoritmos genéticos fueron los correspondientes a las lagunas de Alcahozo, Camino de Villafranca, El Longar, La Veguilla, las Yeguas, y Manjavacas del 23 de julio del 2014.

En todos los trabajos en los que se utilizó este tipo de técnicas, para llevar a cabo la obtención de los algoritmos genéticos, las bases de datos se dividieron en dos grupos, 2/3 de los datos para la construcción de los algoritmos, y el restante 1/3 para la realización de la validación de éstos. Después se realizó el estudio estadístico para poder obtener el error de estimación de cada uno de los parámetros de calidad estudiados o la precisión de la clasificación.

En el estudio de estimación de área inundada se estudió el funcionamiento en la estimación de los píxeles de agua con otro tipo de métodos de aprendizaje automático, como son, dentro del grupo de clasificación supervisada: paralelepípedo, mínima distancia, Mahalanobis y máxima verosimilitud, en el grupo de clasificación no supervisada: k-medias y el Iterative Self-Organizing Data Analysis Technique (ISODATA), y en el caso de los semi-supervisados: ANN y las SVM.

A continuación se explica brevemente en que se basa cada uno de estos métodos:

- Clasificación supervisada

- Paralelepípedo: Se determinan los valores máximos y mínimos estimados a partir de la media y la desviación estándar de cada clase en cada banda. Todos aquellos píxeles que se encuentren dentro del área comprendida por un paralelepípedo son asignados a esa clase, en caso contrario no se les asigna ninguna clase.
- Mínima distancia: En este caso la clasificación consiste en la determinación de las medias de cada clase y la asignación se realiza hacia la clase que presenta menor distancia.
- Máxima probabilidad: Este método asume que los datos presentan una función de distribución normal para asignar la probabilidad de que un píxel cualquiera

pertenezca a cada una de las clases. El pixel se asigna de este modo a la clase a la que es más probable que pertenezca.

- Mahalanobis: Es un método muy parecido al de máxima probabilidad, pero asumiendo que las covarianzas de las clases son iguales.

- Clasificación no-supervisada

- K-medias: Este método determina las medias de las diferentes clases y luego de forma iterativa los pixeles son asociados a las clases más cercanas utilizando técnicas de mínima distancia.
- ISODATA: Este método también es de agrupamiento pero las clases se dividen o se fusionan a partir de las existentes si se cumplen ciertas condiciones. Estos parámetros permiten dividir o fusionar clases y el número de clases son definidas por el usuario.

- Clasificación semi-supervisada

- ANN: Este método imita el funcionamiento de las redes neuronales, es un sistema de interconexión de neuronas que colaboran entre sí para producir un estímulo de salida. Un modelo ANN se basa en tres capas diferentes: capa de entrada (es decir, los datos de entrada incluyen bandas de reflectividad y/o índices de vegetación o agua), una o más capas ocultas, y la capa de salida, es decir, la salida binaria incluye agua (1) o no agua (0).
- SVM: Es una técnica de aprendizaje automático supervisado basado en la teoría del aprendizaje estadístico, que se utiliza para encontrar los límites de diferentes clases.

Una vez comparados todos los métodos utilizados, tanto los clásicos, explicados anteriormente, como los de aprendizaje automático, se seleccionó aquel que presentaba mejores resultados en estimación de pixeles cubiertos por agua para aplicarlo a todo el conjunto de imágenes disponible, y también a otras masas de agua de la zona con el fin de evaluar su funcionamiento en la estimación de pixeles cubiertos por agua.

CAPÍTULO 4

Resultados y discusión

En este capítulo se muestran los resultados obtenidos del propio análisis de los datos verdad-terreno, así como los referentes a los algoritmos propuestos y evaluados, tanto para el estudio de calidad como de área inundada. El capítulo se completa con la discusión que se deriva de todo ello, mientras que los artículos que se encuentran en los anexos recogen los trabajos completos.

4.1. Datos de campo

En la tabla 4.1 se muestran los valores de concentración de clorofila-*a* de la Albufera de Valencia obtenidos del análisis de las muestras. Según los valores establecidos en 1982 por la OCDE (Organización para la Cooperación y el Desarrollo Económico), como se ha comentado en capítulos anteriores, nos encontramos ante una masa de agua hipereutrófica.

Tabla 4.1. Resumen de los estadísticos de los datos de concentración de clorofila-*a*, [Chl-*a*], para cada una de las campañas de muestreo realizadas en la Albufera de Valencia. En cada muestreo se incluyeron 15 puntos distribuidos en toda la superficie del lago.

Fecha	[Chl- <i>a</i>] (mg/m ³)			
	Mínimo	Máximo	Promedio	σ
20/07/2010	45	99	61	13
05/08/2010	24	74	53	13
25/01/2011	32	159	88	30
12/04/2011	147	238	196	30
13/07/2011	50	72	65	7
24/06/2015	80	175	125	35
09/07/2015	50	85	70	10

De entre todas las campañas realizadas los valores más bajos de este parámetro se presentaron en agosto de 2010, con un valor promedio de 53 mg/m³, y el mayor valor absoluto registrado fue de 238 mg/m³, correspondiente al mes de abril de 2011, con un valor promedio en este caso de cerca de 200 mg/m³. Se pudo observar también que las mayores variaciones de clorofila-*a*, descritas por la desviación estándar, se presentaron en los meses de enero, abril y junio, correspondiendo con la apertura de las compuertas, es decir, con el periodo en el que se presenta una mayor variación del flujo de agua en la Albufera de Valencia. En los demás meses se comportó de forma más homogénea, correspondiendo al periodo en el que las compuertas están cerradas, cuando se presenta el menor flujo de agua.

En cuanto a los valores de concentración de partículas en suspensión (tabla 4.2), en las campañas de 2010-2011 la variación de este parámetro se comportó de una forma más homogénea que la clorofila-*a*. Sin embargo, las medidas de 2015, nos indican una

mayor variación espacial de este parámetro. Los mayores valores se observaron en julio de 2015, con un valor promedio de este parámetro de 125 mg/l.

Tabla 4.2. Resumen de los estadísticos de los datos de concentración de partículas en suspensión, [TSS], para cada una de las campañas de muestreo realizadas en la Albufera de Valencia. En cada muestreo se incluyeron 15 puntos distribuidos en toda la superficie del lago.

Fecha	[TSS] (mg/l)			
	Mínimo	Máximo	Promedio	σ
20/07/2010	68	87	78	5
05/08/2010	42	75	60	9
25/01/2011	25	54	44	8
12/04/2011	53	81	65	7
13/07/2011	62	89	78	7
24/06/2015	80	160	120	25
09/07/2015	85	160	125	20

Finalmente, en la tabla 4.3 se pueden observar los valores de transparencia del agua medida con el disco de Secchi. Se aprecia que los valores de la Albufera de Valencia corresponden a los de una masa de agua hipereutrófica, según los valores establecidos por la OCDE (1982). En este caso se obtuvieron medidas mínimas de transparencia de 10 cm y máximas de apenas 40 cm. La variación de este parámetro es escasa, registrándose variaciones máximas de 5 cm a lo largo de los diferentes puntos de muestreo.

Los valores de estos parámetros determinados en los embalses de Escalona y Tous (tabla 4.4) difieren bastante de los de la Albufera de Valencia. Siguiendo la clasificación trófica de la OCDE (1982), los valores presentados por el embalse de Escalona corresponderían a una masa de agua oligotrófica-mesotrófica. Las concentraciones de clorofila-*a* y de partículas en suspensión en este embalse fueron menores de 2 mg/m³ y de 4 mg/l, respectivamente, mientras que se obtuvieron valores de profundidad de visión del Disco de Secchi superiores a 2 m. En el embalse de Tous se registraron valores de concentración de clorofila-*a* inferiores a 0,6 mg/m³ y de transparencia superiores a 4 m. En este último caso los valores corresponderían a una masa de agua oligotrófica.

Tabla 4.3. Resumen de los estadísticos de los datos de transparencia medida con el disco de Secchi (SD), para cada una de las campañas de muestreo realizadas en la Albufera de Valencia.

En cada muestreo se incluyeron 15 puntos distribuidos en toda la superficie del lago.

Fecha	SD (m)			
	Mínimo	Máximo	Promedio	σ
20/07/2010	0,10	0,20	0,14	0,03
05/08/2010	0,14	0,30	0,23	0,05
25/01/2011	0,20	0,40	0,26	0,05
12/04/2011	0,17	0,26	0,21	0,02
13/07/2011	0,22	0,29	0,27	0,02
24/06/2015	0,16	0,22	0,180	0,013
09/07/2015	0,26	0,33	0,30	0,02

Tabla 4.4. Resumen de los estadísticos de los datos de concentración de clorofila-*a*, partículas en suspensión y transparencia medida con el disco de Secchi (SD), para la campaña del 24/07/2015 realizada en los embalses de Tous (18 puntos de muestreo) y Escalona (8 puntos de muestreo).

Embalse	[Chl- <i>a</i>] (mg/m ³)			
	Mínimo	Máximo	Promedio	σ
Tous	0,09	0,60	0,30	0,12
Escalona	0,30	2,30	0,80	0,60
Embalse	[TSS] (mg/l)			
	Mínimo	Máximo	Promedio	σ
Tous	0,6	10	2	2
Escalona	1,3	4	3	0,9
Embalse	SD (m)			
	Mínimo	Máximo	Promedio	σ
Tous	3,8	6,7	5,3	0,7
Escalona	2,3	3,7	3,2	0,5

4.2. Algoritmos para la estimación de parámetros de calidad de agua

4.2.1. Modelado y optimización

Como intento inicial (Doña et al., 2014- *Anexo A*) se intentó encontrar un algoritmo que relacionara la concentración de clorofila-*a* con la banda correspondiente al verde del

sensor Thematic Mapper (TM2), como proponen Domínguez *et al.* (2009), pero los resultados obtenidos no fueron satisfactorios ($R^2=0,40$, $p>0,05$). Para la obtención de los algoritmos se utilizaron los datos de campo correspondientes a la Albufera de Valencia y a las masas de agua del Parque Regional del Sureste. Se utilizaron todas estas masas de agua para tener un intervalo más amplio de datos, ya que debido a la escasa variación de algunos de los parámetros en la Albufera, a veces es difícil encontrar un resultado aceptable solamente con datos de esta masa de agua (López-García y Caselles (1987)) y utilizando este tipo de técnicas más clásicas.

Tras una serie de nuevas pruebas, los mejores resultados se obtuvieron con el cociente entre las reflectividades de las bandas del verde e infrarrojo cercano TM2/TM4, con un coeficiente de correlación de 0,66 y un valor- $p < 0,001$. En la figura 4.1 se muestra como los resultados obtenidos a través del cociente TM2/TM4 pueden ser bastante sensibles a valores de concentración de clorofila-*a* inferiores a 1,5 mg/m³, si bien hay que señalar que estos valores no son habituales en la masa de agua de interés, la Albufera de Valencia.

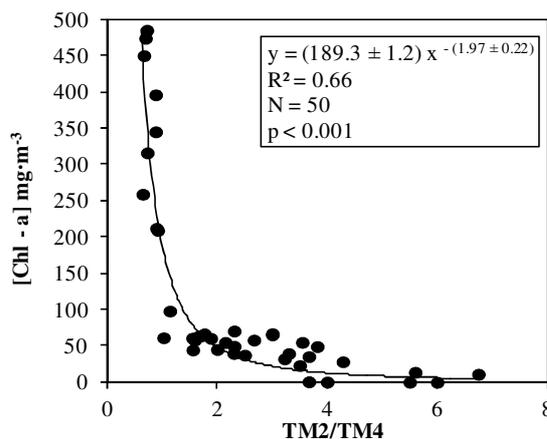


Figura 4.1. Relación entre la concentración de clorofila-*a* [Chl-*a*] y el cociente de las reflectividades de superficie de las bandas 2 y 4 del sensor Thematic Mapper.

Para las partículas en suspensión, se consiguieron resultados aceptables a través de la reflectividad de la banda del infrarrojo cercano (TM4). También en este caso la relación

que mejores resultados proporcionó es de tipo potencial, mejorando a los algoritmos publicados por Domínguez *et al.* (2009) y Ma *et al.* (2005). En la tabla 4.5 se muestra esta comparación entre los resultados obtenidos.

Tabla 4.5. Algoritmos de estimación de concentración de partículas en suspensión con valores de reflectividad de superficie de las bandas del sensor Thematic Mapper (TM).

Autor	Algoritmo	a	b	R ²
Propuesto	$b \cdot (TM4)^a$	$1,07 \pm 0,22$	$15,9 \pm 1,1$	0,92
Domínguez <i>et al.</i> ,(2009)	$a \cdot \ln(TM4) + b$	26 ± 2	30 ± 2	0,84
Ma <i>et al.</i> (2005)	$a \cdot (TM4)^2 + b$	$3,0 \pm 0,3$	18 ± 3	0,73
Ma <i>et al.</i> (2005)	$a \cdot TM4 + b$	$16,9 \pm 1,0$	4 ± 3	0,87

donde a y b son los coeficientes de regresión, R² es el coeficiente de correlación; $p < 0,001$ para todos los algoritmos.

En cuanto a la transparencia, los resultados fueron satisfactorios para la banda que proponíamos (TM2). En la tabla 6 del *Anexo A* se establece una comparación con otros algoritmos publicados que relacionan las bandas del sensor TM con este parámetro. De todos ellos, tan solo se muestran los que proporcionan un coeficiente de correlación mayor de 0,70 y valores-p $< 0,001$.

Para mejorar la modelización se recurrió a técnicas de aprendizaje, en concreto al uso de la programación genética. Esta técnica se utilizó para mejorar los algoritmos de estimación de parámetros de calidad de agua para la Albufera de Valencia. Además se utilizaron las reflectividades de imágenes sintéticas Landsat-5 y 7, obtenidas a partir de la combinación de estos sensores con MODIS (*Anexo B*). En este caso el estudio se centró en la estimación de concentración de clorofila-a y de transparencia. No se incluyó la concentración de partículas en suspensión debido a la falta de datos de campo de este parámetro.

En esta Tesis se utilizó el software Discipulus®, que permite generar las ecuaciones de regresión no lineales que relacionen los diferentes parámetros de interés con las reflectividades. Los resultados obtenidos fueron aceptables para la estimación ambos parámetros, con un coeficiente de correlación de 0,94 en el caso de la clorofila-a

(fig.4.2), y de 0,89 para la transparencia (fig. 4.3), ambos con valores de $p < 0,001$. Los algoritmos obtenidos para cada uno de los casos se muestran en detalle en el *Anexo B*.

También se recurrió a la programación genética en el trabajo presentado en el *Anexo C* para la modelización de los parámetros de calidad, en este caso a partir de las reflectividades simuladas de Sentinel-2. La aplicación de los algoritmos obtenidos a la Albufera de Valencia y los embalses de Tous y Escalona mostraron valores del coeficiente de correlación muy próximos a uno, con valores- $p < 0,001$, para la clorofila-*a*, partículas en suspensión, y transparencia.

4.2.2. Aplicación y validación

En la tabla 4.6 se muestran los resultados de la validación de los algoritmos anteriores con los datos de campo reservados para tal efecto en el trabajo presentado en el *Anexo A*.

Para la concentración de clorofila-*a* se obtuvo un coeficiente de correlación de 0,96 y un error relativo del $\pm 24\%$. En cuanto al algoritmo propuesto para la estimación de concentración de partículas en suspensión, se observó un error de estimación del $\pm 17\%$, con un coeficiente de correlación de 0,91 (tabla 4.6).

Finalmente, en el caso de la transparencia los algoritmos estudiados mostraron resultados aceptables, con valores de R^2 por encima de 0,75 con valores de $p < 0,001$, excepto el caso de Allee et al. (1999). De entre todas las combinaciones de bandas contrastadas, la combinación de TM1 y TM3 fue la que ofreció los mejores resultados, con un error del $\pm 30\%$, viéndose superado por el ajuste potencial con la reflectividad de la banda TM2. Además, este algoritmo basado únicamente en la banda TM2 tiene el atractivo adicional de que permite su aplicación a imágenes del satélite Deimos-1.

La validación de los algoritmos seleccionados aplicados imágenes Deimos-1 mostraron resultados aceptables en estimación de concentración de clorofila-*a* y de partículas en suspensión, con errores del ± 22 y $\pm 15\%$, respectivamente (tabla 4.7).

Tabla 4.6. Resultados estadísticos resultantes de la regresión lineal entre los valores obtenidos con los diferentes algoritmos y los datos de campo para cada uno de los parámetros objeto de estudio, concentración de clorofila-a [Chl-a], concentración de partículas en suspensión [TSS] y transparencia medida con el disco de Secchi (SD).

Autor	Parámetro	Bandas TM	a	b	R ²	Sesgo	MAD	RMSE	MADP
Algoritmo propuesto	[Chl-a] (mg/m ³)	TM2/TM4	0.88 ± 0.03	7 ± 5	0.96	-4	22	30	24 %
Algoritmo propuesto	[TSS] (mg/l)	TM4	0.97 ± 0.06	-2 ± 3	0.91	-3	7	9	17 %
Domínguez et al. (2009)			0.85 ± 0.05	6 ± 3	0.85	-0.00	10	12	25 %
Ma et al. (2005)			0.76 ± 0.06	10 ± 3	0.76	0.3	12	15	30 %
Ma et al. (2005)			0.90 ± 0.05	4 ± 2	0.89	0.2	7	10	17 %
Algoritmo propuesto	SD (m)	TM2	0.87 ± 0.03	0.03 ± 0.04	0.92	-0.07	0.21	0.3	30 %
Wu et al. (2008)		TM1, TM3	0.58 ± 0.03	0.13 ± 0.02	0.89	-0.10	0.22	0.4	30 %
Cózar et al. (2005)		TM1, TM3, TM4	0.67 ± 0.07	0.16 ± 0.06	0.76	-0.04	0.3	0.60	50 %
Olmanson et al. (2008)		TM1/TM3, TM1	0.71 ± 0.03	0.09 ± 0.02	0.80	0.50	0.21	0.3	30 %
Allan et al. (2011)		TM3	0.84 ± 0.04	0.02 ± 0.03	0.79	0.08	0.21	0.3	30 %
Allee et al. (1999)		TM3	0.85 ± 0.04	0.18 ± 0.03	0.90	0.15	0.3	0.4	40 %

(*) p-valores < 0.001; a y b son la pendiente y ordenada en el origen, respectivamente, de la regresión lineal entre los valores predichos (P_i) y los observados (O_i); R² es la bondad del ajuste; el sesgo, la media de la diferencia absoluta (MAD); el error cuadrático medio (RMSE); la media de la diferencia absoluta en porcentaje (MAPD):

$$Sesgo = \sum_{i=1}^n (P_i - O_i) / n; MAD = \sum_{i=1}^n |P_i - O_i| / n; RMSE = [\sum_{i=1}^n (P_i - O_i)^2 / n]^{1/2}; MADP = 100 \cdot MAD \cdot \langle O \rangle.$$

Sin embargo, no fueron buenos los resultados obtenidos para la estimación de transparencia, mostrándose errores del ± 40% y un valor p = 0,4, consecuencia directa de los bajos valores que presenta este parámetro, normalmente por debajo de 0,50 m, y de la incertidumbre en su medida. Esto hace que la estimación de transparencia mediante teledetección en la Albufera de Valencia resulte realmente complicada, como han mostrado algunos estudios como el de López-García y Caselles (1987).

Tabla 4.7. Resultados estadísticos resultantes de la regresión lineal entre los valores obtenidos con imágenes Deimos-1 y los datos de campo de la Albufera de Valencia para cada uno de los parámetros objeto de estudio: concentración de clorofila-*a* [Chl-*a*], concentración de partículas en suspensión [TSS] y transparencia medida con el disco de Secchi (SD).

Variable	< O >	< P >	σ	Sesgo	MAD	RMSE	MADP
[Chl- <i>a</i>] (mg/m ³)	120	140	50	30	30	40	22 %
[TSS] (mg/l)	63	64	11	1.3	8	10	15 %
SD (m)	0,25	0,30	0,09	0,06	0,09	0,10	40 %

(*) p-valores < 0.001 en el caso de [Chl-*a*] y [TSS] y p-valor=0,4 en el caso de SD

Una vez introducidos los algoritmos genéticos (*Anexo B*), los resultados de la comparación con las medidas de campo mejoraron. En cuanto a la estimación de concentración de clorofila-*a* (fig. 4.2) se obtuvieron errores cuadráticos medios de 8 mg/m³ y del \pm 6%, en términos relativos, con un valor de R² de 0,98 con un valor-p < 0,001.

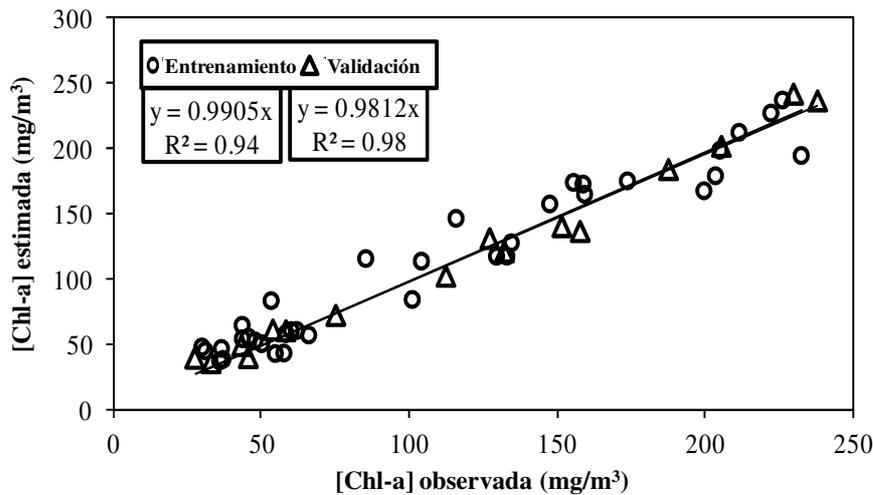


Figura 4.2.- Valores de concentración de clorofila-*a* estimados mediante programación genética vs observados en la Albufera de Valencia (valor-p < 0,001).

Para la transparencia del agua también se obtuvieron resultados aceptables en la validación, con un R^2 de 0,82 y un valor- $p < 0,001$, al comparar los valores estimados y los medidos *in situ* (fig. 4.3). El valor promedio de la transparencia del conjunto de datos reservado para la validación fue de tan sólo 34 cm, con una desviación estándar de 10 cm (tabla 4.8), con lo que se puede apreciar la escasa variación de este parámetro.

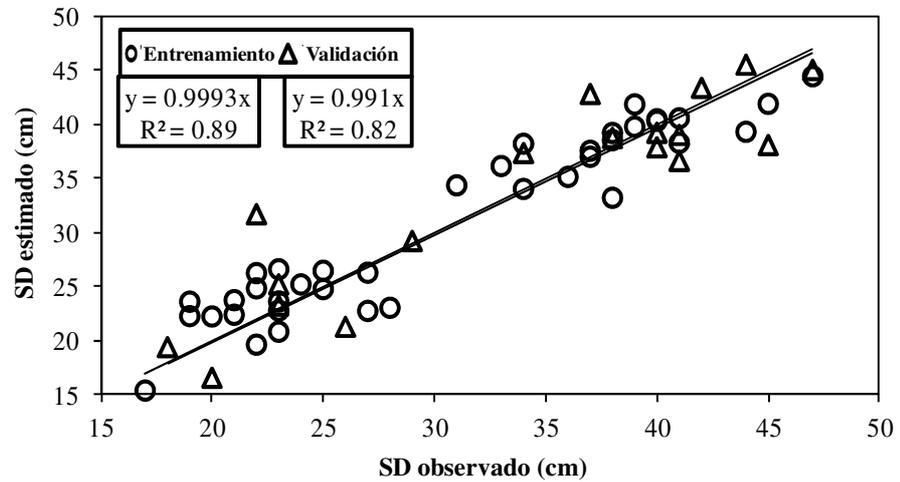


Figura 4.3.- Valores de transparencia (SD) estimados mediante programación genética vs observados en la Albufera de Valencia (valor- $p < 0,001$).

Tabla 4.8. Resultados estadísticos resultantes de la regresión lineal entre los valores obtenidos con imágenes sintéticas Landsat y los datos de campo de la Albufera de Valencia para cada uno de los parámetros objeto de estudio: concentración de clorofila-*a* [Chl-*a*] y transparencia medida con el disco de Secchi (SD).

Variable	<O>	<P>	σ	Sesgo	MAD	RMSE	MADP
[Chl- <i>a</i>] (mg/m ³)	117	116	70	-2	7	8	6 %
SD (cm)	34	34	9	-0,003	3	4	9 %

(*) p-valores < 0.001

Aun así, en este caso, a diferencia de los métodos clásicos, se obtuvo un error en la estimación de la transparencia del $\pm 9\%$ y un valor del error cuadrático medio de sólo 4

cm. Por tanto, la programación genética ayuda a resolver el problema de la dificultad de obtener un algoritmo de transparencia mediante técnicas de regresión (Domínguez *et al.*, 2012), en zonas de escasa variación de este parámetro, sin necesidad de ampliar el intervalo de valores de las bases de datos.

Una vez añadidos a los datos de la Albufera también los dos embalses de estudio, intentando desarrollar un algoritmo para todo este conjunto de masas de agua, los resultados de estimación de concentración de clorofila-*a* mostraron valores de R^2 de 0,98 con un valor- $p < 0,001$ para la base de datos completa. En la tabla 4.9 se recogen los principales estadísticos de la validación de los algoritmos. El error cuadrático medio fue de 6 mg/m³, un $\pm 8\%$ en términos relativos, en estimación de clorofila-*a*. Resultados muy similares se obtuvieron para la concentración de sólidos en suspensión y la transparencia, con errores de estimación inferiores al $\pm 10\%$, y coeficientes de correlación cercanos a 1.

Tabla 4.9. Resultados estadísticos resultantes de la regresión lineal entre los valores obtenidos con imágenes simuladas Sentinel-2 y los datos de campo de la Albufera de Valencia, embalse de Tous y de Escalona, para cada uno de los parámetros objeto de estudio: concentración de clorofila-*a* [Chl-*a*], concentración de partículas en suspensión [TSS] y transparencia medida con el disco de Secchi (SD).

Variable	< O >	< P >	σ	Bias	MAD	RMSE	MADP
[Chl- <i>a</i>] (mg/m ³)	41	40	56	-1,05	3	6	8 %
[TSS] (mg/l)	48	46	64	-1,5	5	8	10 %
SD (m)	2,4	2,5	2,4	0,08	0,12	0,17	5%

(*) p-valores < 0.001

Si los analizamos por separado, los resultados de estimación de la concentración de clorofila-*a* en la Albufera de Valencia fueron muy similares a los obtenidos en el trabajo desarrollado en el *Anexo B*, en este caso con un coeficiente de correlación de 0,9 (valor- $p = 0,001$), el RMSE de 9 mg/m³ y un error de $\pm 7\%$. Sin embargo, los valores obtenidos para los embalses no fueron aceptables con un error superior al $\pm 60\%$,

consecuencia de los bajos valores de concentración de clorofila-*a*, unido al escaso número de datos en estas masas de agua. En cuanto a la transparencia, para la Albufera de Valencia se obtuvo un R^2 de 0,9 con un valor-p igual a 0,01 y un error cuadrático medio de 4 cm, un $\pm 15\%$ en términos relativos, frente a los 20 cm de los embalses, si bien en este caso el error relativo fue de sólo el $\pm 4\%$, debido a los altos valores de transparencia en estas masas de agua.

Por último, en cuanto a los resultados obtenidos en concentración de partículas en suspensión, los resultados de la Albufera de Valencia mostraron un R^2 de 0,8 con un valor-p igual a 0,05 y un error cuadrático medio de 12 mg/l, del $\pm 8\%$ en términos relativos, en sintonía con los obtenidos en el trabajo mostrado en el *Anexo A*. En este caso, el conjunto de datos de los embalses fueron demasiado escasos como para proporcionar estadísticas aceptables, obteniéndose un coeficiente de correlación entre los valores observados y estimados inferior a 0,2 y un valor-p igual a 0,4.

4.3. Algoritmos para la estimación de área inundada

4.3.1. Modelado y optimización

Como se ha explicado en el capítulo de metodología, primero se realizó una comparación entre diferentes métodos para estimar los píxeles correspondientes a cobertura por agua (pixel agua) y secos (pixel no-agua) en nuestra zona de estudio. De las tablas 4.10 a la 4.12 se muestran los resultados que se obtuvieron al aplicar los diferentes métodos en la imagen seleccionada de julio de 2014. Por lo que respecta a los índices de agua, los mejores resultados se obtuvieron a través del NDWI propuesto por McFeeters (1996), que utiliza las bandas del verde y del infrarrojo cercano (tabla 4.10), estableciendo un umbral de cero. Con este índice se obtuvo una precisión global de la clasificación del 90% y un valor del coeficiente kappa de 0,80. En el caso del MNDWI y el NDWI propuesto por Gao (1996), los mejores resultados se obtuvieron con un umbral de 0,5 y para el caso del índice I_CEDDEX con un umbral de -2 (tabla 4.10). En el caso de los índices de vegetación se obtuvieron valores de kappa de 0,4. El análisis visual de los resultados obtenidos en estos últimos permitieron apreciar una sobreestimación de los píxeles de agua, incluso con algunas de las lagunas

completamente secas en julio de 2014, como por ejemplo las lagunas de Manjavacas, Yeguas o Alcahozo.

También se obtuvieron resultados aceptables al aplicar umbrales en las bandas del infrarrojo cercano y medio del ETM+, obteniendo valores del índice kappa de 0,97 y 0,70, respectivamente, para estas bandas (tabla 4.10).

En nuestro caso, aunque los valores de kappa fueron aceptables para la mayoría de los índices y las bandas estudiadas, los errores de omisión y comisión resultaron elevados, y el error total (porcentaje de omisión y comisión) fue igual o superior al $\pm 30\%$ en todos los casos, a excepción de la banda del infrarrojo cercano (ETM+4), en cuyo caso el error total obtenido fue del $\pm 4\%$ (tabla 4.10).

Tabla 4.10. Resumen de los resultados del estudio de la precisión de los diferentes índices de agua y de vegetación y de las bandas simples aplicadas para estudiar la obtención de píxeles de agua.

Método		Precisión global (%)	Kappa	Error de comisión (%)	Error de omisión (%)	Precisión productor (%)	Precisión usuario (%)	Error total (%)
<i>MNDWI</i>	Xu (2006)	90	0,80	30	3	95	70	30
<i>NDWI</i>	Gao (1996)	90	0,60	40	15	85	60	60
<i>NDWI</i>	McFeeters (1996)	90	0,80	25	3	95	75	30
<i>I_CEDIX</i>	Ángel-Martínez (1994)	90	0,70	40	10	90	60	50
<i>NDVI</i>	Rouse et al. (1974)	70	0,40	60	30	70	40	90
<i>SAVI</i>	Huete (1988)	70	0,40	60	30	70	40	90
<i>ETM+4</i>		99	0,97	2	2	98	98	4
<i>ETM+5</i>		90	0,70	40	0	100	60	40

En contraste, los métodos de clasificación supervisada basados en la distancia de Mahalanobis y máxima verosimilitud ofrecieron mejores resultados que los índices de agua (tabla 4.11). En particular, la combinación de las bandas del rojo e infrarrojo cercano fueron la que mejores resultados proporcionaron, con valores kappa de 0,97 y 0,80 para Mahalanobis y máxima verosimilitud, respectivamente. En general, los métodos de clasificación mostraron buenos resultados respecto a los valores obtenidos del parámetro kappa (tabla 4.11).

Tabla 4.11. Resumen de los resultados de la precisión de los métodos de clasificación supervisada y no supervisada aplicados a la estimación de píxeles de agua.

Método		Precisión global (%)	Kappa	Error comisión (%)	Error omisión (%)	Precisión productor (%)	Precisión usuario (%)	Error total (%)
<i>Paralelepípedo</i>	Toda Combinación	<40	<0,20	>50	>50	<50	<50	>60
<i>Mínima distancia</i>	ETM+ 4,5	90	0,67	40	0	100	60	60
<i>Mahalanobis</i>	ETM+ 3,4	99	0,97	5	0	100	95	5
<i>Máxima verosimilitud</i>	ETM+ 3,4	95	0,80	0	30	70	100	30
<i>Máxima verosimilitud</i>	ETM+ 4,5	95	0,70	0	40	60	100	40
<i>K-medias</i>	ETM+ 4,5	80	0,60	10	30	90	75	40
<i>K-medias</i>	ETM+ 4	80	0,60	10	30	90	75	40
<i>ISODATA</i>	ETM+ 3,4	95	0,90	10	5	100	80	15

Sin embargo, y al contrario de lo que ocurría con los índices de agua, una de las lagunas que tenía agua en la fecha de la aplicación de los métodos, la laguna Larga de Villacañas, fue clasificada como seca. También se obtuvieron buenos resultados con el método de clasificación no supervisada ISODATA, con un valor de precisión total de la clasificación del 95% y un valor de kappa de 0,90. El error total obtenido en este caso fue del $\pm 15\%$.

Los estadísticos obtenidos con las redes neuronales artificiales, las máquinas de soporte vectorial y los algoritmos genéticos se muestran en la tabla 4.12. En este caso también se obtuvieron buenos resultados, con valores de la precisión global de las clasificaciones superiores al 90% y valores de kappa por encima de 0,70. Los mejores resultados se obtuvieron para el modelo de red neuronal con 4 capas ocultas, con una precisión global de la clasificación del 95% y un valor de kappa de 0,9. Al aplicar las máquinas de soporte vectorial, por lo general, se obtuvieron valores similares, aunque los resultados mejoraron con la SVM con núcleo cúbico, en este caso con una precisión global del 99% y un valor kappa muy cercano a uno.

Tabla 4.12. Resumen de los resultados de la precisión de los métodos de redes neuronales artificiales (ANN), máquinas de soporte vectorial (SVM) y programación genética (GP) aplicadas a la obtención de píxeles de agua.

Método		Precisión global (%)	Kappa	Error comisión (%)	Error omisión (%)	Precisión productor (%)	Precisión usuario (%)	Error total (%)
ANN	ETM+ 1-5, 7	95	0.90	4	15	85	95	19
SVM lineal	ETM+ 1-5, 7	95	0.80	0	15	100	85	15
SVM lineal	ETM+ 3-5	95	0.87	0	18	100	80	18
SVM Cuadrática	ETM+ 3-5	90	0.70	0	40	100	65	40
SVM Cúbica	ETM+ 1-5, 7	99	0.97	0	6	100	94	6
SVM F Gauss ¹	ETM+ 3-5	90	0.70	0	40	100	59	40
SVM M Gauss ²	ETM+ 3-5	95	0.90	0	20	100	80	20
SVM Course	ETM+ 1-5, 7	95	0.80	0	30	100	75	30
SVM Course	ETM+ 3-5	95	0.85	0	20	100	79	20
GP	ETM+ 4	99	0.98	2	0	100	97	2

A diferencia de las redes neuronales artificiales y las máquinas de soporte vectorial, la programación genética produce una relación entre las entradas y las salidas a través de una función no lineal. En este caso, los mejores resultados se obtuvieron a partir de la reflectividad de la banda del infrarrojo cercano (ETM+4), a través de la expresión:

$$GPR_w = \cos ((0.003 \times ETM + 4)^{1/4}) \quad (4.1)$$

donde GPR_w es la salida del GP para poder clasificar los píxeles de agua. En esta Tesis, los valores de $GPR_w > 0,1$ pertenecían a píxeles de agua, mientras que a los valores $< 0,1$ se asignó la clase no-agua. A través de este algoritmo se obtuvo un valor kappa de 0,98, y un error del $\pm 2\%$ en la clasificación de los píxeles de agua (tabla 4.12).

4.3.2. Aplicación y validación

En base a los resultados obtenidos al comparar los diferentes métodos se optó por el algoritmo genético como método seleccionado para estimar los píxeles de agua y no-agua en un conjunto de imágenes de diferentes fechas. Cuando se aplicó el algoritmo al mismo conjunto de lagunas con el que se elaboró, se obtuvo una precisión global del

clasificador del 95% con un valor de kappa de 0,8. Cuando se trabajó con una base de datos diferente, es decir, cuando se aplicó el algoritmo a un conjunto de lagunas distintas a las empleadas en el desarrollo del algoritmo, la precisión fue del 90%, con un valor kappa de 0,8. Por otra parte, se observó que el algoritmo presentaba dos valores de corte dependiendo de la tipología de la laguna: intervalo [-1 0,1] perteneciente a la clase no-agua y [0,1 1] a la clase agua para lagunas sin presencia de playas y [-1 -0,05] perteneciente a la clase no-agua y [-0,05 1] para el agua, para lagunas con presencia de playas. En el trabajo presentado en el *Anexo D* se pueden consultar todos los detalles relativos a la selección de los umbrales en función de la tipología de las lagunas.

Tabla 4.13. Precisión en la estimación de superficie de agua para las diferentes lagunas que poseen datos de referencia.

Laguna ^{*1}	Umbral	Precisión global (%)	Kappa
<i>Alcahozo (13)</i>	-0,05	70	0,45
<i>Camino de Villafranca (10)</i>	-0,05	94	0,62
<i>Grande de Quero (5)</i>	-0,05	99	0,85
* <i>La Veguilla (11)</i>	0,10	86	0,68
<i>Larga de Villacañas (2)</i>	0,10	90	0,61
* <i>Larga de Villacañas (2)</i>	0,10	80	0,80
<i>Las Yeguas (9)</i>	-0,05	97	0,76
<i>El Longar (1)</i>	-0,05	96	0,53
<i>Manjavacas (12)</i>	0,10	96	0,92
<i>Manjavacas (12)</i>	0,10	86	0,73
* <i>Peñahueca (4)</i>	-0,05	95	0,82
<i>Salicor (7)</i>	-0,05	94	0,15

* Datos de referencia y de satélite simultáneos (mismo día); ^{*1} número de identificación de cada laguna numerada como en la figura 2.3.

Finalmente, el método seleccionado se aplicó a imágenes ETM+ simultaneas o cercanas en tiempo con imágenes de alta resolución de Google Earth™ y a las campañas de campo en las que se midió *in situ* el perímetro de las lagunas de Manjavacas y Alcahozo mediante GPS. Como se muestra en la tabla 4.13, los mejores resultados se observaron en las lagunas Grande de Quero, Larga de Villacañas, Las Yeguas, Manjavacas y Peñahueca, con valores kappa superiores a 0,60 y una precisión global de la clasificación superior al 80%. En el caso de las lagunas Grande de Villafranca y

Mermejuela no se pudo realizar la validación al carecer de medidas de GPS y al ser las imágenes Google Earth™ demasiado distantes en el tiempo.

En la figura 4.4 se muestra la aplicación del algoritmo en la laguna de Manjavacas para la imagen ETM+ del 21 de mayo del 2015, y un detalle del perímetro de la laguna medido con GPS para la misma fecha. En esta figura se puede observar, a modo de ejemplo, el buen funcionamiento del umbral de 0,1 para lagunas sin presencia de playas. En este caso se obtuvo una precisión en la estimación de píxeles de agua del 95% y un valor de kappa de 0,9.

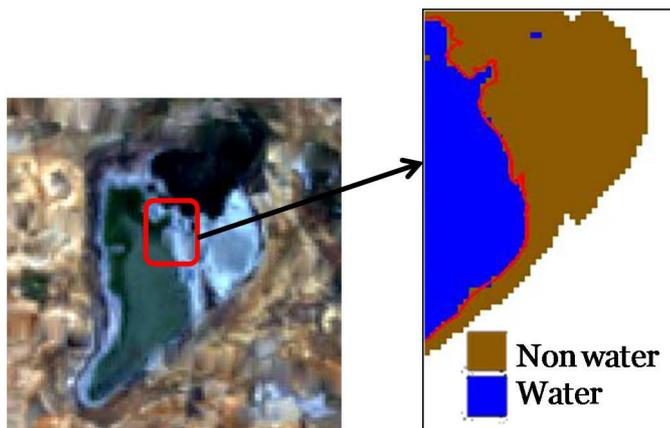


Figura 4.4.- Resultados de la clasificación al utilizar el algoritmo GPRw con un umbral de 0,1 en la laguna de Manjavacas el 21 de mayo de 2015. La línea roja representa la delimitación del perímetro de la masa de agua medido in situ con GPS.

4.4. Elaboración de mapas temáticos. Seguimiento operativo

4.4.1. Seguimiento de la eutrofización de La Albufera de Valencia

Una de las aplicaciones operativas más inmediatas de esta Tesis es el seguimiento espacio-temporal de los parámetros de calidad de aguas. A modo de ejemplo se muestran a continuación los mapas de concentración de clorofila-*a*, partículas en suspensión y transparencia resultado de los diferentes puntos tratados anteriormente.

En la primera parte del estudio se obtuvieron mapas de concentración de clorofila-*a*, partículas en suspensión y transparencia tras la aplicación de los algoritmos clásicos

mencionados anteriormente a las escenas Landsat-5 y Deimos-1 disponibles para el periodo del 2010-2011 en la Albufera de Valencia. Estas imágenes cubren un ciclo anual típico de esta masa de agua. La estacionalidad observada en estos mapas es en gran medida consecuencia del cierre y apertura de las compuertas para cubrir las necesidades para el cultivo de arroz (figs. 4.5- 4.7).

Los altos niveles de concentración de partículas en suspensión (fig. 4.6) están claramente vinculados a valores bajos de transparencia (fig. 4.7), y ambos son una consecuencia de los altos valores de concentración de clorofila-*a* característicos de esta masa de agua (fig. 4.5). Además, en estas figuras se puede observar la gran variación espacial de estos parámetros, es decir, la heterogeneidad espacial de esta masa de agua.

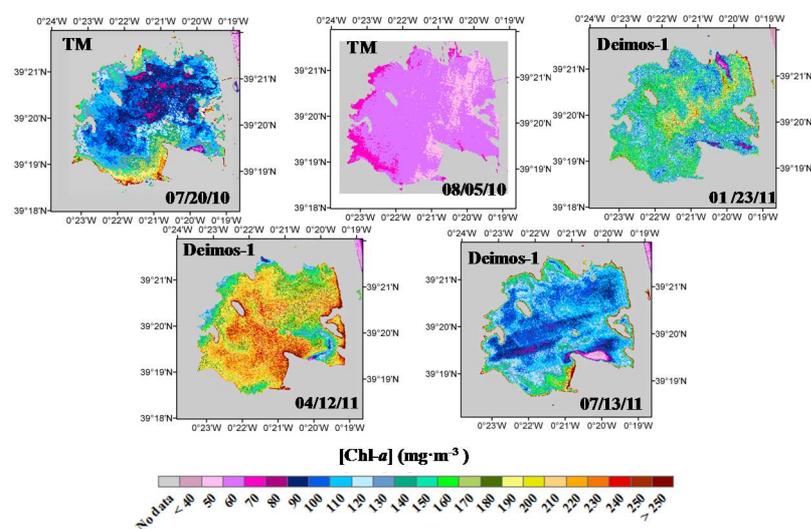


Figura 4.5.- Evolución temporal de la concentración de clorofila-*a* para el periodo 2010-2011.

Los mapas se han elaborado a partir de imágenes Landsat-5 y Deimos-1.

En este caso, la resolución espacial de Landsat y Deimos-1 permite además la detección de canales de entradas y/o salidas de nutrientes en la Albufera de Valencia. Las imágenes Deimos-1 permiten obtener este tipo de mapas cada 3 días gracias a su alta resolución temporal. Por lo tanto, estos dos sensores sirven como herramienta para controlar la calidad del agua y para el análisis de la heterogeneidad espacial de la misma. Además el extenso archivo histórico de imágenes de la serie Landsat posibilita

la realización de estudios temporales, prolongados en el tiempo, lo cual es altamente atractivo para mejorar nuestra comprensión del funcionamiento de las masas de agua.

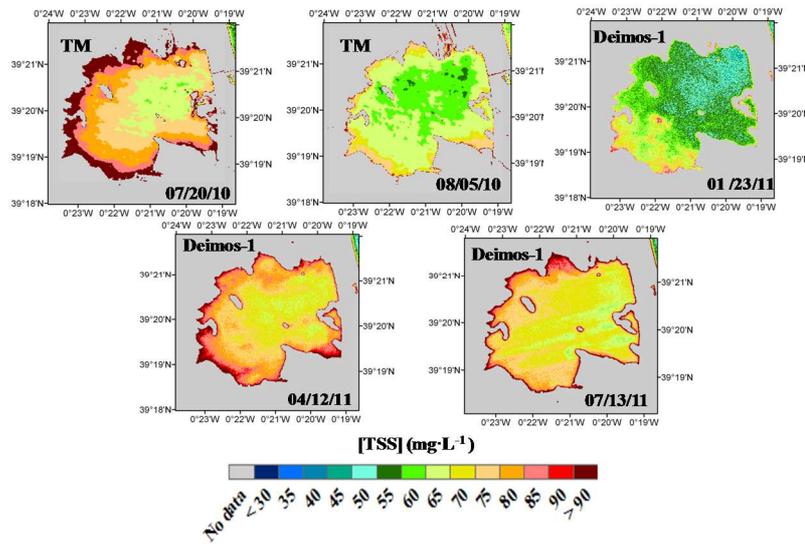


Figura 4.6.- Evolución temporal de la concentración de partículas en suspensión para el periodo 2010-2011. Los mapas se han elaborado a partir de imágenes Landsat-5 y Deimos-1.

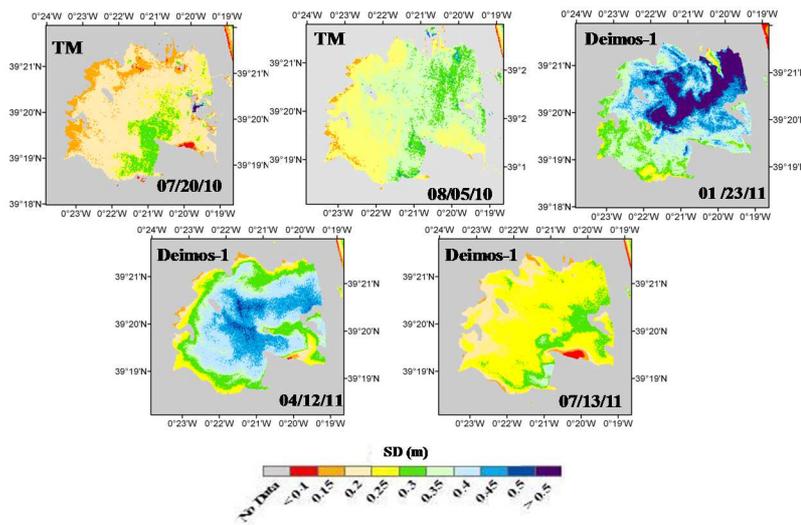


Figura 4.7.- Evolución temporal de la transparencia del disco de Secchi para el periodo 2010-2011. Los mapas se han obtenido a partir de imágenes Landsat-5 y Deimos-1.

A continuación se muestran los mapas obtenidos realizando la fusión de imágenes del Moderate Resolution Imaging Spectroradiometer (MODIS), TM y Enhance Thematic Mapper (ETM +). Como se ha comentado en capítulos anteriores, mientras Landsat ofrece una media resolución espacial, la baja resolución temporal de 16 días es un inconveniente importante para poder realizar el control en tiempo casi real. Esta diferencia se puede solventar mediante el uso de imágenes MODIS que tienen una alta resolución temporal de 1 día, a pesar de su baja resolución espacial (250/500 m). Con todo esto se han generado imágenes sintéticas de Landsat para fechas en las que Landsat no pasa sobre el área de estudio. En la figura 4.8 se muestra un ejemplo de mapas de concentración de clorofila-*a* y de transparencia resultado de la fusión de imágenes Landsat-5-7 y MODIS (libres de nubes) entre los meses de agosto y septiembre del 2006.

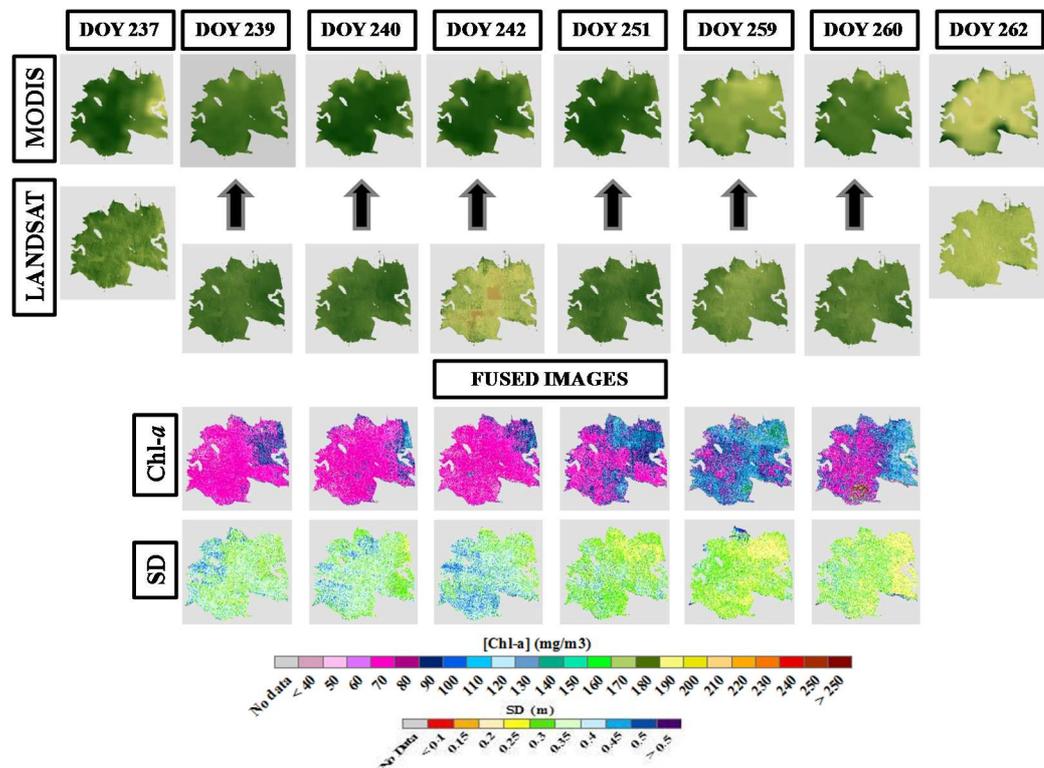


Figura 4.8.- Mapas de concentración de clorofila-*a* y transparencia elaborados a partir de imágenes Landsat sintéticas (fusión Landsat-MODIS) para el periodo agosto-septiembre del 2006 en la Albufera de Valencia. (DOY: día del año)

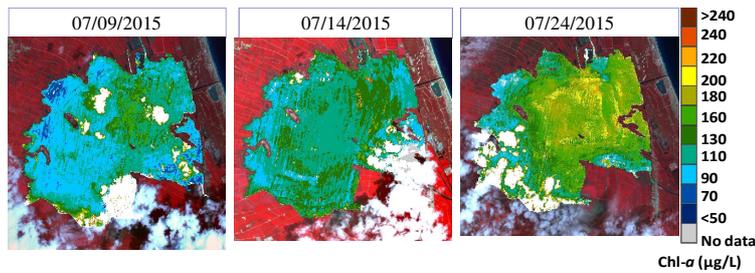


Figura 4.9.- Mapas de concentración de clorofila-*a* elaborados a partir de las imágenes de SPOT take-5 en la Albufera de Valencia.

En las figuras 4.9 - 4.10 se muestran mapas elaborados a partir de imágenes simuladas de Sentinel-2. En ellos se aprecia claramente los altos valores de concentración de clorofila-*a* presentes en esta masa de agua, y los bajos valores de transparencia, así como la propia variación espacial de estos parámetros.

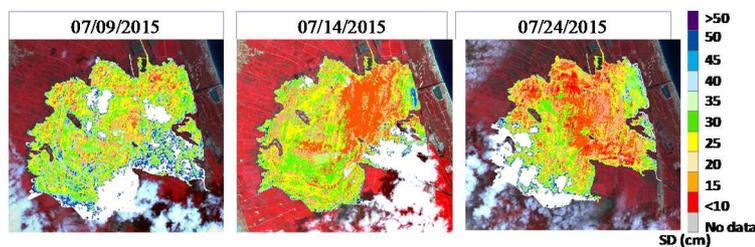


Figura 4.10.- Mapas de transparencia elaborados a partir de las imágenes del SPOT take-5 en la Albufera de Valencia.

4.4.2. Patrones hidrológicos en base al área inundada

En esta Tesis, además de haberse desarrollado trabajos conducentes a la estimación de la calidad del agua, se han desarrollado metodologías para el seguimiento hidrológico de lagunas naturales. Concretamente, el algoritmo seleccionado para la estimación de píxeles de agua se aplicó a las imágenes disponibles para estimar el área inundada de lagunas manchegas desde la primavera de 2013 hasta la de 2015.

Para relacionar el patrón de inundación reconstruido con las variables determinantes de la inundación nos centramos en la laguna de Alcahozo, estableciendo la relación entre la precipitación y la evapotranspiración de referencia (ET_0) con el área de inundación

de la laguna, con el fin de esclarecer la conexión entre la meteorología y el patrón hídrico de la laguna a escala local. La ET_0 se utilizó como indicador del comportamiento e intensidad de la evaporación en nuestra área de estudio. En la figuras 4.11-12 se muestra el área inundada de Alcahozo, como se aprecia, laguna está seca en verano, siguiendo una tendencia estacional. En la figura 4.12 se observa que el área inundada está fuertemente ligada con el patrón de precipitaciones, siendo mayor el área de inundación en periodos más húmedos. Por otra parte, esta laguna presenta un periodo seco en verano, coincidiendo con los valores más altos de ET_0 (fig. 4.11).

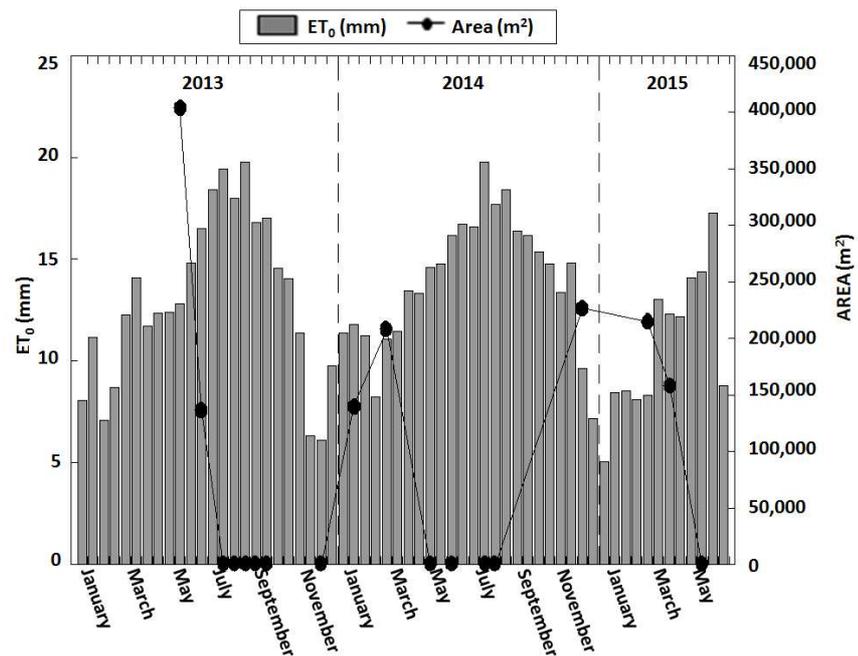


Figura 4.11.- Área inundada en relación con la evapotranspiración de referencia (ET_0) para el período 2013-2015 en la laguna de Alcahozo.

Al realizar el análisis de la superficie de agua de la laguna estimada mediante teledetección y la lluvia acumulada para diferentes periodos (7, 14, 30, y 90 días) se obtuvieron relaciones significativas ($p < 0,01$) para los periodos de 14, 30 y 90 días, mostrando valores de $R^2 \approx 0,7$. En cuanto al análisis de ET_0 también se estudiaron los mismos periodos que en el caso de la precipitación. En este caso los resultados

mostraron relaciones significativas ($p < 0,05$) para todos los periodos estudiados. En cuanto a los valores obtenidos del coeficiente de correlación fue aproximadamente del 70% en todos los casos, a excepción del periodo de 90 días que mostró los valores más bajos ($R^2 = 0,5$).

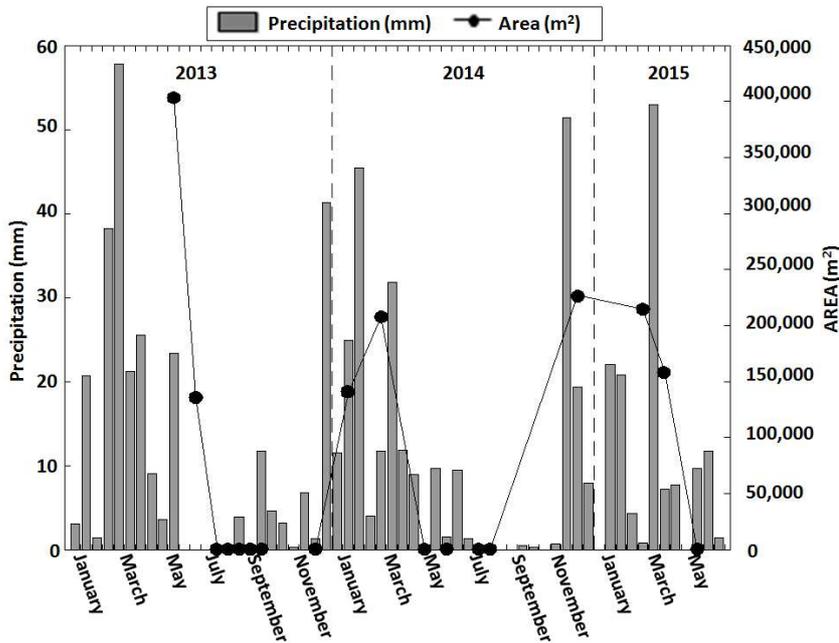


Figura 4.12.- Área inundada en relación con la precipitación para el período 2013-2015 en la laguna de Alcahozo.

4.5. Discusión

Los resultados generales muestran que la teledetección es una herramienta útil para el control de la calidad y la estimación de área de cubierta de agua en zonas de especial interés como son las lagunas someras y los humedales. En términos de calidad de aguas, los resultados son satisfactorios en cuanto a la estimación de concentración de clorofila-*a* y de partículas en suspensión, mediante métodos clásicos, cuando estos se aplican en la Albufera de Valencia los errores cuadráticos medios obtenidos son de $\pm 30 \text{ mg/m}^3$, en estimación de concentración de clorofila-*a*, y de $\pm 9 \text{ mg/l}$, en estimación de concentración de partículas en suspensión. Estos últimos valores son aceptables para esta masa de agua debido a los altos valores que presenta normalmente de estos

parámetros. En términos relativos los errores son del $\pm 24\%$ y $\pm 17\%$ para la concentración de clorofila-*a* y partículas en suspensión, respectivamente. Por otra parte se ha observado que el algoritmo de estimación de concentración de clorofila-*a* (fig. 4.1) es bastante sensible a valores del cociente TM4/TM2 menores a 1,5; por lo que en este caso, las incertidumbres en la estimación de este cociente podrían producir errores elevados en la estimación del parámetro. En el caso de la transparencia, su estimación mediante teledetección resulta más complicada, tal y como indican algunos estudios realizados en esta misma masa de agua, debido a su escaso intervalo de variación, en nuestro caso entre 10 y 40 cm, en sintonía con los 20-47 cm observados en el estudio de Domínguez-Gómez *et al.* (2012).

La resolución temporal de satélites como los de la serie Landsat no es suficiente para realizar un seguimiento de los parámetros de calidad en una masa de agua. Estos sensores poseen una resolución temporal de 16 días, insuficiente para captar con detalle los fenómenos acaecidos en un corto espacio de tiempo en la masa de agua. La importancia de la obtención de algoritmos para la estimación de calidad de agua con este sensor radica en su extenso archivo histórico de imágenes, que nos permite extender en el tiempo los estudios del comportamiento de los diferentes parámetros, con el fin de mejorar nuestro entendimiento de la dinámica de los sistemas y aplicar las acciones necesarias para su posterior conservación. Esta deficiencia en la resolución temporal se ha solventado con la utilización de otros sensores con bandas espectrales similares a las de Landsat-5. Por una parte se han aplicado los algoritmos desarrollados a imágenes Deimos-1, incrementando así la resolución temporal a 3 días, y por otra parte se han introducido técnicas de fusión de datos, consiguiendo así fusionar imágenes MODIS y Landsat-5-7 para obtener imágenes Landsat sintéticas sin la necesidad de que haya captura del sensor por la zona de interés. En ambos casos se han obtenido buenos resultados, similares, por ejemplo, a los mostrados por Chang *et al.* (2014) quienes también aplicaron la fusión de datos para el mismo fin. Las limitaciones de este tipo de técnicas se centran en la problemática de la presencia de nubes en la zona de estudio, como siempre que se trabaja con teledetección pasiva. También presenta limitaciones en cuanto al seguimiento de fenómenos rápidos (Gao *et al.*, 2006), es decir por ejemplo en la formación crecimientos masivos de fitoplancton.

Por otra parte, el problema con el algoritmo de estimación de transparencia en la Albufera se ha solucionado recurriendo a técnicas de aprendizaje automático. Con este tipo de técnicas se ha obtenido un algoritmo de estimación de transparencia con errores del 9%, sin necesidad además de datos adicionales para poder incrementar el intervalo de variación de los datos de campo. Con este tipo de técnicas se mejoran también los resultados de la obtención de concentración de clorofila-*a*, cuyo error es del 6%, mejorando otros algoritmos publicados para esta misma masa de agua (Domínguez *et al.* (2012)) con otros sensores cuya resolución espectral era más adecuada para este tipo de estudios.

Los resultados obtenidos a través de los productos simulados de Sentinel-2 muestran errores globales aceptables del $\pm 8\%$, $\pm 10\%$, $\pm 5\%$, en la estimación de la concentración de clorofila-*a*, concentración de partículas totales en suspensión y de transparencia, respectivamente. Centrándonos en la Albufera, los resultados son similares a los obtenidos anteriormente, es decir, la parte del estudio del desarrollo de los algoritmos genéticos con las imágenes fusionadas. Sin embargo, los bajos valores de clorofila-*a* y concentraciones de sólidos suspendidos observados en los dos embalses estudiados hacen difícil la elaboración de un algoritmo preciso específico para estos cuerpos de agua. Se necesitan medidas adicionales y un análisis más extenso y exhaustivo para poder extraer conclusiones sobre el rendimiento de estos algoritmos en embalses. Estos resultados muestran el potencial de la nueva era de sensores, como Sentinel-2, y de este tipo de técnicas para realizar el control y seguimiento de la evolución espacio-temporal de estos parámetros de calidad de agua.

Los algoritmos genéticos también proporcionan buenos resultados en estimación de la superficie cubierta por agua en lagunas someras, en comparación con los índices de agua y métodos de clasificación que se han analizado, así como con la utilización de bandas simples. El análisis llevado a cabo muestra una sobreestimación de los píxeles de agua por la banda del infrarrojo medio. Esta zona del espectro es especialmente sensible a la presencia de agua, por lo que esta sobreestimación puede deberse a la presencia de charcos o zonas de barrizal en la superficie de las lagunas. El índice de agua NDWI propuesto por Gao (1996), que utiliza las bandas de infrarrojo cercano y

medio, ofrece resultados peores comparados con otros índices estudiados en nuestra zona. Esto podría deberse a la sensibilidad de estas bandas espectrales a la presencia de agua (Bustamante *et al.*, 2006), lo cual también provoca errores en la identificación de píxeles de la cubierta de agua. Estos resultados coinciden con los publicados por Campos *et al.* (2012), quienes exponen que los índices que incluyen banda en el infrarrojo medio pueden actuar de forma diferente a la hora de detectar aguas estacionales y permanentes. Según su estudio, el NDWI que utiliza las bandas del infrarrojo cercano y medio detectaba mejor las masas de agua estacionales, aunque apreciaron sobreestimaciones significativas.

A pesar de que el algoritmo genético obtenido se aplicó a diferentes lagunas de nuestra área de estudio, y se trató de minimizar los efectos estacionales aplicando técnicas de normalización radiométrica en las imágenes, los cambios biofísicos en los cuerpos de agua podrían afectar a los resultados del algoritmo GP. Se ha tratado de minimizar estos factores mediante la aplicación del algoritmo para las diferentes estaciones del año y para diferentes niveles de agua de cada lago. Aunque el algoritmo mostró resultados aceptables, en el futuro será necesario extender la aplicación a una base de datos más completa con el fin de evaluar la robustez de la metodología descrita. Además, el fallo de adquisición de las imágenes Landsat-7 causó errores en la estimación de la superficie total de agua de algunas lagunas, resultando más vulnerables aquellas que presentan formas más irregulares, como es el caso de la laguna de Salicor.

El algoritmo propuesto también se puede aplicar a imágenes de otros sensores tales como Thematic Mapper, lo que nos permitirá llevar a cabo estudios retrospectivos prolongados en el tiempo gracias al extenso archivo histórico disponible de imágenes de este sensor, desde 1985, incluyendo la reconstrucción de series hidrológicas de las lagunas estudiadas u otras, y su comparación con series meteorológicas de la zona para conocer el comportamiento hidrológico de las lagunas. Por otra parte, esta metodología se podría aplicar a otros sensores tales como SPOT-5 o el reciente Sentinel-2, debido a la equivalencia espectral entre las bandas del infrarrojo cercano de estos sensores con ETM+. El futuro satélite de la NASA "Surface Water and Ocean Topographic (SWOT)", cuyo lanzamiento está previsto para 2020, es un buen ejemplo de la

aplicabilidad de este tipo de estudios. Además, la perspectiva de futuro de nuestro trabajo actual es el seguimiento a largo plazo de las lagunas para estudiar sus patrones hidrológicos, y el estudio de calidad de otras masas de agua en mayor detalle para ayudar a desarrollar planes de manejo adecuados para la protección de estos sistemas, así como para su conservación y recuperación bajo la Directivas de Hábitats y Marco del Agua, especialmente en regiones semiáridas como la cuenca mediterránea.

CAPÍTULO 5

Conclusiones

De acuerdo con el Artículo 9 del Reglamento sobre el depósito, evaluación y defensa de la tesis doctoral de la Universitat de València, las conclusiones se han redactado en inglés con el fin de optar a la mención internacional del título de doctor.

The main conclusions drawn from the analyses carried out about water quality and water area estimation using remote sensing techniques obtained in this Thesis are summarized following:

1. Empirical relationships based on classical regression methods have been obtained and tested for three main water quality variables using Landsat Thematic Mapper (TM) imagery and ground measurements from la Albufera de Valencia and three additional ponds dataset. Algorithms based on the reflectance values from the TM bands TM2/TM4, for chlorophyll-*a* concentration, TM4 for total suspended particles concentration, and TM2 for transparency showed the best results for the present study.
2. The algorithm for the estimation of chlorophyll-*a* concentration shows high sensitivity to TM2/TM4 reflectance values lower than 1.5. A small error in the estimation of the reflectance values TM2 or TM4 could lead to a significant uncertainty in the estimation of chlorophyll-*a* concentration.
3. Based on the spectral matching between Deimos-1 and TM, the application of these algorithms to Deimos-1 images will provide continuity to the algorithms designed for TM, improving the spatial resolution and reducing the revisit time to 3 days. The validation using Deimos-1 images with additional field campaigns in La Albufera de Valencia shows acceptable results in terms of chlorophyll-*a* concentration and total suspended particles concentration with error values of $\pm 22\%$ and $\pm 15\%$, respectively.
4. Some discrepancies arose between ground-based and remote sensing data when monitoring water transparency as a consequence of the low values of this parameter and the windy conditions prevailing in La Albufera de Valencia.
5. Data fusion of Landsat and MODIS images with the STAR-FM algorithm solves the problem of the Landsat low revisit time, providing more intensive Earth system observations. We have applied this tool to rapidly derive the heterogeneous maps addressing spatial distributions of chlorophyll-*a* concentration and water transparency, as well as their temporal changes in the Albufera de Valencia.
6. Genetic programming improves significantly the traditional efforts in la Albufera de Valencia by means of linear regression techniques without the necessity of additional databases. Reflectance values from synthetic Landsat images have been used as a basis to develop genetic program models for the estimation of

chlorophyll-*a* concentration and water transparency in la Albufera de Valencia obtaining validation error values of $\pm 6\%$ and $\pm 9\%$, respectively.

7. GP algorithms performed with simulated images of the recent Sentinel-2 show average errors of $\pm 8\%$, $\pm 10\%$, and $\pm 5\%$, in the estimation of chlorophyll-*a* concentration, total suspended particles concentration and transparency, respectively, in la Albufera de Valencia, Tous and Escalona reservoirs.
8. However, the low values of chlorophyll-*a* and suspended solids concentrations observed in the two reservoirs make difficult the training of an accurate algorithm specific for these water bodies. Additional measurements and further analysis are required before extracting any definitive conclusion about the performance of the algorithms in the studied reservoirs.
9. GP algorithms also shows the best performance to discern between water and non-water pixels and estimate the water area cover in shallow lagoons. The estimation of water pixels across a set of 13 shallow saline lakes in La Mancha Húmeda Biosphere Reserve in Spain shows a kappa value of 0.98 and a low error of 2%, both in training and validation.
10. GP algorithms could serve as a tool to monitor water quality in wetlands, which is linked to future water resources management. Also this could serve as a tool to improve our understanding of the temporal trends in seasonal lakes and their dependence on meteorological patterns, especially on rainfall. Such an approach provides a cost-effective way to monitor the flooded area variations related to changes of precipitation and evaporation, as well as to resource exploitation, in order to establish sustainable water resources management plans that preserve the ecological health of wetlands and lakes in semiarid basins with a high water stress, such as the area of La Mancha Húmeda Biosphere Reserve.
11. In brief, our advancements show an operational and cost effective way for monitoring water quality and hydrological patterns of wetlands so as to accomplish the requirements of the Water and Habitats Framework Directives.

Bibliografía

- Allan, M. G., Hamilton, D. P., Hicks, B. J. y Brabyn, L.** Landsat remote sensing of chlorophyll a concentrations in central North Island lakes of New Zealand. *Int. J. Remote Sens.*, 2011, 32, 2037-2055.
- Allee, R. J. y Johnson, J. E.** Use of satellite imagery to estimate surface chlorophyll-a and Secchi disc depth of Bull Shoals Reservoir, Arkansas, USA. *Int. J. Remote Sens.*, 1999, 20, 1057-1072.
- Ángel-Martínez, M.C.** *Aplicación de la Teledetección en la Localización de Superficies de Agua*; CEDEX: Madrid, España, 1994.
- APHA/AWWA/WEF.** Standard Methods for Examination of Water and Wastewater, 19th ed. Amer. Public Health Assoc., Washington, DC, 1992.
- Bai, J., Chen, X., Li, J., Yang, L. y Fang, H.** Changes in the area of inland lakes in arid regions of central Asia during the past 30 years. *Environ. Monit. Assess.*, 2011, 178, 247-256.
- Bilge, F., Yazici, B., Dogeroglu, T. y Ayday, C.** Statistical evaluation of remotely sensed data for water quality monitoring. *Int. J. Remote Sens.*, 2003, 24, 5317-5326.
- Bustamante, J., Díaz-Delgado, R., Aragonés, D. y Pacios, F.** Determining water body characteristics of Doñana shallow marshes through remote sensing. In Proceedings of the IEEE International Conference on Geoscience and Remote Sensing Symposium, Denver, CO, USA, 31 julio-4 agosto 2006; pp. 3662-3663.
- Campos, J.C., Sillero, N. y Brito, J.C.** Normalized difference water indexes have dissimilar performances in detecting seasonal and permanent water in the Sahara-Sahel transition zone. *J. Hydrol.*, 2012, 464-465, 438-446.
- Canty, M.J. y Nielsen, A.A.** Automatic radiometric normalization of multitemporal satellite imagery with the iteratively re-weighted MAD transformation. *Remote Sens. Environ.* 2008, 112, 1025-1036.
- Caselles, V., López García, M. J. y Soria, J. M.** Estudio del estado trófico del lago de la Albufera (Valencia) a partir de imágenes del Landsat-5 (TM). I Reunión Científica del Grupo de Trabajo de Teledetección, Barcelona, 10-11 diciembre, 1986.
- Castañeda, C. y Herrero, J.** Teledetección de cambios en la Laguna de Gallocanta. *Memorias Real Soc. Española Hist. Nat.* 2009, 7, 103-126.

Bibliografía

- Chander, G., Markham, B. L. y Helder, D. L.** Summary of current radiometric calibration coefficients for Landsat MSS, TM, ETM+, and EO-1 ALI sensors. *Remote Sens. Env.*, 2009, 113, 893-903.
- Chang, N. B., Vannah, B. W., Yang, Y. J., y Elovitz, M.** Integrated data fusion and mining techniques for monitoring total organic carbon concentrations in a lake. *Int J Remote Sens.*, 2014, 35, 1064-1093.
- Chavez, P. S. Jr.** Image-Based atmospheric corrections- revisited and improved. *Photogramm. Eng. Remote Sens.*, 1996, 62, 1025-1036.
- Chuvieco, E y Hantson, S.** Plan Nacional de Teledetección de Media Resolución. Procesamiento estándar de imágenes Landsat. Documento técnico de algoritmos a aplicar. 2010.
- DEIMOS Imaging S. L.** DEIMOS-1: Payload Description and Data Characteristics. Disponible en la ftp del Plan Nacional de Teledetección (PNT). 2010.
- Engelbrecht, A. P.** Computational Intelligence: An Introduction.. 2nd Edition. Wiley. 2007.
- European Commission, Directive 2000/60/EC** of the European Parliament and of the Council of 23 October 2000 establishing a framework for Community action in the field of water policy. Official Journal of the European Communities, 2000, 327, pp. 1-72.
- Domínguez, J. A.** Estudio de la calidad del agua de las lagunas de gravera mediante teledetección . Tesis doctoral, Universidad de Alcalá, Alcalá de Henares. 2002.
- Domínguez, J. A., Chuvieco, E. y Sastre, A.** Monitoring transparency in inland water bodies using multispectral images. *Int. J. Remote Sens.*, 2009, 30, 1567-1586.
- Domínguez, J. A., Alonso, C., Alonso A.** Remote sensing as a tool for monitoring water quality parameters for Mediterranean Lakes of European Union water framework directive (WFD) and as a system of surveillance of cyanobacterial harmful algae blooms (SCyanoHABs). *Env.Monit. Assess.*, 2011, 181, 317-334.
- Domínguez-Gómez, J. A., Rodríguez-Pérez, D., Marcos-Martín, C., Chao-Rodríguez Y. y Delgado-Rojas G.** Estudio de aguas continentales mediante teledetección. UNED, España, 2012.

- Duan, H. T., Zhang, Y. Z., Zhan, B., Song, K. S. y Wang, Z. M.** Assessment of Chlorophyll-a concentration and trophic state for Lake Chagan using Landsat TM and field spectral data. *Environ. Monit. Assess.*, 2007, 129, 295-308.
- El-Asmar, H.M. y Hereher, M.E.** Change detection of the coastal zone east of the Nile Delta using remote sensing. *Environ. Earth Sci.* 2011, 62, 769-777.
- Feyisa, G.L., Meilby, H., Fensholt, R. y Proud, S.R.** Automated water extraction index: A new technique for surface water mapping using Landsat imagery. *Remote Sens. Environ.* 2014, 140, 23-35.
- Fisher, A. y Danaher, T.** A water index for SPOT5 HRG satellite imagery, New South Wales, Australia, determined by linear discriminant analysis. *Remote Sens.* 2013, 5, 5907-5925.
- Florín, M. y Montes, C.** Functional analysis and restoration of Mediterranean lagunas in the Mancha Húmeda Biosphere Reserve (Central Spain). *Plant Ecol.* 1999, 109, 97-109.
- Francone, D.** *Discipulus Software Owner's Manual, Version 3.0 DRAFT*; Machine Learning Technologies, Inc.: Littleton, CO, USA, 1998.
- Gao, B.C.** NDWI—A normalized difference water index for remote sensing of vegetation liquid water from space. *Remote Sens. Environ.* 1996, 58, 257-266.
- Gao, F., Masek, J., Schwaller, M. y Hall, F.** On the Blending of Landsat and MODIS Surface Reflectance: Predicting Daily Landsat Surface Reflectance. *IEEE Trans. Geo. & Remote Sens.* 2006, 44, 2207-2218.
- Gardelle, J., Hiernaux, P., Kergoat, L. y Grippa, M.** Less rain, more water in ponds: A remote sensing study of the dynamics of surface waters from 1950 to present in pastoral Sahel (Gourma region, Mali). *Hydrol. Earth Syst. Sci. Discuss.* 2009, 6, 5047-5083.
- Genderen, J. y Pohl, C.** Image Fusion: Issues, Techniques, and Applications. Intelligent Image Fusion. Proc. 1994, EARSel Workshop, Strasbourg, France, 18-26.
- Giardino, C., Pepe, M., Brivio, P. A., Ghezzi, P. y Zilioli, E.** Detecting chlorophyll, Secchi disk depth and surface temperature in a sub-alpine lake using Landsat imagery. *Sci. Tot. Environ.*, 2001, 268, 19-29.

Bibliografía

- Gilabert, M. A., Conese, C. y Masselli, F.** An atmospheric correction method for the automatic retrieval of surface reflectances from TM images. *Int. J. Remote Sens.*, 1994, 15, 2065-2086.
- Gosálvez, R., Gil-Delgado, J. A., Vives-Ferrándiz, C., Sánchez, G. y Florín, M.** Seguimiento de aves acuáticas amenazadas en lagunas de la Reserva de la Biosfera de La Mancha Húmeda (España central). *Polígonos* 2012, 22, 89-122.
- Härmä, P., Vepsäläinen, J., Hannonen, T., Pyhälä, T., Kämäri, J., Kallio, K., Eloheimo, K. y Koponen, S.** Detection of water quality using simulated satellite data and semi-empirical algorithms in Finland. *Sci. Tot. Environ.* 2001, 268, 107- 121.
- Huete, A.R.** A soil-adjusted vegetation index (SAVI). *Remote Sens. Environ.* 1988, 25, 295-309.
- Jain, S.K., Singh, R.D., Jain, M.K. y Lohani, A.K.** Delineation of flood-prone areas using remote sensing techniques. *Water Resour. Manag.* 2005, 19, 333-347.
- Kloiver, S. M., Brenzonik, P. L., Olmanson, L. G. y Bauer, M. E.** A procedure for regional lake water clarity assessment using Landsat multispectral data. *Remote Sens. Environ.*, 2001, 82, 38-47.
- Laguna, C., Gosálvez, R., Sánchez, G., Falomir, J., Velasco, A., Florín, M., Gil-Delgado, J. y Chicote, A.** Climate change footprint in the Mancha húmeda biosphere reserve. In Proceeding of the Energy and Environment Knowledge Week, Toledo, España, 28-29 octubre 2013; pp. 183-185.
- Landis, J.R. y Koch, G.G.** The measurement of observer agreement for categorical data. *Biometrics*, 1977, 33, 159-174.
- Li, W., Du, Z., Ling, F., Zhou, D., Wang, H., Gui, Y., Sun, B. y Zhang, X.** A comparison of land surface water mapping using the normalized difference water index from TM, ETM+ and ALI. *Remote Sens.* 2013, 5, 5530-5549.
- Lira, J.** Segmentation and morphology of open water bodies from multispectral images. *Int. J. Remote Sens.* 2006, 27, 4015-4038.
- López-García, M. J. y Caselles, V.** Use of Thematic Mapper data to assess water quality in Albufera lagoon of Valencia (Spain). In 13th Annual Conference of the Remote Sensing Society. 7-11 septiembre 1987 (Nottingham: Remote Sensing Society), pp.510-

519.

- Ma, R., y Dai, J.** Investigation of chlorophyll-a and total suspended matter concentrations using Landsat ETM and field spectral measurement in Taihu Lake, China. *Int. J. Remote Sens.*, 2005, 26, 13, 2779-2787.
- Maglione, P.** Coastline extraction using high resolution WorldView-2 satellite imagery. *Eur. J. Remote Sens.* 2014, 685–699.
- Mancino, G., Nolè, A., Urbano, V., Amato, M. y Ferrara, A.** Assessing water quality by remote sensing in small lakes: the case study of Monticchio lakes in southern Italy. *iForest*, 2009, 2, 154-161.
- Masek, J.G., Vermote, E.F., Saleous, N., Wolfe, F., Huemmrich, F., Gao, F., Kutler, J. y Lim, T.K.** A Landsat surface reflectance data set for North America, 1990-2000, *Geosci. Remote Sens. Letters*, 2006, 3, 68-72.
- Mayo, M., Gitelson, A., Yazobi, Y. Z. y Ben-Avraham, Z.** Chlorophyll distribution in Lake Kinneret determined from Landsat Thematic Mapper data. *Int. J. Remote Sens.*, 1995, 16, 175-182.
- McCullough, I. M., Loftin C.S. y Sader S.A.** High-frequency remote monitoring 15 of large lakes with MODIS 500 m imagery. *Remote Sens. Environ.* 2012, 124, 234-241.
- McFeeters, S.K.** The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. *Int. J. Remote Sens.* 1996, 17, 1425–1432.
- 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. *Remote Sens.* 2013, 5, 3544–3561.
- Moser, L., Voigt, S. y Schoepfer, E.** Monitoring of critical water and vegetation anomalies of Sub-Saharan West-African wetlands. In Proceedings of the IEEE International Conference on Geoscience and Remote Sensing Symposium, Québec, Canada, 13-18 July 2014; pp. 3842–3845.
- Organización de Cooperación y Desarrollo Económico (OECD).** Eutrophication des Eaux. Methodes de surveillance, d'évaluation et de lutte (Eutrophication of water, monitoring, assessment and control). Paris, Francia, 1982.

Bibliografía

- Olmanson, L. G., Bauer, M.E. y Brezonik, P.L.** A 20-year Landsat water clarity 22 census of Minnesota's 10,000 lakes. *Rem. Sens. Environ.*, 2008, 112, 4086-4097.
- Ouma, Y.O. y Tateishi, R.** A water index for rapid mapping of shoreline changes of five East African Rift Valley lakes: An empirical analysis using Landsat TM and ETM+ data. *Int. J. Remote Sens.* 2006, 27, 3153-3181.
- Peña, R., Ruiz, A. y Domínguez, J. A.** Mapping of photosynthetic pigments in Spanish inland waters using MERIS imagery. *Envisat Symposium*, 6-10 Sep 2004, Salzburg, Austria.
- Picazo, A., Rochera, C., Vicente, E., Miracle M.R. y Camacho A.** Determination of photosynthetic pigments by spectrophotometric methods in stratified lakes: a critical analysis based on comparisons with HPLC determinations in a model lake. *Limnetica*, 2013, 32, 139-158.
- Pohl, C., y Van Genderen, J.** Multisensor image fusion in remote sensing: Concepts, methods, and applications. *Int. J. Remote Sens.* 1998, 19, 823-854.
- Romo, S., García-Murcia, A., Villena, M. J., Sánchez, V. y Ballester A.** Phytoplankton trends in the lake of Albufera de Valencia and implications for its ecology, management, and recovery. *Limnetica*, 2008, 27, 11-28.
- Rouse, J.W., Haas, R.H., Deering, D.W. y Sehell, J.A.** *Monitoring the Vernal Advancement and Retrogradation (Green Wave Effect) of Natural Vegetation*; Remote Sensing Center. Report RSC 1978-4; Texas A&M University: College Station, TX, USA, 1974.
- Sakamoto, T., Van Nguyen, N., Kotera, A., Ohno, H., Ishitsuka, N. y Yokozawa, M.** Detecting temporal changes in the extent of annual flooding within the Cambodia and the Vietnamese Mekong Delta from MODIS time-series imagery. *Remote Sens. Environ.* 2007, 109, 295-313.
- Schroeder, T.A., Wulder, M.A., Healey, S.P. y Moisen, G.G.** Mapping wildfire and clearcut harvest disturbances in boreal forests with Landsat time series data. *Remote Sens. Environ.* 2011, 115, 1421-1433.
- Seifert, J.W.** Data mining: An overview. In *National Security Issues*; Pegarkov, D.D., Ed.; Nova Science Publishers, Inc.: New York, NY, USA, 2006; pp. 201-217.
- Serrano M. L., Camacho, A., Vicente, E. y Peña, R.** Estudio por teledetección de la

- evolución del estado trófico de tres embalses del ámbito de la Confederación Hidrográfica del Júcar en el periodo estival de los años 1994 y 1995. *Limnetica*, 1997, 13, 5-14.
- Soliman, G. y Soussa, H.** Wetland change detection in Nile swamps of southern Sudan using multitemporal satellite imagery. *J. Appl. Remote Sens.* 2011, 5, 053517.
- Soria, J. M. y Vicente, E.** Estudios de los aportes hídricos al parque natural de la Albufera de Valencia. 2002, *Limnetica*, 2, 105-115.
- Strickland, J. D. H. y Parsons, T. P.** A practical handbook for seawater analysis. (Ottawa: Fisheries Research Board), 1968.
- Sun, F., Sun, W., Chen, J. y Gong, P.** Comparison and improvement of methods for identifying waterbodies in remotely sensed imagery. *Int. J. Remote Sens.* 2012, 33, 6854–6875.
- Vermote, E.F., Tanre, D., Deuze, J.L., Herman, M. y Morcrette, J.J.,** Second simulation of the satellite signal in the solar spectrum, 6S: an overview. *IEEE Trans. Geosci. & Remote Sens.*, 1997, 35, 675–685.
- Vicente E. y Miracle, M. R.** The coastal Lagoon Albufera de Valencia: an ecosystem under stress. *Limnetica*, 1992, 8, 87-100.
- Vidal, D., Anza, I., Taggart, M.A., Pérez-ramírez, E., Crespo, E., Hofle, U. y Mateo, R.** Environmental factors influencing the prevalence of a *Clostridium botulinum* type C/D mosaic strain in nonpermanent Mediterranean wetlands. *Appl. Environ. Microbiol.* 2013, 79, 4264–4271.
- Work, E.A. y Gilmer, D.S.** Utilization of satellite data for inventorying prairie ponds and lakes. *J. Photogramm. Eng. Remote Sens.* 1976, 42, 685–694.
- Xu, H.** Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery. *Int. J. Remote Sens.* 2006, 27, 3025–3033.
- Yang, J., Weisberg, P.J., Bristow, N.A.** Landsat remote sensing approaches for monitoring long-term tree cover dynamics in semi-arid woodlands: Comparison of vegetation indices and spectral mixture analysis. *Remote Sens. Environ.* 2012, 119, 62–71.

Anexo A

Artículo “Empirical relationships for monitoring water quality of lakes and reservoirs through multispectral images”

En este primer trabajo, realizado en colaboración con el Dr. J. A. Domínguez y publicado en IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing en mayo de 2014, se describe como se han utilizado relaciones empíricas para la estimación de los parámetros de calidad de agua estudiados a partir de datos de Landsat Thematic Mapper (TM). En este caso, se utilizaron datos *in situ* de varias masas de agua españolas que cubrían una variedad de estados tróficos diferentes, abarcando de oligotróficos a hipereutróficos. Aplicando estos algoritmos a imágenes Deimos-1, aprovechando su equivalencia espectral con TM, se mejoró la resolución temporal a 3 días y se aumentó la resolución espacial a 22 m.

La revista IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing poseía un factor de impacto de **3,03** y ocupaba la posición relativa **7/28**, dentro de la categoría de Remote Sensing de la base de datos ISI Web of Knowledge en el año de publicación.

Empirical Relationships for Monitoring Water Quality of Lakes and Reservoirs Through Multispectral Images

Carolina Doña, Juan M. Sánchez, Vicente Caselles, Jose Antonio Domínguez, and Antonio Camacho

Abstract—Remote sensing techniques can be used to estimate water quality variables such as chlorophyll *a*, total suspended particles, and water transparency. This paper describes empirical algorithms for the estimation of these variables using Landsat Thematic Mapper (TM) data. Ground data were taken from several Spanish lakes covering a variety of trophic statuses, ranging from oligotrophic to hypereutrophic. The studied lakes were the Albufera de Valencia and lakes and ponds of the Southeast Regional Park in Madrid. Empirical equations were obtained to estimate chlorophyll *a* from the ratio in reflectance values between bands 2 and 4 of TM ($R^2 = 0.66$, $p < 0.001$), transparency [Secchi disk (SD)] from reflectance in band 2 ($R^2 = 0.80$, $p < 0.001$), and total suspended particles from reflectance in band 4 ($R^2 = 0.92$, $p < 0.001$). The spectral equivalence between TM and the recent satellite Deimos-1 was also tested. By applying the proposed algorithms to this new sensor, the temporal resolution is improved by up to 3 days, which increases spatial resolution to 22 m. The algorithms were validated using three Deimos-1 scenes of the Albufera de Valencia together with ground measurements. Results of this validation showed root-mean-square errors (RMSEs) of $40 \text{ mg} \cdot \text{m}^{-3}$ for Chl-*a* (data range: $32\text{--}238 \text{ mg} \cdot \text{m}^{-3}$), $10 \text{ mg} \cdot \text{L}^{-1}$ for total suspended solid (TSS) (data range: $25\text{--}89 \text{ mg} \cdot \text{L}^{-1}$), and 0.10 m for SD (data range: $0.17\text{--}0.40 \text{ m}$). In any case, these results show the potential of Deimos-1 as a substitute of TM in water quality monitoring in small/medium water bodies, providing continuity to three decades of TM imagery.

Index Terms—Albufera de Valencia, Deimos-1, lakes, reservoirs, Thematic Mapper (TM), trophic status.

I. INTRODUCTION

THE trophic status of lakes and reservoirs has been studied for years as water quality can deteriorate due to eutrophication. Evidence of eutrophication, caused by the excess of nutrients such as phosphorus and nitrogen, includes the increase in turbidity and the change in water colour toward green due to an increase in phytoplankton abundance. The Organization for

Economic Co-operation and Development (OECD) [1] has defined eutrophication as the water enrichment in nutritive substances that generally leads to changes such as the increase in algae production and other aquatic plants, and the deterioration of water quality and aquatic ecosystem. The European Water Framework Directive [2] establishes that all countries of the European Union have to avoid the deterioration, to improve and retrieve the state of the superficial water bodies and maintain their good ecological status. It also establishes that the protected areas should be maintained and that management must be based on scientific data.

Remote sensing techniques are useful for the study of water quality of lakes and reservoirs. Different sensors have been used with this aim within the last few years. For example, Landsat Thematic Mapper (TM) images have largely been used to determine variables such as chlorophyll *a* concentration [3]–[5], water transparency [6], [7], or both [8]–[10], as well as suspended particles (seston) concentration in the water [11], [12]. The spatial resolution of the TM sensor is appropriate for the study of small- to medium-sized water bodies, but its spectral resolution is not enough to determine some water quality variables. Spectral resolution of more recent sensors, such as MODIS, onboard TERRA and AQUA, and MERIS, onboard ENVISAT, is more suitable to determine the different water quality variables [13, 14], but in these cases, the spatial resolution is too coarse to analyze medium- or small-sized water bodies [10]. Comparative studies with the three mentioned sensors [15], [16] have shown the potential of TM to monitor water quality variables such as transparency and suspended particles in the water, despite its low spectral resolution. Images from the new satellite Deimos-1 (or Spain-DMC) are now available. This sensor operates in three of the TM optical bands, improving the temporal resolution by up to 3 days, and increasing the spatial resolution to 22 m compared to 30 m of TM. This first Spanish satellite is operated by Deimos Imaging S.L. and is part of the Spanish contribution for the Disaster Monitoring Constellation (DMC) [17].

López-García and Caselles [11] were the first to study the trophic status of the Albufera de Valencia lake (Valencia, Spain) from Landsat-TM data. These authors obtained empirical models for the chlorophyll *a* and seston using TM band 1 (452–518 nm), but they could not define a transparency model. Domínguez *et al.* [12] studied the trophic status of lakes/ponds located in the Southeast Regional Park, 20 km southeast of Madrid, Spain, where they proposed models for chlorophyll *a* and transparency using TM band 2 (528–609 nm), and for the concentration of total suspended particles using TM band 4 (776–904 nm).

Manuscript received March 27, 2013; revised December 11, 2013; accepted January 06, 2014. Date of publication February 02, 2014; date of current version June 13, 2014. This study was jointly supported by the “Ministerio de Economía y Competitividad” by aid granted (Projects CGL2007-64666-CL1, CGL2007-29819E, CGL2010-17577-CL1 and CGL2012-38909) and “Generalitat Valenciana” (Project PROMETEO/2009/086).

C. Doña and V. Caselles are with the Department of Earth Physics and Thermodynamics, Faculty of Physics, University of Valencia, E-46100 Burjassot, Valencia, Spain (e-mail: Carolina.Dona@uv.es).

J. M. Sánchez is with the Department of Applied Physics, University of Castilla-La Mancha, E-13400 Almadén, Ciudad Real, Spain.

J. A. Domínguez is with the Department of Mathematical and Fluids Physics, Science School, Spanish Open University (UNED), E-28040 Madrid, Spain.

A. Camacho is with the Cavanilles Institute of Biodiversity and Evolutionary Biology, University of Valencia, E-46100 Burjassot, Valencia, Spain.

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/JSTARS.2014.2301295

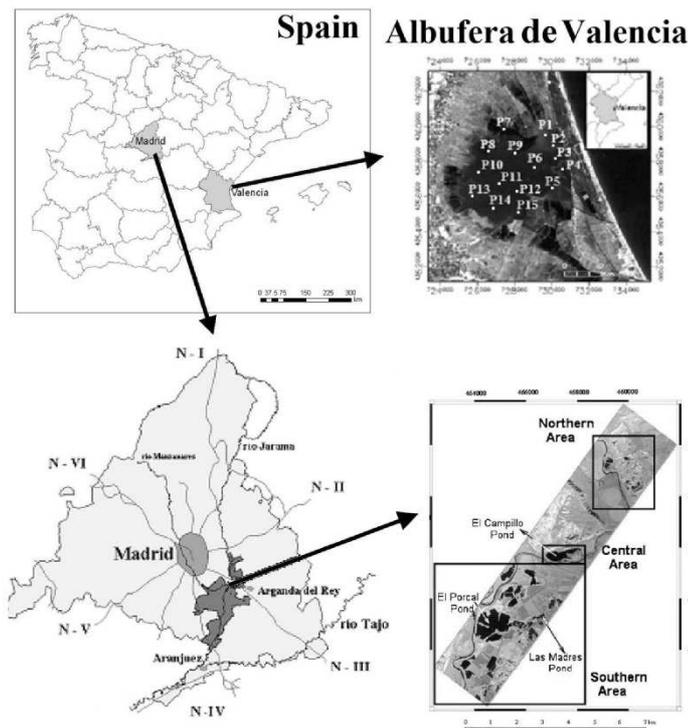


Fig. 1. Study areas.

The primary aim of our work is to develop workable algorithms for the estimation of water quality variables that can be applied to some Spanish water bodies, using the Landsat TM bands proposed by Domínguez *et al.* [12]. For this, data from two different sites, the Albufera de Valencia and lakes at the Southeast Regional Park in Madrid, were used. A secondary goal is the application of these algorithms to a set of Deimos-1 images of the Albufera de Valencia, to show the potential of this new sensor to monitor the trophic status of highly dynamic inland water bodies. With this work, we intend to implement a simple and operative tool that can be available for the monitoring of multiple variables related to the ecological status of Spanish small-medium water bodies, according to the European Water Framework Directive [2].

II. STUDY SITE AND MATERIALS

A. Study Area

Albufera de Valencia is a coastal freshwater lake located on the Mediterranean coast, 10 km south of Valencia, Spain (Fig. 1). This lake is connected to the sea by one artificial and two natural channels, where the water flow is regulated by gates keeping the water level appropriate for rice cultivation, as this is the main crop activity in the area. Water inflows to the Albufera come

mainly from irrigation channels bringing waters from the Turia and Júcar rivers, and ravines which carry rain water. It also receives treated wastewater and storm water from the urban and industrial areas nearby. The Albufera Lake is a hypereutrophic system with registered concentrations of chlorophyll *a* of up to $717 \text{ mg} \cdot \text{m}^{-3}$ [18], with average annual concentrations higher than $100 \text{ mg} \cdot \text{m}^{-3}$. After its recognition as a Natural Park in 1985, and its inclusion 4 years later as part of the Ramsar list of international protected wetlands, Albufera de Valencia has been the target of numerous limnological studies [19].

The other water bodies included in this study are located in the Southeast Regional Park, a natural reserve located 20 km south-east of Madrid, Spain. These water bodies, along with others in the area, are artificial and created as a result of mining activities to extract gravel from a river. The trophic status of these water bodies spans from oligotrophic to hypereutrophic largely as a result of several human impacts. For example, a restoration plan was implemented for Las Madres Pond in the 1990s which resulted in the improvement of water quality driving to a mesotrophic system that supports recreational and educational uses. However, Campillo Pond has not received any efforts at restoration since mining operations were concluded and displays eutrophic conditions. Finally, Porcal Pond, for which mining operations still continue, suffers a high water quality degradation

TABLE I
GROUND-BASED DATA RANGE FROM [11] AND [12] DATABASES

Lake	[Chl- <i>a</i>] (mg·m ⁻³)	[TSS] (mg·L ⁻¹)	Secchi disk (m)
Las Madres	0.6–2.8	1.2–8.9	3.3–4.3
Campillo	10.7–38.2	12.1–18.0	1.0–1.3
Porcal	39.0–81.5	20.2–15.6	0.5–0.8
Albufera	210–717	37–80	0.15–0.28

and is currently a hypereutrophic system. These water bodies do not have any channelized inflow or outflow of surface water. Moreover, the low amount of inorganic solid particles in these lakes allows high light penetration that can be used profitably by phytoplankton, leading to high chlorophyll *a* concentrations when nutrients are available [12].

B. Ground-Based Data

Ground-measured data used in this work correspond to chlorophyll *a* concentration [Chl-*a*] (mg·m⁻³), water transparency (Secchi disk transparency SD, in m), and total suspended particles concentration [TSS] (mg·L⁻¹). Three different databases were used: 1) data of the Albufera from [11] corresponding to July 22, 1985 (20 sampling points spread all over the lake, see Table I); 2) data of the Southeast Regional Park lakes provided by Domínguez *et al.* [12] (18 sampling points in Porcal, 10 in Las Madres, and 14 in Campillo, see Table I; and 3) data from a new field campaign carried out in the Albufera in 07/20/2010, 08/05/2010, 01/23/2011, 04/12/2011, and 07/13/2011 (15 sampling points covered within 3 h, centred in the sensor overpass time, Fig. 1).

Chlorophyll *a* concentration, used as an indicator of algal biomass, was determined spectrophotometrically after pigment extraction in acetone according to Strickland and Parsons [20]. The concentration of total suspended solids [TSS], which includes dead, inert, and degraded particles, was determined by filtering the water samples through pre-weighted AH—934 glass—fibre filters, then [TSS] was obtained after drying for 5 h at 103 °C by considering the final weight and the volume of the initial sample [21]. Water transparency was measured with a Secchi disk (SD). This method has been shown to be acceptable for measuring transparency [22] and it can also be used to determine the trophic status of freshwater [1]. Transparency can be affected by water colour, algae, and suspended particles. Obviously, higher total suspended particles and chlorophyll *a* concentrations correspond to lower transparencies.

C. Satellite Data

Images used in this work correspond to Landsat-5 and Deimos-1 satellites. TM images are available online at the United States Geological Survey web (USGS) (<http://glovis.usgs.gov/>), with the exception of the 1985 Albufera image that was provided by the Thermal Remote Sensing group of the University of Valencia. Deimos-1 images were provided by Deimos Imaging.

The extensive USGS archive of TM images allows the user to conduct temporal studies of the quality of water bodies for the

TABLE II
MAIN FEATURES OF LANDSAT 5-TM AND DEIMOS-1

Sensor	Band	Resolution			Radiometric (bits)	Cover area (km ²)
		Spectral (µm)	Spatial (m)	Temporal (days)		
Landsat 5-TM	1	0.452–0.518	30			185 x 170
	2	0.528–0.609	30			
	3	0.626–0.693	30			
	4	0.776–0.904	30	16	8	
	5	1.567–1.784	30			
	6	10.45–12.42	120			
	7	2.097–2.349	30			
Deimos-1	2	0.520–0.600	22			300 x 300
	1	0.630–0.690	22	3	8	
	0	0.770–0.900	22			

TABLE III
SENSORS AND SCENES USED IN THIS WORK

Images	Dates
Landsat-5 (Path/row: 198/33)	07/22/1985; 07/20/2010; 08/05/2010
Landsat-7 (Path/row: 199/33)	09/26/2006
Deimos-1	07/19/2010; 01/23/2011; 04/12/2011; 07/13/2011

TABLE IV
RANGES OF GROUND-BASED DATA FOR THE SAMPLING CAMPAIGNS 2010/2011 FROM THE ALBUFERA DE VALENCIA

Date	[Chl- <i>a</i>] (mg·m ⁻³)	[TSS] (mg·L ⁻¹)	Secchi disk (m)
07/20/2010	45–99	68–87	0.10–0.20
08/05/2010	24–74	42–75	0.14–0.30
01/25/2011	32–159	25–54	0.20–0.40
04/12/2011	147–238	53–81	0.17–0.26
07/13/2011	50–72	62–89	0.22–0.29

past three decades. Data acquisition by TM halted on November 2011 as a consequence of a failure in an electronic component, leaving three decades of historical registers. The spectral band equivalence between TM and Deimos-1 provides continuity to the algorithms designed for TM. The main features of both sensors are listed in Table II.

Remote sensing data of the Southeast Regional Park lakes were provided by Domínguez *et al.* [12].

Furthermore, an extra ETM+ image of the Albufera Lake was used to test different atmospheric correction methods. Sensors and dates of the images used in this work are shown in Table III.

III. METHODOLOGY

A. Remote Sensing Data Processing

Pre-processing of the Landsat and Deimos-1 images consisted of radiometric calibration and atmospheric correction. Radiometric calibration was needed to convert data measured by the sensor [Digital number (DN)] in radiance, L ($W \cdot m^{-2} \cdot sr^{-1} \cdot \mu m^{-1}$), for every spectral band “*i*”

$$L_i = G_i \cdot DN_i + B_i \quad (1)$$

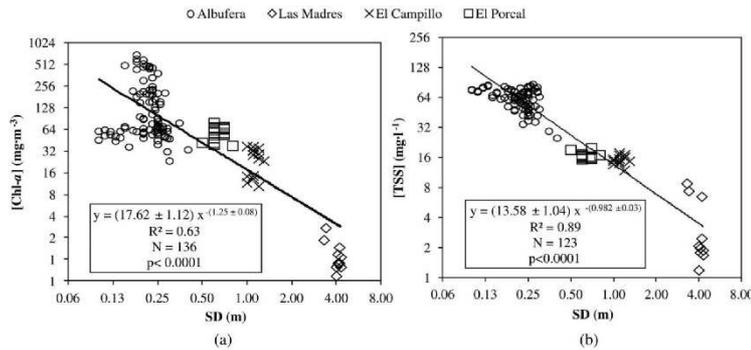


Fig. 2. Relationship between [Chl-*a*] (a) and [TSS] (b) with SD depth.

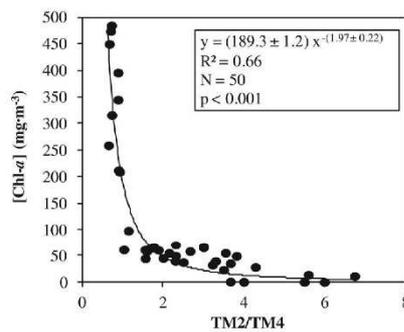


Fig. 3. Power relationship between [Chl-*a*] and the ratio TM2/TM4 reflectance values.

where G_i and B_i are the calibration coefficients. For the Landsat images, these coefficients are taken from Chander *et al.* [23], whereas they are included in the image metadata in the case of the Deimos-1 images.

Different atmospheric correction methods were first evaluated in order to select the best model to remove the atmospheric contribution. For this, *in-situ* radiometric data of the Albufera water were collected, concurrent with a Landsat-7 overpass, using an ASD-FR spectroradiometer [24]. The following atmospheric correction methods were applied to bands 1–4 of the ETM+ scene: methods based in the Dark Object Subtraction (DOS): 1) Chavez [25], [2] Gilabert *et al.* [26]; and analytical methods: 3) ATCOR3 [27], and 4) MODTRAN5 [28]. Reflectance results after correction were compared with the *in-situ* reflectance data. Results pointed out that DOS-based methods (Chavez [25] and Gilabert *et al.* [26]) showed the best performance and thus lowest errors (percentage error values lower than 10%). This is in agreement with other comparative studies on water bodies, which show that these simple methods provide a reasonable correction in TM bands 1–3 [29]. Particularly, Domínguez *et al.* [12] obtained the best results for the atmospheric correction through the Gilabert *et al.* [26] method using

the TM bands 1–4. In this work, the Chavez [24] method was finally applied since a blue band, necessary in the method proposed by Gilabert *et al.* [26], is not available in Deimos-1.

Finally, water/land pixels were discriminated using a simple threshold in the near infrared band for each scene [12].

B. Algorithm Development

Different algorithms to determine [Chl-*a*], [TSS], and SD were obtained from ground-based data and TM reflectance values through simple regression techniques. The entire dataset was split into two groups, one for the algorithm development (2/3 data) and the second for model testing (1/3 data). Linear, power, logarithmic, exponential, and second-third order polynomial functions were evaluated to obtain the best fit between ground-based data and TM reflectance. In this work, we focused on bands TM2 and TM4 since these match Deimos-1 bands, and the absolute maximum reflectance of the chlorophyll *a* falls within the TM2 (528–609 nm) spectral range. Also, the [TSS] spectral response shows a relative maximum at 810 nm, within the spectral range of band 4 (776–904 nm) [12]. Nevertheless, some other band combinations proposed by different authors were also tested.

These sensors have an adequate spectral resolution to monitor water quality variables like TSS or SD, but their bands are too broad to provide unequivocal measurements of chlorophyll because other optical properties of water, including CDOM, suspended matter, and other plant pigments, also affect measured light reflectance within the bands. Some studies, such as Olmanson *et al.* [16], obtained acceptable results for Chl-*a* because water clarity in the lakes they studied was strongly correlated with phytoplankton concentrations. In our work, we tested whether we could obtain acceptable results in the estimation of this variable in our study area.

Finally, regressions showing the best performance based on statistic parameters were selected. Statistical analyses for the comparison between algorithm predictions and field measurements were done as described in [30], including the root-mean-square error (RMSE), the mean absolute difference (MAD), the mean absolute percent difference (MAPD), and the biased estimator (BIAS).

TABLE V
ALGORITHMS OF TOTAL SUSPENDED PARTICLES ESTIMATION WITH TM BANDS REFLECTANCE DATA

Author	Algorithm	a	b	R ²
Proposed algorithm	b · (TM4) ^a	1.07 ± 0.22	15.9 ± 1.1	0.92
Domínguez <i>et al.</i> [12]	a · ln(TM4) + b	26 ± 2	30 ± 2	0.84
Ma <i>et al.</i> [31]	a · (TM4) ² + b	3.0 ± 0.3	18 ± 3	0.73
Ma <i>et al.</i> [31]	a · TM4 + b	16.9 ± 1.0	4 ± 3	0.87

a and b are the regression coefficients, R² is the fit goodness, N = 47 and all algorithms presented p-values <0.001.

TABLE VI
ALGORITHMS OF WATER TRANSPARENCY ESTIMATION WITH TM BANDS REFLECTANCE DATA

Author	Algorithm	a ₁	a ₂	a ₃	a ₄	b	R ²
Proposed algorithm	b · (TM2) ^{-a₁}	1.82 ± 0.12	–	–	–	4.5 ± 1.2	0.80
Allan <i>et al.</i> [5]	b · a ₁ · (TM4)	0.24 ± 0.04	–	–	–	1.27 ± 0.15	0.70
Wu <i>et al.</i> [32]	exp[b · (a ₁ · TM1) - (a ₂ · TM3)]	0.27 ± 0.05	0.65 ± 0.06	–	–	1.3 ± 0.2	0.77
Cozar <i>et al.</i> [33]	b · (a ₁ · TM1) - (a ₂ · TM3) - (a ₃ · TM4)	0.16 ± 0.05	0.35 ± 0.07	0.09 ± 0.04	–	2.17 ± 0.19	0.76
Olmanson <i>et al.</i> [6]	exp[(a ₁ · TM1/TM3) - (a ₂ · TM1) - b]	2.2 ± 0.2	1.10 ± 0.09	–	–	0.58 ± 0.14	0.80
Lavery <i>et al.</i> [34]	b · (a ₁ · TM3) - (a ₂ · TM1/TM3)	0.56 ± 0.07	0.42 ± 0.13	–	–	2.5 ± 0.3	0.60
Mancino <i>et al.</i> [10]	b · (a ₁ · TM3/TM2) + (a ₂ · TM1/TM2) - (a ₃ · TM1) + (a ₄ · TM2/TM1)	2 ± 1	3.6 ± 0.8	0.67 ± 0.07	0.04 ± 0.09	1 ± 1	0.80
Allan <i>et al.</i> [5]	exp[b · a ₁ · lnTM3]	1.73 ± 0.15	–	–	–	0.51 ± 0.16	0.79
Allee <i>et al.</i> [8]	b · (a ₁ · TM3) + (a ₂ · TM3 ²) - (a ₃ · TM3 ³)	5.4 ± 0.4	1.59 ± 0.19	0.15 ± 0.02	–	6.2 ± 0.3	0.90

a_i and b are the regression coefficients, R² is the fit goodness, N = 50 and all algorithms presented p-values <0.001.

TABLE VII
STATISTICAL ANALYSIS OF THE LINEAR FIT BETWEEN PREDICTED AND OBSERVED [CHL-a], [TSS] AND SD USING THE TEST DATA FOR DIFFERENT ALGORITHMS

Author	Variable	TM bands	a	b	R ²	Bias	MAD	RMSE	MADP (%)
Proposed algorithm	[CHL-a] (mg·m ⁻³)	TM2/TM4	0.88 ± 0.03	7 ± 5	0.96	-4	22	30	24
Proposed algorithm	[TSS] (mg·L ⁻³)	TM4	0.97 ± 0.06	-2 ± 3	0.91	-3	7	9	17
Domínguez <i>et al.</i> [12]			0.85 ± 0.05	6 ± 3	0.85	-0.0003	10	12	25
Ma <i>et al.</i> [31]			0.76 ± 0.06	10 ± 3	0.76	0.3	12	15	30
Ma <i>et al.</i> [31]			0.90 ± 0.05	4 ± 2	0.89	0.2	7	10	17
Proposed algorithm	SD (m)	TM2	0.87 ± 0.03	0.03 ± 0.04	0.92	-0.07	0.21	0.3	30
Wu <i>et al.</i> [32]		TM1, TM3	0.58 ± 0.03	0.13 ± 0.02	0.89	-0.10	0.22	0.4	30
Cózar <i>et al.</i> [33]		TM1, TM3, TM4	0.67 ± 0.07	0.16 ± 0.06	0.76	-0.04	0.3	0.60	50
Olmanson <i>et al.</i> [6]		TM1/TM3, TM1	0.71 ± 0.03	0.09 ± 0.02	0.80	0.50	0.21	0.3	30
Allan <i>et al.</i> [5]		TM3	0.84 ± 0.04	0.02 ± 0.03	0.79	0.08	0.21	0.3	30
Allee <i>et al.</i> [8]		TM3	0.85 ± 0.04	0.18 ± 0.03	0.90	0.15	0.3	0.4	40

p-values <0.001; a and b are the slope and the intercept, respectively, of the linear regression between the predicted (P_i) and the observed (O_i) data values; R² is the fit goodness; BIAS; MAD; RMSE; MAPD: $BIAS = \sum_{i=1}^n (P_i - O_i) / n$; $MAD = \sum_{i=1}^n |P_i - O_i| / n$; $RMSE = [\sum_{i=1}^n (P_i - O_i)^2 / n]^{1/2}$; and $MADP = 100 \cdot MAD \cdot \langle O \rangle$.

C. Application of Algorithms to Deimos-1 Imagery

The algorithms extracted for TM were then applied to the Deimos-1 images of the Albufera de Valencia (Table III). Results were validated by comparison with ground measured data for the different water quality variables. Statistical analyses of the comparison between algorithm

predictions and field measurements were made as explained above. Once validated, the proposed algorithms were applied to a set of TM and Deimos-1 images to obtain maps of the spatial distribution of the above mentioned variables in the Albufera de Valencia for different seasons.

TABLE VIII
STATISTICAL ANALYSIS OF THE LINEAR FIT BETWEEN DEIMOS-1 AND TM PREDICTED [Chl-*a*], [TSS] AND SD FOR JULY 19–20, 2010 IN THE ALBUFERA DE VALENCIA

Variable predicted	$x = \text{TM predicted}$ $y = \text{Deimos predicted}$	R^{2a}	RMSE	p-value
[Chl- <i>a</i>] ($\text{mg} \cdot \text{m}^{-3}$)	$y = (0.990 \pm 0.004) \cdot x$	0.87	8	< 0.05
[TSS] ($\text{mg} \cdot \text{L}^{-1}$)	$y = (0.972 \pm 0.004) \cdot x$	0.78	5	< 0.01
SD (m)	$y = (1.044 \pm 0.005) \cdot x$	0.73	0.015	< 0.01

R^{2a} is the adjusted fit goodness.

TABLE IX
STATISTICAL ANALYSIS OF THE LINEAR FIT BETWEEN PREDICTED AND OBSERVED [Chl-*a*], [TSS], AND SD USING THE VALIDATION DATA FOR DEIMOS-1 SCENES OF THE ALBUFERA DE VALENCIA LAKE

Variable	<O>	<P>	σ	Bias	MAD	RMSE	MADP (%)
[Chl- <i>a</i>] ($\text{mg} \cdot \text{m}^{-3}$)	120	140	50	30	30	40	22
[TSS] ($\text{mg} \cdot \text{L}^{-1}$)	63	64	11	1.3	8	10	15
SD (m)	0.25	0.30	0.09	0.06	0.09	0.10	40

$\langle O \rangle$ is the average observed data values; $\langle P \rangle$ is the average predicted data values; σ the standard deviation; BIAS; MAD; RMSE; MADP: $\text{BIAS} = \sum_{i=1}^n (P_i - O_i)/n$; $\text{MAD} = \sum_{i=1}^n |P_i - O_i|/n$; $\text{RMSE} = [\sum_{i=1}^n (P_i - O_i)^2/n]^{1/2}$; and $\text{MADP} = 100 \cdot \text{MAD} \cdot \langle O \rangle$.

IV. RESULTS AND DISCUSSION

In this section, we present the steps of the application of the methodology explained above.

A. Ground-Based Data

Data from the Albufera field campaign 2010/11 confirm that this water body is a hypereutrophic system with minimal transparencies of 0.10 m (SD) and maximum values of $238 \text{ mg} \cdot \text{m}^{-3}$ of [Chl-*a*] (chlorophyll *a* concentration) and $89 \text{ mg} \cdot \text{L}^{-1}$ of [TSS] (total suspended solids concentration), respectively (Table IV) [1].

Fig. 2 shows the relationship between the in-situ data of SD and [Chl-*a*] [Fig. 2(a)] and [TSS] [Fig. 2(b)] for the whole dataset. The different water bodies can be easily discerned in Fig. 2; e.g., the highest values of SD corresponding to Las Madres Lake or the highest values of both [Chl-*a*] and [TSS] for the Albufera. Also note the small values of water transparency in Albufera.

B. Algorithm Retrieval

1) *Chlorophyll a Concentration*: In an initial attempt, we tried to relate Chl-*a* levels directly to TM2 as a single input variable, following Domínguez *et al.* [12], but results were not satisfactory due to the low fitting value and high errors obtained ($R^2 = 0.40$). Then, further band combinations were evaluated, and the best correlation between TM reflectances and [Chl-*a*] measurements was shown for the ratio TM2/TM4 in a power function. A regression coefficient of $R^2 = 0.66$ was obtained in this case, $p < 0.001$ (Fig. 3), in comparison with the R^2 values lower than

0.50 obtained for the rest of the band combinations tested. In Fig. 3, we can also see that the algorithm is highly sensitive for TM2/TM4 values lower than 1.5, therefore, a small error in the estimation of TM2 or TM4 could allow considerable errors in the estimation values of [Chl-*a*].

The proposed algorithm was then applied to the second group of data reserved for testing and results were compared with the ground measured data. Table VII summarizes the regression parameters and statistical analyses of the comparison between modelled and observed [Chl-*a*]. A regression coefficient of 0.96 and a MADP value of $\pm 24\%$ were obtained. In our case, the error values obtained were acceptable, this may be because in this water bodies the Chl-*a* is highly correlated with TSS such as we mentioned above for Olmanson *et al.* [16].

2) *Total Suspended Particles Concentration*: In this case, band 4 of TM showed the best results for the estimation of [TSS]. Besides the power function, algorithms published by Domínguez *et al.* [12] and Ma *et al.* [31] showed acceptable results, too (Table V).

Regarding the comparison between modeled and measured values of [TSS] for the testing dataset, note that the power algorithm presented the best results with an MADP of $\pm 17\%$ and a regression coefficient of 0.91 (Table VII).

3) *Transparency (SD)*: A large number of models to estimate SD from TM reflectance information can be found in the appropriate literature. Only algorithms presenting a R^2 value over 0.70 are included in the statistical analysis summarized in Table VI. Results showed that equations proposed in this study, Allan *et al.* [5], Wu *et al.* [32], Cózar *et al.* [33], Mancino *et al.* [10], Lavery *et al.* [34], Olmanson *et al.* [6], McCullough *et al.* [7], and Allee and Johnson [8] yielded good results in SD estimation from different combinations of TM bands (1–4).

Regarding the comparison between observed and modeled values, now in terms of SD, all the algorithms listed in Table VI showed R^2 values over 0.75, except that proposed by Allan *et al.* [5]. Algorithms using TM2 band, proposed algorithm, TM3 band [5], and TM1–TM3 band combination [32], [6], and [7], showed the best values of MADP ($\pm 30\%$), and the overall best results were shown by the power fit function with TM2 reflectance. Although the algorithms published by Olmanson *et al.* [6] and Allan *et al.* [5] using TM1 and TM3 bands also show good results, we ultimately chose the power TM2 algorithm since this spectral band is integrated in Deimos-1 and the model can then be operationally applied to imagery from this sensor.

C. Application of Algorithms to Deimos-1 Imagery

The algorithms previously developed for the estimation of the different water quality variables were applied to the Deimos-1 images. In order to test the feasibility of Deimos-1 in comparison with TM, the water quality variables in the Albufera monitored by both sensors, with only 1 day of delay (July 19 and 20 for Deimos-1 and TM, respectively), were compared. Results showed slope values very close to one and regression coefficients over 0.70 and a paired *t*-test showed no significant difference at *p*-values < 0.05 (Table VIII).

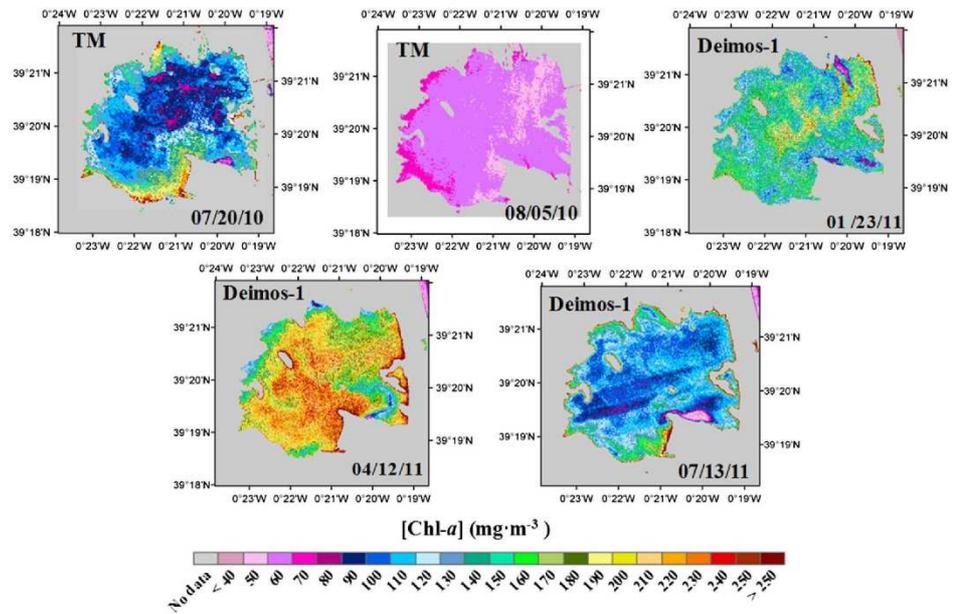


Fig. 4. Temporal evolution of $[Chl-a]$ in the Albufera for the period 2010–2011. Maps obtained from TM and Deimos-1 images.

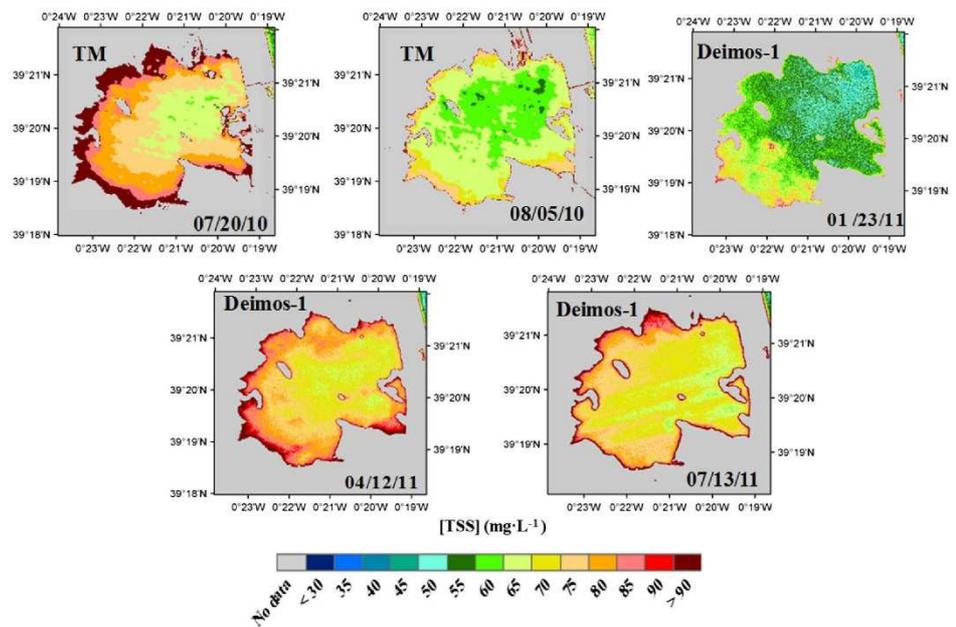


Fig. 5. Evolution of $[TSS]$ in the Albufera for the period 2010–2011. Maps obtained from TM and Deimos-1 images.

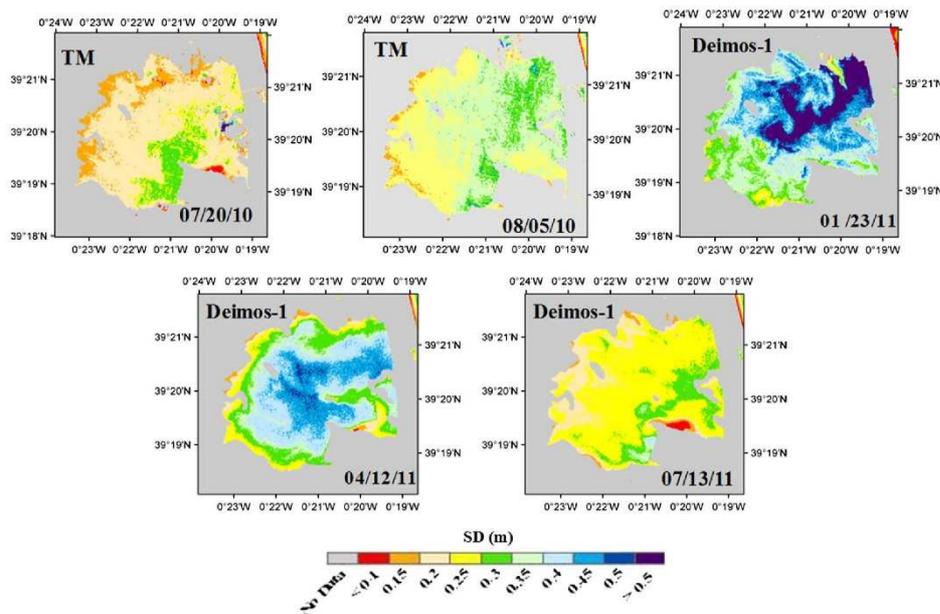


Fig. 6. Evolution of transparency in the Albufera for the period 2010–2011. Maps obtained from TM and Deimos-1 images.

Finally, results of the comparison between predicted and observed data are shown in Table IX. Acceptable results are shown in the case of the [Chl-*a*] and [TSS] estimations, with MADP values of ± 22 and $\pm 15\%$, respectively (Table IX), whereas SD results are quite poor ($R^2 = 0.03$, $p = 0.015$) with the MADP reaching $\pm 40\%$. This is a direct consequence of the uncertainty of the measurement itself, together with the low values of transparency characteristic of the Albufera where Secchi depth remains below 0.5 m. This makes the Albufera a challenging water body to perform SD estimations from satellite. Some previous studies, such as, López-García and Caselles [11] have attempted to do so, but did not find a correlation between TM bands and this parameter for the same reasons given above.

Finally, as a demonstration of the real potential of remote sensing, maps of [Chl-*a*], [TSS], and SD were obtained by applying the previous algorithms to the Landsat-5 and Deimos-1 scenes available for the 2010–2011 period in the Albufera. These images cover a typical annual cycle for this water body. Seasonality observed in these maps is largely a consequence of the gate opening to achieve the water requirements for the rice cultivation in this area. For instance, values of [Chl-*a*] peak in April since gates normally remain closed during this month [35] (Figs. 4–6).

High levels of [TSS] (Fig. 5) are clearly linked to low values of transparency (Fig. 6) and both are a consequence of the high values of [Chl-*a*] (Fig. 4). Also, the horizontal heterogeneity along the Albufera lake is evident (Figs. 4–6).

The high spatial resolution of TM and Deimos-1 allows the detection of weak inflows or outflows of water, which is enhanced by the ability of Deimos-1 to capture images every 3 days. Therefore, these two sensors are a valuable tool to monitor the water quality, to analyze the horizontal heterogeneity, and to perform temporal studies of lakes and reservoirs such as the Albufera de Valencia.

V. CONCLUSION

In this work, empirical relationships were obtained and tested for each one of the water quality variables under study, using reflectance values from TM spectral bands as inputs: the ratio TM2/TM4 for [Chl-*a*], TM4 for [TSS], and TM2 for transparency. These algorithms were validated for the Albufera de Valencia, using additional field campaigns concurrent with Deimos-1 overpasses, and results were acceptable in terms of [Chl-*a*] and [TSS] estimation with MADP values of $\pm 22\%$ and $\pm 15\%$, respectively. The algorithm for the estimation of Chl-*a* showed high sensitivity in its estimation for TM2/TM4 values lower than 1.5. Also, the high correlation between Chl-*a* and TSS allows us to obtain acceptable error values in the estimation of Chl-*a* with TM and Deimos-1 sensors. However, some discrepancies arose when monitoring water transparency as consequence of the low values of this parameter and the windy conditions prevailing in the Albufera, leading to a large relative error in the determination of the SD.

In our case, the use of TM imagery is especially attractive to study the evolution of the water quality of the Albufera over the last three decades. Furthermore, the spectral band similarity with Deimos-1 will provide continuity to the algorithms designed for TM, following its recent failure, improving the spatial resolution and reducing the revisit time to 3 days. Users have free access to the Landsat archive, available online in the USGS, and some Deimos-1 imagery will also be available in the National Remote Sensing Plan [36] of Spain.

Future work will focus on the expansion of the database used for the model construction and the application of these models to a wider variety of lakes and reservoirs, as well as further quantifying the model uncertainty, with the aim of it serving as a tool for the monitoring of the ecological status of water bodies with regard to the objective of the European Water Framework Directive [2].

ACKNOWLEDGMENTS

The authors would like to thank the comments made by Prof. M. A. Gilabert, M. Mira, M. Bisquert (Valencia University), M. Murphy (Imperial College London) and by the anonymous referees since they have contributed to improve the paper.

REFERENCES

- [1] "OECD," *Eutrophication des Eaux. Methodes de surveillance, d'évaluation et de lutte (Eutrophication of water, monitoring, assessment and control)*. Paris, France, 1982.
- [2] European Commission, "Directive 2000/60/EC of the European Parliament, and of the Council of 23 October 2000 establishing a framework for Community action in the field of water policy," *Official J. Eur. Communities*, vol. 327, pp. 1–72, 2000.
- [3] M. Mayo, A. Gitelson, Y. Z. Yacobi, and Z. Ben-Avraham, "Chlorophyll distribution in Lake Kinneret determined from Landsat Thematic Mapper data," *Int. J. Remote Sens.*, vol. 16, no. 1, pp. 175–182, 1995.
- [4] H. Duan, Y. Zhang, B. Zhang, K. Song, and Z. Wang, "Assessment of Chlorophyll-*a* concentration and trophic state for Lake Chagan using Landsat TM and field spectral data," *Environ. Monit. Assess.*, vol. 129, pp. 295–308, Oct. 2006.
- [5] M. G. Allan, D. P. Hamilton, B. J. Hicks, and L. Brabyn, "Landsat remote sensing of chlorophyll-*a* concentrations in central North Island lakes of New Zealand," *Int. J. Remote Sens.*, vol. 32, no. 7, pp. 2037–2055, 2011.
- [6] L. G. Olmanson, M. E. Bauer, and P. L. Brezonik, "A 20-year Landsat water clarity census of Minnesota's 10,000 lakes," *Remote Sens. Environ.*, vol. 112, no. 11, pp. 4086–4097, 2008.
- [7] I. M. McCullough, C. S. Loftin, and S. A. Sader, "Combining lake and watershed characteristics with Landsat TM data for remote estimation of regional lake clarity," *Remote Sens. Environ.*, vol. 123, pp. 109–115, 2012.
- [8] R. J. Allee and J. E. Johnson, "Use of satellite imagery to estimate surface chlorophyll-*a* and Secchi disc depth of Bull Shoals Reservoir, Arkansas, USA," *Int. J. Remote Sens.*, vol. 20, no. 6, pp. 1057–1072, 1999.
- [9] P. A. Brivio, C. Giardino, and E. Zilioli, "Determination of Chlorophyll concentration changes in Lake Garda using an image-based radiative transfer code for Landsat TM images," *Int. J. Remote Sens.*, vol. 22, no. 2–3, pp. 487–502, 2001.
- [10] G. Mancino, A. Nolè, V. Urbano, M. Amato, and A. Ferrara, (2009, Jul. 30). "Assessing water quality by remote sensing in small lakes: The case study of Monticchio lakes in southern Italy," *iForest* vol. 2, pp. 154–161 [Online]. Available: <http://www.sisef.it/forest/show.php?id=507>.
- [11] M. J. López-García and V. Caselles, "Use of Thematic Mapper data to assess water quality in Albufera lagoon of Valencia (Spain)," presented at the 13th Annu. Conf. Remote Sens. Soc., Nottingham, UK., Sep. 7–11, 1987, pp. 510–519.
- [12] J. A. Domínguez, E. Chuvieco, and A. Sastre, "Monitoring transparency in inland water bodies using multispectral images," *Int. J. Remote Sens.*, vol. 30, no. 6, pp. 1567–1586, Mar. 2009.
- [13] I. M. McCullough, C. S. Loftin, and S. A. Sader, "High-frequency remote monitoring of large lakes with MODIS 500 m imagery," *Remote Sens. Environ.*, vol. 124, pp. 234–241, 2012.
- [14] R. Peña, A. Ruiz, and J. A. Domínguez, "Mapping of photosynthetic pigments in Spanish inland waters using MERIS imagery," in *Proc. Envisat Symp.*, Salzburg, Austria, Sep. 6–10, 2004 [Online]. Available: <http://hercules.cedex.es/ecosistemas/TeleCongresos/186penamartinez.pdf>.
- [15] P. Härmä, J. Vepsäläinen, T. Hannonen, T. Pyhäläthi, J. Kämäri, K. Kallio, K. Floheimo, and S. Koponen, "Detection of water quality using simulated satellite data and semi-empirical algorithms in Finland," *Sci. Tot. Environ.*, vol. 268, pp. 107–121, 2001.
- [16] L. G. Olmanson, P. L. Brezonik, and M. E. Bauer, "Evaluation of medium to low resolution satellite imagery for regional lake water quality assessments," *Water Resour. Res.*, vol. 47, p. W09515, 2011.
- [17] DMC International Imaging [Online]. Available: http://www.dmcii.com/?page_id=6770.
- [18] E. Vicente and M. R. Miracle, "The coastal Lagoon Albufera de Valencia: An ecosystem under stress," *Limnetica*, vol. 8, pp. 87–100, 1992.
- [19] J. M. Soria and E. Vicente, "Estudios de los aportes hidricos al parque natural de la Albufera de Valencia, (Studies of water flow contribution to the natural park of la Albufera de Valencia)," *Limnetica*, vol. 2, pp. 105–115, 2002.
- [20] J. D. H. Strickland and T. P. Parsons, *A Practical Handbook for Seawater Analysis*. Ottawa, Canada: Fisheries Research Board, 1968.
- [21] APHA/AWWA/WEF, *Standard Methods for Examination of Water and Wastewater*, 19th ed. Washington, DC, USA: Amer. Public Health Assoc., 1992.
- [22] R. Margalef, *Limnology* [in Spanish], Barcelona, Spain: Omega, 1983.
- [23] G. Chander, B. L. Markham, and D. L. Helder, "Summary of current radiometric calibration coefficients for Landsat MSS, TM, ETM+, and EO-1 ALI sensors," *Remote Sens. Environ.*, vol. 113, pp. 893–903, 2009.
- [24] J. A. Domínguez, C. Alonso, and A. Alonso, "Remote sensing as a tool for monitoring water quality parameters for Mediterranean Lakes of European Union water framework directive (WFD) and as a system of surveillance of cyanobacterial harmful algae blooms (SCYanoHABs)," *Environ. Monit. Assess.*, vol. 181, no. 1–4, pp. 317–334, 2011.
- [25] P. S. Chavez Jr., "Image-based atmospheric corrections-revisited and improved," *Photogramm. Eng. Remote Sens.*, vol. 62, no. 9, pp. 1025–1036, 1996.
- [26] M. A. Gilabert, C. Conese, and F. Maselli, "An atmospheric correction method for the automatic retrieval of surface reflectances from TM images," *Int. J. Remote Sens.*, vol. 15, no. 10, pp. 2065–2086, 1994.
- [27] M. Neubert and G. Meinel, "Atmospheric and terrain correction of IKONOS imagery using ATCOR3," in *Proc. ISPRS Hanover Workshop*, Germany, 2005.
- [28] A. Berk, G. P. Anderson, P. K. Acharya, and E. P. Shettle, "MODTRAN 5 User's Manual," Hascos AFB, MA: Air Force Res. Lab., Space Veh. Dir., Air Force Mater. Comm., 2008.
- [29] D. G. Hadjimitsis, C. R. I. Clayton, and V. S. Hope, "An assessment of the effectiveness of atmospheric correction algorithms through the remote sensing of some reservoirs," *Int. J. Remote Sens.*, vol. 25, no. 18, pp. 3651–3674, 2004.
- [30] C. J. Willmott, "Some comments on the evaluation of model performance," *Bull. Amer. Met. Soc.*, vol. 11, pp. 1309–1313, 1982.
- [31] R. Ma and J. Dai, "Investigation of chlorophyll-*a* and total suspended matter concentrations using Landsat ETM and field spectral measurement in Taihu Lake, China," *Int. J. Remote Sens.*, vol. 26, no. 13, pp. 2779–2787, 2005.
- [32] G. Wu, J. De Leeuw, A. K. Skidmore, H. H. T. Prins, and Y. Liu, "Comparison of MODIS and Landsat TM5 images for mapping tempo-spatial dynamics of Secchi disk depths in Poyang Lake National Nature Reserve, China," *Int. J. Remote Sens.*, vol. 29, pp. 2183–2198, 2008.
- [33] A. Cózar, C. M. García, J. A. Gálvez, S. A. Loiselle, L. Bracchini, and A. Cognetta, "Remote sensing imagery analysis of the lacustrine system of Ibera wetland (Argentina)," *Ecol. Model.*, vol. 186, no. 1, pp. 29–41, 2005.
- [34] P. Lavery, C. Patiaratchi, A. Wyllie, and P. Hick, "Water quality monitoring in estuarine waters using the landsat thematic mapper," *Remote Sens. Environ.*, vol. 46, pp. 268–280, 1993.
- [35] S. Romo, A. García-Murcia, M. J. Villena, V. Sánchez, and A. Ballester, "Phytoplankton trends in the lake of Albufera de Valencia and implications for its ecology, management, and recovery," *Limnetica*, vol. 27, no. 1, pp. 11–28, 2008.
- [36] *National Plan of Remote Sensing (Spain)* [Online]. Available: <http://www.ign.es/PNT/>.



Carolina Doña received the B.Sc. degree in physics and the M.Sc. degree in remote sensing from the University of Valencia, Valencia, Spain, in 2008 and 2012, respectively. Currently, she is working toward the Ph.D. degree in remote sensing with the Thermal Remote Sensing Group, Department of Earth Physics and Thermodynamics, Faculty of Physics, University of Valencia.

His research interest focused in optical remote sensing applied to water quality and environmental management.



Juan M. Sánchez received the B.Sc., M.Sc., and Ph.D. degrees in physics from the University of Valencia, Valencia, Spain, in 2003, 2005, and 2008, respectively.

Currently, he is a Lecturer with the Applied Physics Department, University of Castilla-La Mancha, Almadén, Spain. His research interest focuses on the thermal infrared remote sensing in general and the surface-energy-flux retrieval in particular. He has published 23 papers in international journals, 5 book chapters and more than 50 conference papers. He has

participated in 18 national and international projects, and he is referee in 11 international journals.

Dr. Sánchez was awarded the Nobert Gerbier-MUMM International Award 2010 by the World Meteorological Organization.



Vicente Caselles received the B.Sc., M.Sc., and Ph.D. degrees in physics from the University of Valencia, Valencia, Spain, in 1979, 1980, and 1983, respectively.

Currently, he is a Professor in earth physics and the Head of the Thermal Remote Sensing Group, Department of Earth Physics and Thermodynamics, Faculty of Physics, University of Valencia. He has 35 years expertise in the physical processes involved in both temperature measurement and evapotranspiration using remote sensing techniques, which has been documented through 10 books, 20 doctoral theses, 100 papers in international journals, 60 conference papers, and 30 reports. He was collaborating with the European Space Agency as a Member of the Advisory Group for the Land-Surface Processes and Interactions Mission. He was the Chairman of the Spanish Remote Sensing Society and is currently the Manager of Human Resources and Researchers Mobility General Direction at the Spanish Ministry of Economy and Competitiveness.

Dr. Caselles received the Norbert Gerbier-MUMM International Award for 2010, conferred by the Executive Council of the World Meteorological Organization.



Jose Antonio Domínguez received the B.Sc. and M.Sc. degrees in physics from the University of Valladolid, Valladolid, Spain, in 1992 and 1994, respectively; and the M.Sc. degree in remote sensing, GIS, and cartography; and the Ph.D. degree in environmental sciences from the University of Alcalá, Madrid, Spain, in 1998 and 2003, respectively.

Currently, he is an Honorary Lecturer at Spanish Open University, Madrid, Spain. From 1999 to 2013, he is a Researcher at the Remote Sensing and Applied Hydrology and Limnology Area (Ecology of Continental Aquatic Systems), Center for Hydrographic Studies, CEDEX. He has 15 years expertise in remote sensing applied to water quality and environmental management. He has published more than 50 papers and presentations on remote sensing and/or aquatic media both in national and international journals and conferences and a book on water quality using remote sensing, published by the Spanish Open University, in 2011. He worked as PI in several ESA projects about the study of freshwater using remote sensing, as a Member of the Scientific Committee for Determination of water requirements in Wetlands, and as a Coordinator of the Thematic Group Inland Water at the Spanish Remote Sensing Society.



Antonio Camacho was born in 1965 in Spain. He received the B.Sc. degree in biology and the Ph.D. degree in ecology and microbial ecology from the University of Valencia, Valencia, Spain, in 1987 and 1997, respectively.

He is a Head of Research at the Aquatic Ecology Group of the Cavanilles Institute for Biodiversity and Evolutionary Biology and Associate Professor with the Department of Microbiology and Ecology, the University of Valencia, Valencia, Spain. Currently, his main research and university teaching concerns are

associated with basic and applied aspects of aquatic ecology. Basic research aspects are related to the functional ecology of lakes and wetlands, microbial ecology of such ecosystems, and the effect of climate change on Mediterranean and Antarctic lakes and wetlands. Applied aspects of his research mainly concern the development of methodologies for the evaluation of ecological status of lotic ecosystems and the conservation and restoration of aquatic ecosystems and water quality. He has been a Visiting Scientist at some Universities and Research Centres mainly from the USA, France, The Netherlands, and Austria. He led five Antarctic expeditions in which multinational teams with researchers of 15 different countries made joint research on the functioning of terrestrial freshwater Antarctic ecosystems. He is a co-author of more than 100 research papers in peer-review scientific journals and books, as well as of near 200 papers presented in scientific meetings. Currently, he is a President of the Iberian Limnological Association, which associates near 500 limnologists from Spain and Portugal.

Dr. Camacho obtained the award for the best Ph.D. thesis on aquatic microbiology for the biennium 1996–1997 from the Spanish Society for Microbiology. In 2004, he also got the Award on Environmental Innovation from Foundation 3M.

Anexo B

Artículo “Integrated satellite data fusion and mining for monitoring lake water quality status of the Albufera de Valencia in Spain”

Este artículo se realizó en colaboración con los profesores N. B. Chang y J. Delegido y se publicó en Journal of Environmental Management en enero de 2015. En este trabajo se mejoró la resolución temporal de las imágenes Landsat realizando la fusión de imágenes de satélite con imágenes MODIS, así se pudieron generar imágenes sintéticas de Landsat para fechas en las que no pasaba sobre el área de estudio. Por último, se desarrollaron algoritmos genéticos para realizar la estimación de parámetros de calidad de agua tales como la concentración de clorofila-*a* y transparencia de agua en la Albufera de Valencia, utilizando como entrada los datos de reflectividad de superficie. Esta metodología permitió mejorar el esfuerzo tradicional por medio de técnicas de regresión lineal, como se mostraba en el trabajo anterior.

La revista Journal of Environmental Management poseía un factor de impacto de **3,13** y ocupaba la posición relativa **54/225**, dentro de la categoría de Environmental Sciences de la base de datos ISI Web of Knowledge, en el año de publicación.



Integrated satellite data fusion and mining for monitoring lake water quality status of the Albufera de Valencia in Spain



Carolina Doña^a, Ni-Bin Chang^{b,*}, Vicente Caselles^a, Juan M. Sánchez^c, Antonio Camacho^d, Jesús Delegido^e, Benjamin W. Vannah^b

^a Department of Earth Physics and Thermodynamics, University of Valencia, Burjassot, Valencia, Spain

^b Department of Civil, Environmental, and Construction Engineering, University of Central Florida, Orlando, FL, USA

^c Department of Applied Physics, University of Castilla-La Mancha, Almadén, Ciudad Real, Spain

^d Cavanilles Institute of Biodiversity and Evolutionary Biology, University of Valencia, Burjassot, Valencia, Spain

^e Image Processing Laboratory, University of Valencia, Paterna, Valencia, Spain

ARTICLE INFO

Article history:

Received 5 June 2014

Received in revised form

24 November 2014

Accepted 2 December 2014

Available online 17 January 2015

Keywords:

Water quality

Lake management

Remote sensing

Data fusion

Data mining

Machine learning

ABSTRACT

Lake eutrophication is a critical issue in the interplay of water supply, environmental management, and ecosystem conservation. Integrated sensing, monitoring, and modeling for a holistic lake water quality assessment with respect to multiple constituents is in acute need. The aim of this paper is to develop an integrated algorithm for data fusion and mining of satellite remote sensing images to generate daily estimates of some water quality parameters of interest, such as chlorophyll *a* concentrations and water transparency, to be applied for the assessment of the hypertrophic Albufera de Valencia. The Albufera de Valencia is the largest freshwater lake in Spain, which can often present values of chlorophyll *a* concentration over 200 mg m⁻³ and values of transparency (Secchi Disk, SD) as low as 20 cm. Remote sensing data from Moderate Resolution Imaging Spectroradiometer (MODIS) and Landsat Thematic Mapper (TM) and Enhance Thematic Mapper (ETM+) images were fused to carry out an integrative near-real time water quality assessment on a daily basis. Landsat images are useful to study the spatial variability of the water quality parameters, due to its spatial resolution of 30 m, in comparison to the low spatial resolution (250/500 m) of MODIS. While Landsat offers a high spatial resolution, the low temporal resolution of 16 days is a significant drawback to achieve a near real-time monitoring system. This gap may be bridged by using MODIS images that have a high temporal resolution of 1 day, in spite of its low spatial resolution. Synthetic Landsat images were fused for dates with no Landsat overpass over the study area. Finally, with a suite of ground truth data, a few genetic programming (GP) models were derived to estimate the water quality using the fused surface reflectance data as inputs. The GP model for chlorophyll *a* estimation yielded a R^2 of 0.94, with a Root Mean Square Error (RMSE) = 8 mg m⁻³, and the GP model for water transparency estimation using Secchi disk showed a R^2 of 0.89, with an RMSE = 4 cm. With this effort, the spatiotemporal variations of water transparency and chlorophyll *a* concentrations may be assessed simultaneously on a daily basis throughout the lake for environmental management.

© 2014 Elsevier Ltd. All rights reserved.

1. Introduction

Eutrophication in lakes and coastal waters has long been a common environmental problem in many countries around the world. Eutrophication mostly results from an anthropogenic supply of nutrients, mainly nitrogen and phosphorus, from urban and agricultural runoff (Bricker et al., 1999). These nutrients enhance

algal growth, thereby reducing water transparency and deteriorating water quality (Bricker et al., 1999). Remote sensing reflectance bands can be used in aquatic environments to detect these pollution impacts. Chlorophyll *a* concentrations (Chl-*a* hereafter), an indicator of phytoplankton biomass that exhibits a positive correlation with a nutrient increase, are normally determined from bio-optical observations in relatively open waters (O'Reilly et al., 1998).

Remotely sensed Chl-*a* is particularly challenging in turbid, highly productive waters (Siegel et al., 2000; Gitelson et al., 2007,

* Corresponding author. Tel.: +1 407 754 7521.

E-mail address: nchang@ucf.edu (N.-B. Chang).

2008), such as the coastal lake Albufera de Valencia (Spain). In many turbid water bodies, total suspended solids (TSS), either organic (mainly phytoplankton) or inorganic (sediment particles), determine water column optics. Whereas the latter can be depicted by water transparency, the former can be reflected by Chl-*a*. Monitoring water transparency informs the eutrophic status of the ecosystem because the growth of phytoplankton decreases light penetration. Paired monitoring of water transparency and Chl-*a* can thereby reveal the general condition of the trophic status in a lake when the lack of transparency is mostly due to phytoplankton. With these indicators, water bodies can be identified through the oligotrophic to the hypereutrophic range with accuracy. The Organization for Economic Co-operation and Development (OECD, 1982) proposed different ranges of the trophic classification determined in terms of total phosphorus (TP), Chl-*a* and water transparency measured with Secchi disk (SD), which have been widely accepted.

Encouraged by the enforcement of the European Water Framework Directive (WFD) in 2000, which establishes that all EU countries have to monitor the ecological status of their surface water bodies in order to achieve the good ecological status, there is a growing interest in using remote sensing techniques to monitor water quality variables. Remote sensing may allow a highly consistent monitoring to come with essential spatial coverage and temporal resolution when compared to the sporadic and intermittent field sampling work. Lake Albufera de Valencia, located in eastern Spain, is a hypertrophic water body that presents average Chl-*a* concentrations of up to 100 mg m⁻³, and peaks are higher than 200 mg m⁻³. Therefore it requires intensive monitoring to retrieve its water quality conditions, leading to the generation of specific measures for environmental management (Vicente and Miracle, 1992). Right after the Albufera de Valencia and its surroundings were classified as a Natural Park in 1985. López-García and Caselles (1987) carried out the first remote sensing study in this water body. The main goal of this study was to monitor the Chl-*a* concentration, seston concentration, and water transparency with the aid of Thematic Mapper (TM) sensor onboard Landsat-5 satellite (López-García and Caselles, 1987). To retrieve the general water quality information, they applied regression techniques for estimating each parameter.

Since then, many water quality monitoring studies have focused on mapping different water quality parameters using remote sensing technologies, with varying purposes (e.g. Serrano et al., 1997; Härmä et al., 2001; Peña et al., 2004; Duan et al., 2006; Allan et al., 2011; McCullough et al., 2012). Landsat imageries were extensively used in those studies because of its relatively high spatial resolution of 30 m and its long period of archived data, which allows one to conduct multitemporal change detection of

targeted water bodies (e.g., Mayo et al., 1995; Allee and Johnson, 1999; Olmanson et al., 2008). Yet, the sensor onboard Landsat has a limited spectral resolution and this, together with the low temporal resolution (16 days), makes it difficult to determine some of the water quality parameters accurately and timely. Sensors such as MODIS and MERIS, that have higher temporal-spectral resolution, are more suitable to monitor different water quality parameters with time-sensitive reflective spectral signatures (Peña et al., 2004; McCullough et al., 2012); but the spatial resolution of their images is too coarse to analyze medium or small size water bodies (Mancino et al., 2009).

To study highly eutrofied lakes, such as the Albufera de Valencia, the development of the integrated data fusion and mining (IDFM) algorithm with the involvement of both Landsat and MODIS imageries might be a promising tool to get through the hurdle of near real-time monitoring (Chang et al., 2014a,b). A comparison between the main features of both sensors (i.e., Landsat and MODIS) gives rise to some insight about the possibility of sensor fusion, data fusion, or even information fusion (Table 1). To fuse the images collected by MODIS and Landsat, the data fusion algorithms working at the pixel level, such as the Temporal Adaptive Reflectance Fusion Model (STAR-FM) algorithm, may be adopted to enhance spatial, spectral, and temporal properties (Gao et al., 2006). Nevertheless, bio-optical or empirical algorithms, such as statistical regression and computational intelligence algorithms, are required to help classify and interpret fused remote sensing data linking the reflectance over radiance bands of the sensors with ground truth data (Peña et al., 2004; Allan et al., 2011; Domínguez et al., 2011; McCullough et al., 2012; Alonso-Fernández et al., 2013; Chang et al., 2014a,b).

The aim of this paper is to apply the IDFM algorithm developed by Chang et al. (2014a,b) to estimate and assess the dynamics of Chl-*a* concentration and water transparency at the Albufera de Valencia, addressing the water quality status under the agricultural runoff impact. With such understanding of the changing water quality status, water management policy may be examined for sustainable development. In this context, fused STAR-FM Landsat–MODIS images were analyzed by a suite of genetic programming (GP) models, in order to retrieve the water quality status in the Albufera de Valencia. This endeavor leads us to explore the following science questions: 1) Can data fusion techniques be used to fill in the temporal data gaps left by Landsat and capture the spectral features of water transparency and Chl-*a* concentrations simultaneously during the study period? 2) Can we perform data mining via GP models to effectively retrieve water transparency and Chl-*a* concentrations based on the fused images at the same time? We hypothesized that the IDFM technique can estimate multiple water quality parameters at the same time, improving the spatial and temporal resolution of the water quality maps for holistic, near real-time environmental monitoring and assessment.

Table 1
Comparison between the main features of sensors MODIS and Landsat.

Sensor	Fused bands	Spectral resolution (nm)	Temporal resolution (days)	Spatial resolution (m)
Landsat TM/ETM+	1	450–520	16	30
	2	520–600		
	3	630–690		
	4	760–900		
	5	1550–1750		
	7	2080–2350		
	MODIS TERRA	3		
4		545–565		
1		620–670		
2		841–876		
6		1628–1652		
7		2105–2155		

2. Materials & methods

2.1. Study area

Albufera de Valencia is a shallow lake located in the Mediterranean coast of Valencia, Spain (39° 20'N, 0° 20'W) (Fig. 1). It is the largest natural water body on the Iberian Peninsula, and its surface is 23.2 km² with an average depth of 1.2 m (Romo et al., 2008). In 1985, the lake and its surroundings were recognized as a Natural Park. In 1989, the lake was included in the Ramsar international list of protected wetlands. After that, the lake has been a popular site for numerous limnological studies (Vicente and Miracle, 1992).

The lake is connected to the Mediterranean Sea by two natural and one artificial channels called “golas”, whose management is

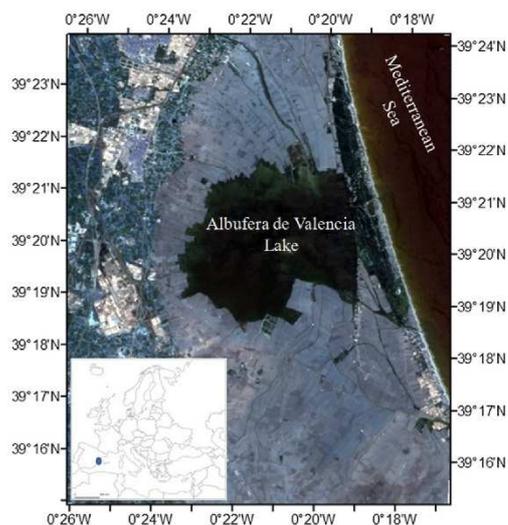


Fig. 1. Study area. Landsat true color image (RGB (bands 3, 2 and 1)).

linked to the different rice growing seasons since water level in the surrounding rice fields is regulated through the lake level. Rice is the dominant crop in the area and the hydrologic cycle of the Albufera is intimately linked with the irrigation schedule and agricultural runoff. Therefore water management policy depends largely on these operations. The channel gates are open in January–March to allow the lake water level to rise mostly during this period for irrigation. During the rice growing season (April–September) the gates remain closed, with a negligible water flux into the lake. The gates open again in September–October to allow rice fields to become dry for rice harvesting. Finally, the gates are closed again in November–December to allow the flooding of the harvested rice fields to happen, favoring nutrient mineralization (Romo et al., 2008). Water inflows to the Albufera come mainly from irrigation channels bringing waters from the Turia and Júcar rivers, and several ravines which carry rain water. It also receives some treated wastewater from the urban and industrial areas nearby.

Even today, despite the improvements in the sanitary sewer systems, small amounts of untreated wastewater are still discharged into this water body. Due to these collective impacts primarily associated with changes in land use patterns, the water body of the Albufera has been polluted into a hypertrophic status with high values of Chl-*a* exceeding 100 mg m^{-3} , and low values of water transparency with Secchi Disk depths (SDDs) commonly below 20 cm (Soria and Vicente, 2002). The high trophic level is nowadays maintained not only by the external nutrient inputs, but also by the strong internal nutrient load accumulated in the lake sediments. The different sources of nutrient inputs and the differences in the sediment nutrient deposits throughout the lake, together with the effect of wind and currents, determine a certain spatial heterogeneity in the distribution of water quality variables such as the Chl-*a* concentration and the water transparency. This variability can hardly be monitored with enough spatial-temporal resolution via regular limnological monitoring techniques, which justifies the use of remote sensing techniques.

2.2. Ground-based data

Several field campaigns were carried out in the Albufera de Valencia during 2006 and 2011. A total of 6 sampling efforts were used in this study, including 5 sampling efforts in 2006 and one in 2011. Ground-based data were collected to reflect Chl-*a* concentration and water transparency (Tables 2 and 3). Although more sampling records were available later on in 2011, the corresponding MODIS images contained some clouds in the nearby area such that we focused only on these cloud-free days for remote sensing analysis.

In these 2006 field campaigns, Chl-*a* concentrations were measured by in-situ fluorometry with later re-calibration with laboratory data (Domínguez et al., 2011). In the 2011 field campaigns, however, samples were collected and then analyzed in the laboratory using spectrophotometric methods following the method proposed by Strickland and Parson (1968) and adapted as described in Picazo et al. (2013). Both methods fluorometry and spectrophotometry show a good agreement. Yet an underestimation of the fluorimeter measurements were observed, which could be due to the presence of phycocyanin in the water body, whose light absorption does not result in equal molar ranges of Chl-*a* fluorescence. Additional factors can also affect this measurement such as the degradation byproducts of Chl-*a*. Nevertheless, both methods are comparable if the fluorimeter can be re-calibrated with Chl-*a* data estimated in the laboratory (Pinto et al., 2001; Lawrenz and Richardson, 2011). The water transparency was measured as the SDD. This method has been shown to be acceptable for measuring water transparency (Margalef, 1983) and can also be used to determine the trophic level of inland water bodies (OECD, 1982).

In the field campaigns, the Chl-*a* concentration measurements (Table 2), ranged from a minimum value of 30 mg m^{-3} , in July 2006, to a maximum of 240 mg m^{-3} in April 2011. These values are far higher than the water quality standard established by the WFD for this type of water bodies, which evaluates the achievement of good ecological status. In the case of the water transparency (Table 3), collected values of SDD did not exceed 47 cm. Therefore, the Albufera de Valencia is deemed as a hypertrophic lake requiring intensive water quality monitoring as required by the WFD to achieve the environmental goals.

2.3. Remote sensing data collection

Landsat TM and ETM+ as well as MODIS, onboard Landsat 5, Landsat 7 and Terra platforms, respectively, were the sensors used to conduct this study. All the images are available, free of charge, in the webpage of the United States Geological Survey. Landsat images were downloaded from the Glovis visor (<http://glovis.usgs.gov/>) and the MODIS images from the web interface of the MODIS reprojection tool (<https://mrtweb.cr.usgs.gov/>). With this tool the

Table 2

Average, maximum, minimum values, standard deviation and number of sample points, for Chl-*a* concentration registered in the Albufera de Valencia during the sampling campaigns carried out in 2006 and 2011.

	Chl- <i>a</i> (mg m^{-3})				<i>n</i>
	Avg	Max	Min	Std dev	
06/22/2006	140	210	60	110	2
06/24/2006	80	100	50	20	4
07/13/2006	40	60	30	10	16
08/11/2006	55	66	44	9	7
09/26/2006	130	158	104	17	10
04/12/2011	200	240	150	30	15

Table 3
Average, maximum, minimum values, standard deviation and total number of samples for water transparency registered in the Albufera de Valencia during the sampling campaigns carried out in 2006 and 2011.

	Water transparency (cm)				n
	Avg	Max	Min	Std dev	
06/22/2006	36	41	29	6	6
06/24/2006	39	44	36	4	3
07/13/2006	40	47	33	4	15
08/11/2006	39	44	34	3	7
09/26/2006	24	28	20	3	10
04/12/2011	14	20	10	3	15

MODIS Level-2G, Level-3, and Level-4 land data products are referenced to a global tiling scheme, and tiles are non-overlapping (Dwyer and Schmidt, 2006), i.e. these images does not present the bow tie effect. This means that the scans are partially overlapping at off nadir angles and this effect occurs at the border of each scene. This effect is presented in MODIS level 1 and 2 products. These products can be corrected using the MODIS Reprojection Tool Swat, but not level 2G products. Even though the bow tie effect is corrected, sometimes these images present distortion in the pixel values. We checked every single image to screen out this kind of issue. Finally, we used a total of 28 images (19 from MODIS and 9 from Landsat). The downloaded images correspond to the synchronous ground data sampling campaigns and the extra images necessary to carry out the data fusion processing (Table 4).

2.4. Remote sensing image processing

The steps needed to carry out image processing for the MODIS and Landsat images individually before performing the data fusion are shown in Fig. 2. The selected MODIS product was the MOD-09GA/GQ daily reflectance data with 500/250 m resolution, while the Landsat images were the radiometrically and geometrically terrain corrected product – L1T, with 30 m resolution. This type of Landsat data is depicted in digital numbers and need to be transformed to surface reflectance values. The atmospheric layer between the surface and the satellite sensors disturbs the signal and an atmospheric correction of the Landsat images was necessary to remove this atmospheric disturbance. In this work, the 6S radiative transfer code was used. It is an improved version of 5S and was developed according to the radiative transfer theory (Vermote et al., 1997; Masek et al., 2006). In addition, the MODIS images were resampled to a pixel size of 30 m, applying the nearest neighbor resample method. Both types of images were resampled to UTM zone 31N, and then a water mask was applied to separate the water-land pixels. Finally, Landsat and MODIS images were obtained with the same pixel size (30 m) for the subsequent data fusion over the entire study area. After finishing the image processing, all the images were visually compared to check if they were co-registered correctly.

2.5. Data fusion

The application of data fusion for remotely sensed images involves the algorithmic combination of spectral, temporal, or spatial characteristics from two or more images (Genderen and Pohl, 1994; Caselles et al., 1998). The fused or synthetic image possesses the best characteristics of their parent images to provide higher monitoring power. Given that data fusion can focus on one or more properties (spectral, temporal, spatial, etc.) offered by the parent images, an obvious motivation for data fusion comes from the fact that the fused images may incorporate more information than the

original images (Pohl and Van Genderen, 1998). There are a number of methods for fusing images. One way to differentiate between data fusion techniques is to classify methods according to when data fusion is performed. According to Pohl and Van Genderen (1998), the three primary categories of data fusion by processing level are: pixel level, feature level, and decision level.

STAR-FM, which was used in our analysis, is the fusion software at the pixel level produced by the National Aeronautics and Space Administration of the United States (Gao et al., 2006). The software can be downloaded from the Landsat Ecosystem Disturbance Adaptive Processing System web (LEDAPS) (<http://ledaps.nascom.nasa.gov/tools/tools.html>). It was designed for the fusion of satellite images such as Landsat and MODIS. The methodology incorporates fusion at the pixel level between images that are spectrally similar and spatially correlated. Fusion at the pixel level is primarily done by the combination of physical attributes characterized by the pixel, such as spatial or spectral information. Since spectral similarities are required, the techniques used can be applied to any sensors sharing overlapping, spectrally similar bands, such as MODIS ocean color and MERIS (Chang et al., 2014b). For fusion between Landsat and MODIS images, at least a Landsat–MODIS input pair for the same day are required, followed by a MODIS image corresponding to the next day for which the fused Landsat images can be generated via data fusion. Although STAR-FM algorithm can work based on a single Landsat–MODIS input pair (pre or post conditions, depending on whether the pair image date is from before or after the predicted image date), we used two input pairs (pre and post conditions) to obtain better daily images (Gao et al., 2006). Thus, a suite of images are required to complete the data fusion: a pair (Landsat–MODIS) to build a fusion horizon in every 16 days apart and a series of daily MODIS images to fit in the gap of 14 days in the middle. The predicted images to fit in the gap of 14 days in the middle are called fused or synthetic Landsat image.

It should be noted that a primary limitation of STAR-FM is cloud-cover. If clouds obstruct the area, then the reflectance of the ground cannot be observed. Furthermore, applying this methodology to water may not be feasible under highly dynamic conditions, such as swift currents, rapid algal bloom growth, and highly windy conditions. The effect of these conditions is exacerbated when selecting a MODIS and Landsat input pair that is taken hours apart on the same day. These are, however, not common conditions for the studied lake.

2.6. Water quality mapping

Genetic programming (GP) developed based on genetic algorithms is a subclass of evolutionary computation techniques designed to search for the best fit to perform a user-defined task. GP can decode system behaviors based on empirical data for symbolic regression, uncover relationships, and make inferences using association, path analysis, classification, clustering, and forecasting (Seifert, 2004). One principal advantage of GP is that the solution methodology aims to learn the relationship between the inputs and outputs without any prior knowledge or preconceptions. Thus, the burden of the discovery process is primarily handled by GP, with reduced contribution from preprocessing of data such as random grouping of data for model calibration and validation by the user. In this study, the user-defined task is to develop a GP model or program that uses the inputs of fused surface reflectance data associated with common bands to predict the outputs, which are Chl-a concentration and water transparency.

The fundamental concepts for GP are rooted in evolutionary computing, which relies on Darwinian principles of survival to have the fittest programs reproduce, while subject to genetic operations including crossover, mutation, and reproduction. In GP, a program

Table 4
MODIS and Landsat images used in this work for the days of years (DOY) 2006 and 2011.

DOY 2006	173	175 ^a	194 ^a	198	221	223 ^a	237	239	240	242	251	259	260	262	269 ^a	278	DOY 2011	99	102 ^a	116
MODIS	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	MODIS	X	X	X
LANDSAT	X			X	X		X							X	X	X	LANDSAT	X		X

^a Days with ground-based data.

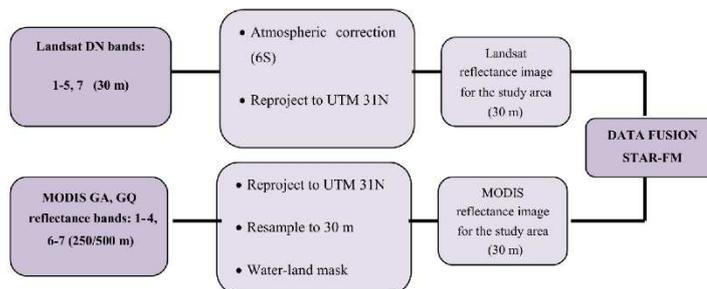


Fig. 2. Steps needed to carry out the image processing of the MODIS and Landsat images.

is a structured set of functions that are ordered as a hierarchical tree (Koza, 1992). The fittest program is a set of instructions comprising highly nonlinear functions that carry out the objective to achieve the goodness-of-fit in an optimal and accurate manner. The end result yields a well-fitted nonlinear equation emergent from natural selection processes and genetic operations selected from a suite of fittest programs.

The ground truth data set was divided into two subsets for GP model training (2/3 of total data) and validation (1/3 of total data), each of which is a supervised learning process leading to the determination of the best thirty nonlinear equations based on the best fit over the training and validation data subsets. Mapping of the Chl-*a* concentrations and the water transparency with their respective reflectance values of each of fused band may be fulfilled in two independent GP analyses.

Discipulus[®] (Francone, 1998) was the software used in this study to perform GP. The program operates using the above stated

procedures; however, programs are selected for genetic operations using tournament selection, which selects a program of highest fitness based upon two or more randomly chosen programs from the population. The resulting white-box GP model may allow one to draw inferences upon the mathematical processes relating the inputs and outputs, but such interpretations are often difficult to make for complex models, which can be further muddled with code bloat.

2.7. Statistical evaluation

Statistical analyses for the comparison between predictions and field observations were conducted using four indices as described in Willmott (1982), as shown below (Eqs. (1)–(4)). Low values of these statistical indices imply better performance in the estimation of these variables.

1. the root of mean square error (RMSE), in which (P_i) and (O_i) are the predicted and the observed values, respectively, and n is the total number of observations:

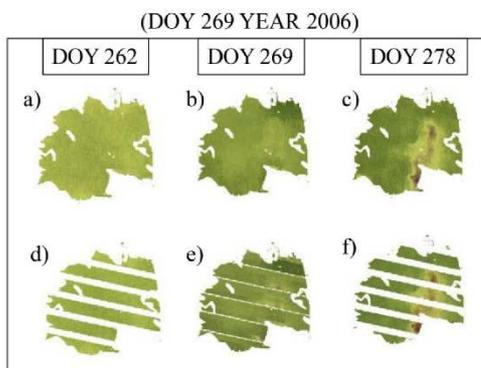


Fig. 3. Original ETM+ images for the DOY 262 and 278 with (a, c) and without (d, f) pretreatment and, the corresponding synthetic images obtained for the DOY 269 with (b) and without (e) pretreatment.

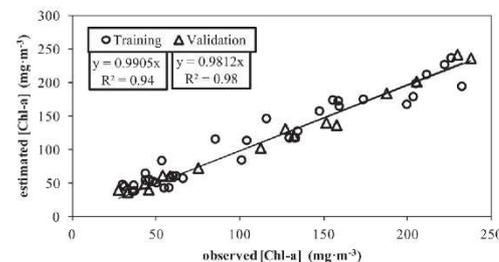


Fig. 4. Estimated versus observed Chl-*a* concentrations based on the derived GP model with respect to the training and validation data sets.

Table 5
Summary of the statistical assessment between predicted and observed Chl-*a* and water transparency [SDD] using the validation data set.

Statistic	[Chl- <i>a</i>] (mg m ⁻³)	SDD (cm)
Observed values	120	30
Std Dev _{obs}	70	10
Predicted values	120	34
Std Dev _{pred}	70	9
BIAS	-2	-0.003
MAD	7	3
RMSE	8	4
MADP (%)	6	9

Note: Std Dev_{obs} and Std Dev_{pred} stand for the standard deviation of observed and predicted values, respectively.

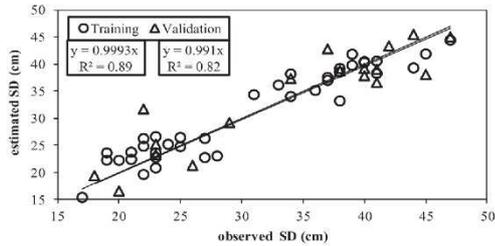


Fig. 5. Estimated versus observed water transparency (Secchi Disk depths) based on the proposed GP model with respect to the training and validation data sets.

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (P_i - O_i)^2}{n}} \quad (1)$$

2. the mean absolute difference (MAD)

$$MAD = \frac{\sum_{i=1}^n |P_i - O_i|}{n} \quad (2)$$

3. the mean absolute deviation percent (MADP)

$$MADP = \frac{\sum_{i=1}^n |P_i - O_i|}{\sum_{i=1}^n O_i} \quad (3)$$

4. the biased estimator (BIAS):

$$BIAS = \frac{\sum_{i=1}^n (P_i - O_i)}{n} \quad (4)$$

3. Results and discussion

3.1. Data fusion

For the purpose of demonstration, we show the fused image in September 2006 (*days of year* – DOY – 269) in Fig. 3. Note that Landsat 7 images have scan line errors resulting in strips which require pretreatment before fusion. To solve the ETM+ gap problem, we applied a bilinear interpolation method for pretreatment (ESRI, 2013). For the purpose of comparison, the pre-condition date is the DOY 262 with and without pretreatment (Fig. 3a, d), respectively, and the post-condition date is the DOY 278 with and without pretreatment (Fig. 3c, f), respectively. Final synthetic images with and without pretreatment are shown in Fig. 3b, e. All this is indicative that the fused results were heavily affected by the problem with the strips in the Landsat ETM+ images.

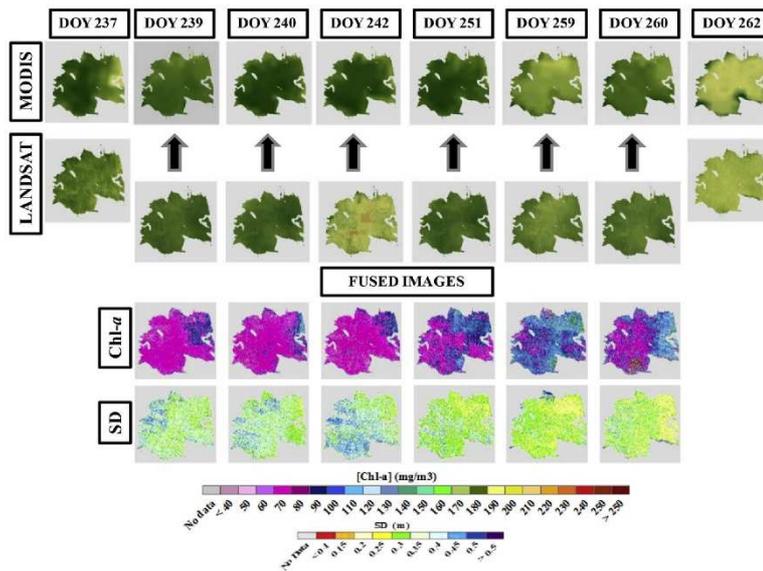


Fig. 6. Data fusion of Chl-*a* concentration and transparency (SD) cloud free maps generated for the study period of August–September 2006.

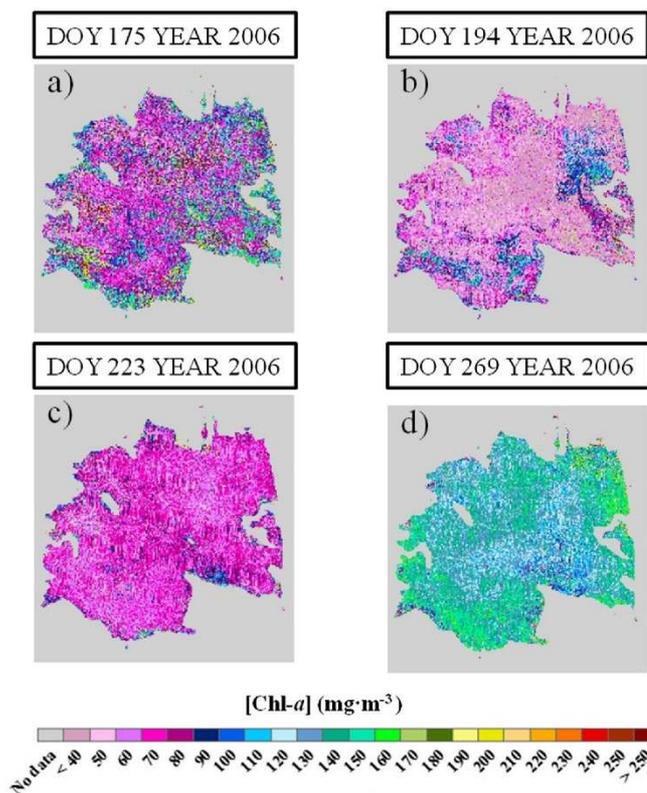


Fig. 7. Chl-*a* concentration maps of the Albufera de Valencia, given after applying the obtained algorithm to the fused images.

3.2. Data mining with GP models

For the estimation of Chl-*a* concentration and water transparency, Discipulus[®] software package was used to help generate the nonlinear regression equations with respect to the ground-truth data associated with each of the two water quality constituents of interest (Chl-*a* concentration and water transparency). Each of the nonlinear regression analyses derived by the GP modeling process is based on two subsets of data: one for the training and the other for the validation of the proposed GP models selected by Discipulus[®]. In the case of the Chl-*a* mapping, we had 52 samples in association with their respective fused reflectance values. 36 out of the 52 data points were applied for the training, whereas the rest of 16 were utilized for model validation. The best GP model was screened and selected for the Chl-*a* mapping (see Appendix 1). It is observed that the derived GP model for mapping Chl-*a* is credible since it has adjusted R^2 values of 0.94 and 0.98, associated with the training and validation processes, respectively (Fig. 4). The statistical values obtained for the validation (Table 5) showed the accurate estimation of Chl-*a* concentrations with an MADP of 6% and an RMSE of 8 mg m^{-3} relative to the average values of ground-truth data of 120 mg m^{-3} .

In the case of the estimation of water transparency, the best GP model was screened and selected for the water transparency mapping (see Appendix 2). The derived GP model for mapping water transparency is still credible since it has adjusted R^2 values of 0.89 and 0.82, associated with the training and validation processes, respectively (Fig. 5). The average value of the water transparency based on the validation data set is 34 cm along with a standard deviation of around 10 cm (Table 5). In addition, the statistics of the derived GP model for the water transparency mapping based on the validation data set presented an RMSE of 4 cm and an MADP of $\pm 9\%$. Given the range in the water transparency values it might be difficult to obtain an algorithm throughout regression techniques (Domínguez-Gómez et al., 2012). However, GP modeling techniques help solve this issue without the requisite of acquiring additional data sets to increase the range of the water transparency values.

Regression techniques were also evaluated to highlight the potential of GP models in this water body. In the case of the Chl-*a* estimation, regression coefficients lower than 0.50 were obtained for all the band combinations tested, with exception of band 2 which presented a R^2 value of 0.6. Regarding the water transparency (SDD) estimation, we could not define a good relationship

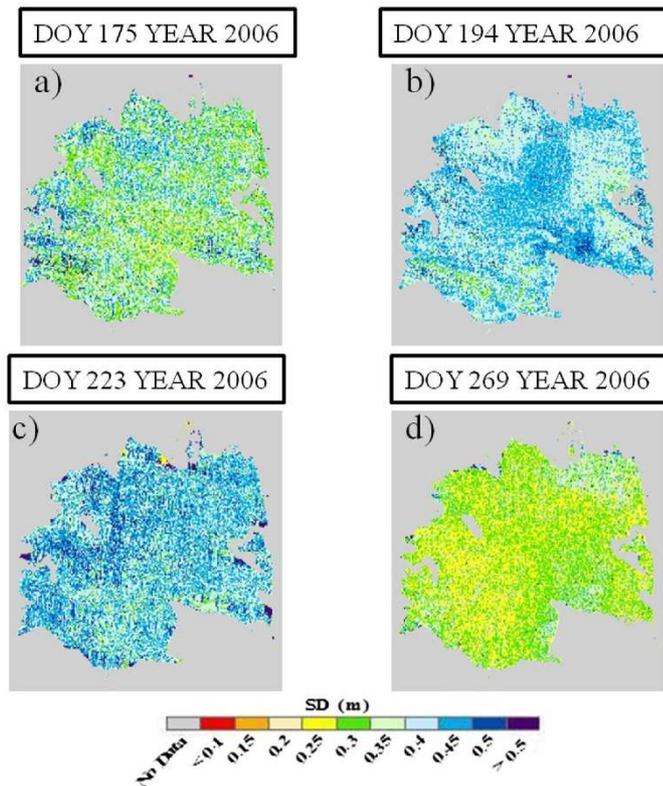


Fig. 8. Water transparency maps (SD, in m) of Albufera de Valencia, obtained by applying the developed algorithm to the fused images.

between the reflectance bands and this variable with this kind of regression techniques, as already pointed out by López-García and Caselles (1987). The main cause of this fact in the Albufera de Valencia is that the very low water transparency values and the high fluctuations of the groundwater data complicate the model calibration. This was evidenced by Doña et al. (2014) that carried out a companion study in the same water body obtaining a relationship between TM band 2 reflectance and SDD by means of regression techniques.

3.3. Generation of concentration maps

With the GP models available based on daily fused images, mapping the spatiotemporal variations of Chl-*a* concentration and water transparency on a daily basis became feasible. For illustration, the multi-sensor GP models developed above were applied based on the proposed data fusion process using surface reflectance values collected from mostly cloud-free days in August and September of 2006. The Chl-*a* concentration and water transparency SD maps are shown in Fig. 6 for the dates missing high-resolution Landsat images:

One set of snapshots corresponding to July–September 2006 was selected for the purpose of demonstration (Fig. 7). It is

indicative that the seasonality effect of the Albufera de Valencia is dependent to the different rice growing seasons (Romo et al., 2008). In July–September, the channel gates are normally closed during the rice growing period. This implies that water circulation is very low. Yet the level of water circulation would increase after the harvesting season in September–October when the gates are opened again. Romo et al. (2008) observed that the maximum values of Chl-*a* concentration in the Albufera de Valencia may be tied to the critical months of April and September coincidentally. When the gates are closed, it results in a longer period of hydrologic stability during the rice culture stage. Since no discharge is allowed into the sea for cleaning-up part of the phytoplankton, Chl *a* concentrations increase over time due to the lower level of water circulation from May to Aug (Fig. 7a–d), and the nutrient needs are mainly sustained by the internal nutrient load recycling. Note that Chl *a* concentrations in the first image (Fig. 7a) are even higher than those of the second image (Fig. 7b) due to rainfall effects. In addition, the gap filling procedure to solve the issue with Landsat 7 stripes might introduce effects on the different variable maps (Fig. 7c).

To confirm this hypothesis associated with Fig. 7b, rainfall data were collected for these months via the Integrated Information Water System (SIA) geographic visor tool available in the webpage

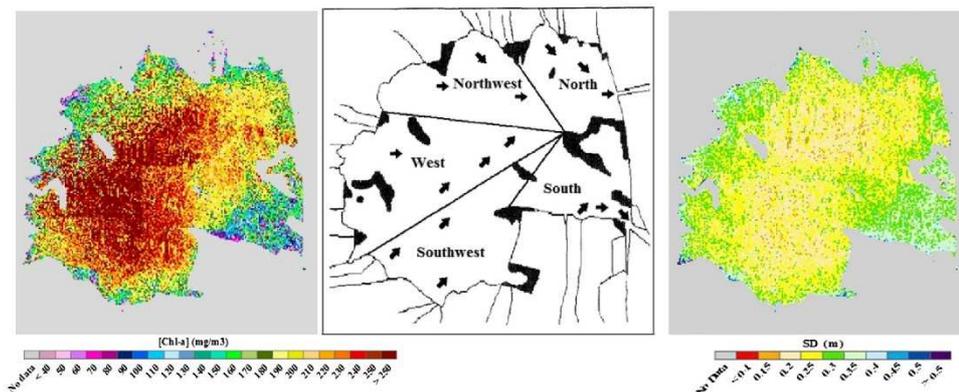


Fig. 9. Chlorophyll *a* concentration map (left), water transparency map (right) and lake sectors proposed by Soria and Vicente (2002) (center).

of the Spanish Ministry of Agriculture, Food and Environment. We found a significant storm 50 mm day^{-1} in south part of the Natural Park in the DOY 266 which made the water quality worsen in the September image (DOY 269).

Water transparency maps (Fig. 8) corroborate that low SDD values generally predominate in this water body, with the lower values of SDD corresponding to the higher values of Chl-*a* concentration, as expected. With these two figures (Figs. 7 and 8), their spatial coupling and temporal correlation are well described even in this heterogeneous environment. To evidence this contrast, the different hydrological sectors of the lake proposed by Soria and Vicente (2002) may be adopted (see the middle panel of Fig. 9). These authors differentiated the lake into 5 sectors, depending on the observed water circulation, quantity of flux, quality of water and other features.

The left and right panels of Fig. 9 show the contrast maps of Chl-*a* concentrations and water transparency in the DOY 112 of 2011, respectively. Based on the map of the different sectors presented in the central panel of the figure, it can be observed that the Northwest and West sectors present the higher values of Chl-*a* concentration and lower water transparency, due to low water circulation. Conversely, the lower values of Chl-*a* concentrations appear, with higher water transparency, in the northeast and southeast areas due to the proximity of the gates near these areas, allowing higher water flux at this region. The southwest area could be considered a transition zone, as shown in Fig. 9.

4. Conclusions

Our study demonstrates that the IDFM method fusing the Landsat and MODIS images, with the aid of STAR-FM algorithm, is useful to improve the temporal resolution of Landsat images. This solves the lengthy data gap posed by Landsat, providing more intensive Earth system observations. The GP models were produced for the estimation of Chl-*a* concentrations and water transparency simultaneously, based on the synthetic Landsat reflectance values. Satisfactory results were obtained for the estimation of Chl-*a* and water transparency, with high correlations between the predicted values and the observed values at both training and validation stages. This IDFM drastically

improves the traditional effort by means of linear regression techniques. The robustness of these GP models can be improved with the addition of more ground-truth data for training when exposing the GP models to a wider range of Chl-*a* concentrations and water transparency data.

With the aid of IDFM, we applied such a tool to rapidly derive the heterogeneous maps addressing spatial distributions of Chl-*a* concentration and water transparency, as well as their temporal changes in the Albufera de Valencia. In the case of our study in the Albufera de Valencia, the comparison of the spatial and temporal changes of water quality resulting in strong water management implications. It clearly offers a powerful insight that can be used to improve management practices that allow us to increase the ecological health of the lake, as required by the WFD. This advancement provides a cost-effective way for monitoring water quality in lakes, which is linked to future water resources management. Note that the cloud-cover is the primary limitation of this approach. If clouds obstruct the study area, STAR-FM cannot determine the surface reflectance. Future work may be directed toward developing a cloud removal algorithm for restoring more useful remote sensing images.

Acknowledgments

This study was jointly supported by the "Ministerio de Economía y Competitividad" by aid granted projects CGL2010-17577-CLL, CGL2013-46862-C2-1/2-P, PROMETEUII/2014/086 to VC and CGL2012-38909 to AC. The authors would like to thank Mr. Yan-Oing Dong for providing the code to solve the issue of Landsat-7 data.

Appendix 1. Chlorophyll *a* GP algorithm

The resulting GP algorithm derived using the fused band data is given below. Due to the equation's lengthy complexity, variables X_1 to X_{25} have been introduced to simplify it. The variables $v_0, v_1, v_2, v_3, v_4,$ and v_5 correspond to the surface reflectance values of the synthetic images, where the spectral range of every image band correspond to: $v_0 = \text{Landsat B1}, v_1 = \text{Landsat B2}, v_2 = \text{Landsat B3}, v_3 = \text{Landsat B4}, v_4 = \text{Landsat B5}$ and $v_5 = \text{Landsat B7}$.

$$\text{Chl } a = 60.9 - (153 * (F2XM1(X_{25})))$$

where X_1 to X_{25} are as follows:

$$X_1 = \sin(2v_4 - 2v_5)$$

$$X_2 = v_4 - v_5$$

$$X_3 = (X_2 + X_1 - 1.20 \text{fix}(X_2 + 3X_1))^2$$

$$X_4 = v_4 - v_2 + X_3$$

$$X_5 = \text{abs}\left(\frac{16X_4}{v_0} + 31.8\right)$$

$$X_6 = 2v_{45} - 2\text{fix}(v_4 - 18.9X_5)$$

$$X_7 = \frac{2X_4}{v_0} - \text{fix}(X_6)$$

$$X_8 = \frac{2}{X_3} + 1.08$$

$$X_9 = X_7 - \text{fix}(v_4 - 18.9X_5)$$

$$X_{10} = v_4 + (X_2 + X_1 - 1.20 \text{fix}(X_2 + 3X_{21}))^2$$

$$X_{11} = X_9 - \frac{\left(\frac{X_{10}}{v_4 + X_8} + X_8\right)}{X_4} + 3.98$$

$$X_{12} = v_4 + v_5 - 18.9X_5 + X_{11}$$

$$X_{13} = 2 * (v_1 + \sin(v_4 + \text{fix}(X_6)))$$

$$X_{14} = \left(v_4 + 2v_5 - 18.9X_5 + \frac{X_{13}}{X_{12}} + X_{11}\right)$$

$$X_{15} = \frac{\left(\frac{X_{10}}{v_4 + X_8} + X_8\right)}{X_4}$$

$$X_{16} = 8 * (v_1 + \sin(v_4 + \text{fix}(X_6)))$$

$$X_{17} = X_{15} + \frac{F2XM1\left(4v_5 + \frac{X_{16}}{X_{17}}\right)}{X_{14}} - 3.98$$

$$X_{18} = 18.9X_5 - 2v_5 - v_4 - \frac{X_{13}}{X_{12}}$$

$$X_{19} = \text{fix}(X_6) + \text{fix}(v_4 - 18.9X_5) + X_{17}$$

$$X_{20} = 4 * \left(2v_1 + 11.4 \cos\left(X_{18} - \frac{2X_4}{v_0} + X_{19}\right) * X_{14}\right)$$

$$X_{21} = \frac{2.17 * \left(2v_5 + \frac{X_{20}}{X_{14}}\right)^2 * X_{14}}{v_0^2}$$

$$X_{22} = v_2 + X_{21}$$

$$X_{23} = 57.1 F2XM1\left(\frac{4X_{22}}{v_3} - 5.26\right)$$

$$X_{24} = \text{abs}\left(v_2 - v_4 = X_{23} - \frac{18.9X_{14}}{v_2 + \frac{X_{21}}{v_0} - 5.51}\right)$$

$$X_{25} = v_5 - 75.6 * \text{abs}\left(\frac{4X_{24}}{v_{03}}\right)$$

Note: $\text{abs}()$ is the absolute value of the parenthetic expression. The $\text{fix}()$ expression rounds the value down to the nearest integer value. $F2XM1()$ is an exponential function; if the absolute value of the parenthetic expression is less than or equal to 1, then the $F2XM1$ instruction calculates two raised to the $f(x)$ power, minus one. This expression is as follows:

$$F2XM1 = (2^x - 1); \quad \text{abs}(x) \leq 1$$

Appendix 2. Water transparency GP algorithm

The resulting GP algorithm derived using the fused band data is given below. Due to the equation's lengthy complexity, variables X_1 to X_{10} have been introduced to simplify it. The variables $v_0, v_1, v_2, v_3, v_4,$ and v_5 correspond to the surface reflectance values of the synthetic images, where the spectral range of every image band correspond to: $v_0 = \text{Landsat B1}, v_1 = \text{Landsat B2}, v_2 = \text{Landsat B3}, v_3 = \text{Landsat B4}, v_4 = \text{Landsat B5}$ and $v_5 = \text{Landsat B7}$.

$$SD = X_{10} - 6X_6 - \frac{2X_3}{X_1}$$

where X_1 to X_{10} are as follows:

$$X_1 = \cos(X_3^2)$$

$$X_2 = 1.2 * v_2 - v_0$$

$$X_3 = \sin(9.45v_3 - v_5 - 7.36)$$

$$X_4 = v_5 * (2X_6^2 + 3.29)$$

$$X_5 = \text{abs}\left(\frac{3X_3}{X_1} + X_4\right)$$

$$X_6 = \cos\left(\frac{F2XM1(X_2)}{v_4}\right)$$

$$X_7 = \left(\frac{1.05 * \left(\frac{2X_5}{X_1} - 3.98 \right)}{X_6} \right)$$

$$X_8 = \frac{(4X_6 - X_7)^2}{X_5}$$

$$X_9 = \frac{1.09X_5}{4.0 * X_6 - X_8} - 4.35X_6 + X_8$$

$$X_{10} = (v_1 + 1.09 * abs(X_9))^{1/2}$$

Note: *abs()* is the absolute value of the parenthetic expression. The *fix()* expression rounds the value down to the nearest integer value. *F2XMI()* contains a conditional nested within a higher-tiered conditional. *F2XMI()* is an exponential function; if the absolute value of the parenthetic expression is less than or equal to 1, then the *F2XMI* instruction calculates two raised to the *f(x)* power, minus one.

This expression is as follows:

$$F2XMI = (2^x - 1); \quad abs(x) < 1$$

References

- Allan, M.G., Hamilton, D.P., Hicks, B.J., Brabyn, L., 2011. Landsat remote sensing of chlorophyll-a concentrations in central North Island lakes of New Zealand. *Int. J. Remote Sens.* 32, 2037–2055.
- Allee, R.J., Johnson, J.E., 1999. Use of satellite imagery to estimate surface chlorophyll-a and Secchi disc depth of Bull Shoals Reservoir, Arkansas, USA. *Int. J. Remote Sens.* 20, 1057–1072.
- Alonso-Fernández, J.R., Díaz-Muñiz, C., García-Nieto, P.J., De Cos Juez, F.J., Sánchez-Lasheras, F., Roqueñi, M.N., 2013. Forecasting the cyanotoxins presence in fresh waters: a new model based on genetic algorithms combined with the MARS technique. *Ecol. Eng.* 53, 68–78.
- Bricker, S.B., Clement, C.G., Pirhalla, D.E., Orlando, S.F., Farrow, D.R.G., September 1999. National Estuarine Eutrophication Assessment, Effects of Nutrient Enrichment in the Nation's Estuaries. NOAA, 71 pp.
- Caselles, V., Artiago, M.M., Hurtado, E., Coll, C., Brava, A., 1998. Mapping actual evapotranspiration by combining Landsat TM and NOAA-AVHRR images: application to the Barrax area, Albacete, Spain. *Remote Sens. Environ.* 63, 1–10.
- Chang, N.B., Vannah, B.W., Yang, Y.J., Elowitz, M., 2014a. Integrated data fusion and mining techniques for monitoring total organic carbon concentrations in a lake. *Int. J. Remote Sens.* 35, 1064–1093.
- Chang, N.B., Vannah, B.W., Yang, Y.J., 2014b. Comparative sensor fusion between hyperspectral and multispectral satellite sensors for monitoring microcystin distribution in Lake Erie. *IEEE J. Sel. Top. Appl. Earth Obs.* 7, 2426–2442.
- Dominguez, J.A., Alonso, C., Alonso, A., 2011. Remote sensing as a tool for monitoring water quality parameters for Mediterranean Lakes of European Union water framework directive (WFD) and as a system of surveillance of cyanobacterial harmful algae blooms (SCyanoHABs). *Environ. Monit. Assess.* 181, 317–334.
- Dominguez-Gómez, J.A., Rodríguez-Pérez, D., Marcos-Martín, C., Chao-Rodríguez, Y., Delgado-Rojas, G., 2012. Estudio de aguas continentales mediante teledetección. UNED, Spain.
- Doña, C., Sánchez, J.M., Caselles, V., Domínguez, J.A., Camacho, A., 2014. Empirical relationships for monitoring water quality of lakes and reservoirs through multispectral images. *IEEE J. Sel. Top. Appl. Earth Obs.* 7, 1632–1641.
- Duan, H., Zhang, Y., Zhang, B., Song, K., Wang, Z., 2006. Assessment of chlorophyll-a concentration and trophic state for Lake Chagan using Landsat TM and field spectral data. *Environ. Monit. Assess.* 129, 295–308.
- Dwyer, J., Schmidt, G., 2006. The MODIS reprojection tool. In: Qu, J.J., Gao, W., Kafatos, M., Murphy, R.E., Salomonson, V.V. (Eds.), *Earth Science Satellite Remote Sensing*. Springer Berlin Heidelberg, pp. 162–177.
- Environmental System Research Institute (ESRI), 2013. <http://bbs.esrichina-bj.cn/ESRI/viewthread.php?tid=52263&extra=&page=4> (accessed by Oct. 2013).
- European Water Framework Directive (WFD), European Commission. Directive 2000/60/EC of the European Parliament and of the Council of 23 October 2000 establishing a framework for Community action in the field of water policy. *Off. J. Eur. Commun.* 327, 1–72.
- Francone, D.F., 1998. *Discipulus Software Owner's Manual*. Version 3.0 DRAFT. Machine Learning Technologies, Inc., Colorado.
- Gao, F., Masek, J., Schwaller, M., Hall, F., 2006. On the blending of Landsat and MODIS surface reflectance: predicting daily Landsat surface reflectance. *IEEE Trans. Geosci. Remote Sens.* 44, 2207–2218.
- Genderen, J., Pohl, C., 1994. Image fusion: issues, techniques, and applications. *Intelligent image fusion*. In: *Proc. EARSeL Workshop*, Strasbourg, France, pp. 18–26.
- Gitelson, A.A., Dall'Olmo, G., Moses, W., Rundquist, D.C., Barrow, T., Fisher, T.R., Gurlin, D., Holz, J., 2008. A simple semi-analytical model for remote estimation of chlorophyll-a in turbid waters: validation. *Remote Sens. Environ.* 112, 3582–3593.
- Gitelson, A.A., Schalles, J.F., Haladik, C.M., 2007. Remote chlorophyll-a retrieval in turbid, productive estuaries: Chesapeake Bay case study. *Remote Sens. Environ.* 109, 464–472.
- Härmä, P., Vepsäläinen, J., Hannonen, T., Pyhäläthi, T., Kämäri, J., Kallio, K., Eloheimo, K., Koponen, S., 2001. Detection of water quality using simulated satellite data and semi-empirical algorithms in Finland. *Sci. Total Environ.* 268, 107–121.
- Koza, J., 1992. *Genetic Programming: on the Programming of Computers by Means of Natural Selection*. The MIT Press, Cambridge, MA.
- Lawrenz, E., Richardson, T.L., 2011. How does the species used for calibration affect chlorophyll a measurements by in situ fluorometry? *Estuar. Coasts* 34, 872–883.
- López-García, M.J., Caselles, V., 1987. Use of thematic mapper data to assess water quality in Albufera lagoon of Valencia (Spain). In: *Proc. 13th Ann. Conf. of Remote Sensing Society*, Nottingham, UK, pp. 510–519.
- Mancino, G., Nolè, A., Urbano, V., Amato, M., Ferrara, A., 2009. Assessing water quality by remote sensing in small lakes: the case study of Monticchio lakes in southern Italy. *iForest* 2, 154–161.
- Margalef, R., 1983. *Limnologia*. Omega, Barcelona, Spain.
- Masek, J.G., Vermote, E.F., Salous, N., R Wolfe, Hall, F.G., Huemmrich, F., Gao, F., Kutler, J., Lim, T.K., 2006. A Landsat surface reflectance data set for North America, 1990–2000. *Geosci. Remote Sens. Lett.* 3, 68–72.
- Mayo, M., Gitelson, A., Yacobi, Y.Z., Ben-Avraham, Z., 1995. Chlorophyll distribution in Lake Kinneret determined from Landsat thematic mapper data. *Int. J. Remote Sens.* 16, 175–182.
- McCullough, I.M., Loftin, C.S., Sader, S.A., 2012. High-frequency remote monitoring of large lakes with MODIS 500 m imagery. *Remote Sens. Environ.* 124, 234–241.
- O'Reilly, J.E., Maritorena, S., Mitchell, B.G., Siegel, D.A., Carder, K.L., Garver, S.A., Kahru, M., McClain, C., 1998. Ocean color chlorophyll algorithms for SeaWiFS. *J. Geophys. Res.* 103, 24937–24953.
- OECD, 1982. *Eutrophication des Eaux. Methodes de surveillance, d'évaluation et de lutte (Eutrophication of Water, Monitoring, Assessment and Control)*. France, Paris.
- Olmanson, L.G., Bauer, M.E., Brezonik, P.L., 2008. A 20-year Landsat water clarity census of Minnesota's 10,000 lakes. *Remote Sens. Environ.* 112, 4086–4097.
- Peña, R., Ruiz, A., Domínguez, J.A., 2004. Mapping of photosynthetic pigments in Spanish inland waters using MERIS imagery. In: *Envisat Symp.* Salzburg, Austria.
- Picazo, A., Rochera, C., Vicente, E., Miracle, M.R., Camacho, A., 2013. Determination of photosynthetic pigments by spectrophotometric methods in stratified lakes: a critical analysis based on comparisons with HPLC determinations in a model lake. *Limnetica* 32, 139–158.
- Pinto, A.M.F., Von Sperling, E., Moreira, R.M., 2001. Chlorophyll-a determination via continuous measurement of plankton fluorescence: methodology development. *Water Res.* 35, 3977–3981.
- Pohl, C., Van Genderen, J., 1998. Multisensor image fusion in remote sensing: concepts, methods, and applications. *Int. J. Remote Sens.* 19, 823–854.
- Romo, S., García-Murcia, A., Villena, M.J., Sánchez, V., Ballester, A., 2008. Phytoplankton trends in the lake of Albufera de Valencia and implications for its ecology, management, and recovery. *Limnetica* 27, 11–28.
- Seifert, J., 2004. *Data Mining: an Overview*. CRS report for congress.
- Serrano, M.L., Camacho, A., Vicente, E., Peña, R., 1997. Estudio por teledetección de la evolución del estado trófico de tres embalses del ámbito de la Confederación Hidrográfica del Júcar en el período estival de los años 1994 y 1995. *Limnetica* 13, 5–14.
- Siegel, D.A., Wang, M., Maritorena, S., Robinson, W., 2000. Atmospheric correction of satellite ocean color imagery: the black pixel assumption. *Appl. Opt.* 39, 3582–3591.
- Soria, J.M., Vicente, E., 2002. Estudios de los aportes hídricos al parque natural de la Albufera de Valencia (Studies of water flow contribution to the natural park of the Albufera de Valencia). *Limnetica* 2, 105–115.
- Strickland, J.D.H., Parsons, T.P., 1972. *A Practical Handbook for Seawater Analysis*. Fisheries Research Board, Ottawa, Canada.
- Vermote, E.F., Tanre, D., Deuze, J.L., Herman, M., Morcrette, J.J., 1997. Second simulation of the satellite signal in the solar spectrum, 6S: an overview. *IEEE Trans. Geosci. Remote Sens.* 35, 675–685.
- Vicente, E., Miracle, M.R., 1992. The coastal Lagoon Albufera de Valencia: an ecosystem under stress. *Limnetica* 8, 87–100.
- Willmott, C.J., 1982. Some comments on the evaluation of model performance. *Bull. Am. Meteor. Soc.* 11, 1309–1313.

Anexo C

Ponencia “Monitoring inland water quality status using images from the SPOT-5 take-5 experiment”

Este trabajo se realizó en colaboración con los doctores A. Picazo, C. Rochera y J. M. Galve, fue presentado en el ESA Living Planet Symposium en mayo de 2016 y publicado en el libro de actas en agosto de 2016. En este trabajo se recurrió a imágenes simuladas de Sentinel-2, ya que ofrece la oportunidad de continuar con la tarea de control de calidad de las aguas continentales, gracias a su ciclo de revisita de 5 días y la resolución espacial de 10-30 m. En el marco de las actividades preparatorias de Sentinel-2, la ESA desarrolló el experimento SPOT take 5 ofreciendo imágenes simuladas de Sentinel-2. Con los datos de estas imágenes se desarrollaron algoritmos para estimar los parámetros de calidad del agua en tres cuerpos de agua pertenecientes a la cuenca del río Júcar. Una vez más, se hizo uso de la programación genética para generar ecuaciones de regresión no lineales entre las medidas de campo y los valores de reflectividad.

MONITORING INLAND WATER QUALITY STATUS USING IMAGES FROM THE SPOT-5 TAKE-5 EXPERIMENT

Carolina Doña ⁽¹⁾, Juan Manuel Sánchez ⁽²⁾, Vicente Caselles ⁽¹⁾, Antonio Camacho ⁽³⁾, Antonio Picazo ⁽³⁾, Carlos Roquera ^(3,4), Joan Miquel Galve ⁽¹⁾

⁽¹⁾ Earth Physics and Thermodynamics Department, Faculty of Physics, University of Valencia, 46100 Burjassot (Spain), Email: Carolina.Dona@uv.es ; Vicente.Caselles@uv.es ; Joan.Galve@uv.es

⁽²⁾ Department of Applied Physics, Polytechnic School, University of Castilla-La Mancha, 16071 Cuenca (Spain), Email: JuanManuel.Sanchez@uclm.es

⁽³⁾ Cavanilles Institute of Biodiversity and Evolutionary Biology, University of Valencia, 46100 Burjassot (Spain), Email: Antonio.Camacho@uv.es ; carlos.roquera@uv.es ; antonio.picazo-mozo@uv.es

⁽⁴⁾ Regional Centre for Water Studies (CREA), University of Castilla-La Mancha, 13071 Ciudad Real (Spain)

ABSTRACT

The aim of this work is to develop algorithms to estimate the water quality from SPOT-5 images in three water bodies belonging to the Júcar river basin. Images from the SPOT5 take 5 experiment were used, in the framework of the Sentinel-2 preparatory activities. Based on the spectral matching between both VNIR sensors, SPOT-5 was used to simulate Sentinel-2 products. Several experimental campaigns were carried out concurrent with SPOT-5 overpasses or close in date. Chlorophyll-a concentration [Chl-a], transparency (SD) and total suspended particles concentration [TSS] were measured. Genetic programming models were used to generate nonlinear regression equations between ground measurements and reflectance values from the SPOT-5 spectral bands. Results showed mean errors of $\pm 8\%$, $\pm 5\%$ and $\pm 10\%$ in the estimation of [Chl-a], SD and [TSS], respectively. These results show the potential of Sentinel-2 to monitor and study the spatio-temporal trend of these water quality parameters.

1. INTRODUCTION

The trophic state of lakes and reservoirs has been studied for years since water quality can deteriorate due to the eutrophication phenomenon. This is a critical issue in the nexus of water supply, environmental management, and ecosystem conservation. This issue is of great importance due to the requirements of new environmental initiatives, such as the European Water Framework Directive (WFD) in 2000, which establishes that all EU countries have to maintain their water bodies, achieving a good ecological status with the aid of a suite of favourable conservation actions. Remote sensing techniques have been shown useful to estimate water quality such as chlorophyll-a, total suspended particles and water transparency.

Since the entry into force of this WFD, many water quality monitoring studies have focused on mapping different water quality parameters using remote sensing technologies[1]-[11]. Landsat imagery has been widely used due to its medium spatial resolution of 30 m and its

historical archive, which allows to conduct multitemporal studies of water quality in water bodies [7]-[9]. However, Landsat poor revisit time (16 days) and coarse spectral resolution makes it difficult to determine certain water quality parameters timely and accurately. Sensors with a better temporal-spectral resolution such as MODIS and MERIS are more suitable to monitor several water quality parameters [4], [10]. But in this case the problem is the coarse spatial resolution of their images, which is too coarse to analyze medium or small size water bodies [11]. Some studies take the advantage of the spectral equivalence between Landsat and MODIS (daily revisit time) or MERIS (3-day revisit time) to generate synthetic Landsat images through fusion techniques [12]-[13]. By applying these techniques maps of the different parameters can be obtained without Landsat overpass. Different algorithms have been recently proposed for the estimation of water quality variables using TM data and Deimos-1 (22-m spatial resolution, and around 3-day revisit time) [14]. These algorithms were derived and tested using ground measurements from a set of lakes and ponds in Spain.

Sentinel-2 offers the opportunity to continue with this inland water quality monitoring task, thanks to its 5-day revisit cycle and 10-30 m spatial resolution. In the framework of the Sentinel-2 preparatory activities ESA developed the SPOT5 take 5 experiment. From early April to the end of August 2015, SPOT-5 satellite was relocated in a 5-day orbit, before being decommissioned. Based on the spectral matching between both VNIR sensors, SPOT-5 was used to simulate Sentinel-2 products and show the benefits of its high spatial resolution to monitor small water bodies.

The aim of this work is to develop algorithms to estimate the water quality parameters from SPOT-5 images in three water bodies belonging to the Júcar river basin. Several experimental campaigns were carried out concurrent with SPOT-5 overpasses or close in date to accomplish this aim. Genetic programming models were used to generate nonlinear regression

equations between ground measurements and reflectance values from the SPOT-5 spectral bands.

2. MATERIALS AND METHODS

In this work we focused on the Jucar river basin, located in the central part of the Mediterranean coast of Spain, in the province of Valencia. The area of interest is centered at 39.30 °N, -0.45° W. A few reservoirs plus La Albufera de Valencia Lake are contained in the SPOT5 scenes provided (Fig. 1). La Albufera de Valencia Lake is a hypereutrophic system with concentrations of chlorophyll-a reaching values up to 717 µg/L [15], and annual averages higher than 100 µg/L. After its recognition as a Natural Park in 1985, and its inclusion 4 years later as part of the Ramsar list of international protected wetlands, La Albufera de Valencia has been the target of numerous limnological studies [16]. Escalona and Tous reservoirs have been also studied in this work. They are oligotrophic-mesotrophic water bodies presenting low chlorophyll-a concentrations values and high values of transparency.

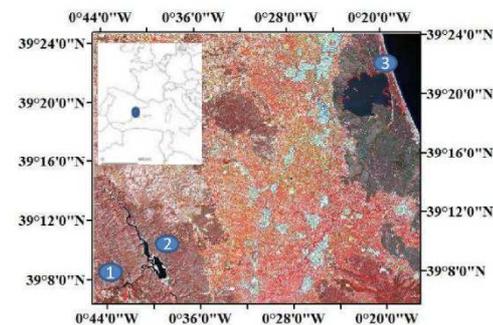


Figure 1. Study area. SPOT-5 false colour image (RGB: bands 3, 2 and 1). 1) Escalona, 2) Tous, 3) La Albufera de Valencia Lake.

The ground-truth data used in this work consist in transparency, chlorophyll-a ([Chl-a]) and total suspended solids concentration ([TSS]). Several field campaigns were carried out concurrent with SPOT-5 overpasses or close in date. Measurements in the Albufera de Valencia Lake were taken for 24th June and 9th July, 2015. Data from the single date 24th July, 2015 were available in Escalona and Tous reservoirs. Transparency was measured in situ with a Secchi disk (SD) whereas water samples were collected in a set of points (Fig.2) and then analyzed at the laboratory. The [Chl-a] used as an indicator of algal biomass, was determined spectrophotometrically after pigment extraction in acetone according to Strickland and Parsons [17]. The [TSS], which includes degraded, dead and non-organic particles, was determined by filtering

the water samples through pre-weighted AH - 934 glass - fibre filters. Finally, [TSS] was obtained after drying for 5 h at 103 °C by considering the final weight and the volume of the initial sample.

The Chl-a concentration measurements (Tab. 1), ranged from a minimum value of 50 µg/L, in July 9th, to a maximum of 175 µg/L in June 24th. These values are far higher than the water quality standard established by the WFD for this type of water bodies, which evaluates the achievement of good ecological status. In the case of the water transparency (Tab. 1), SD values did not exceed 33 cm. Therefore, the Albufera de Valencia is deemed as a hypertrophic lake requiring intensive water quality monitoring as required by the WFD to achieve the environmental goals.

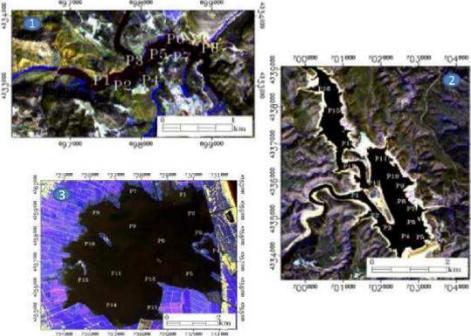


Figure 2. Sample points in: 1) Escalona, 2) Tous and 3) La Albufera de Valencia Lake.

Values were significantly lower for the two reservoirs. Escalona presented values of [Chl-a] and [TSS] lower than 2 µg/L and 4 mg/L, respectively, and values of transparency over 2 m. Water quality was even better in Tous reservoir, with typical oligotrophic values of [Chl-a] lower than 0.6 µg/L and values of transparency over 4 m (Tab. 1).

Images from SPOT-5 were used in this work. The L2A surface reflectance products were downloaded from <https://spot-take5.org/>. All data are processed by THEIA land data centre and distributed with a free and open policy via the joint ESA-CNES web portal spot-take5.org. The images used in this work are concurrent with the field campaigns.

Genetic programming (GP) models were used to generate nonlinear regression equations between ground measurements and reflectance values from the SPOT-5 spectral bands. This method is developed based on genetic algorithms. It is a subclass of data mining computation techniques such as the Artificial Neural Networks (ANN) and Support Vector Machines (SVM).

One of the principal advantages of the GP in comparison with other data-mining techniques is that is a white box.

Table 1. Average, maximum, minimum values, standard deviation and total number of samples for Chlorophyll-a, total suspended solids and transparency registered in the study areas during the sampling campaigns carried out in 2015.

Site	Albufera	Albufera	Tous	Escalona
Date	06/24/15	07/09/15	07/24/15	07/24/15
N	13	15	18	8
Chl-a (µg/L)				
Average	125	70	0.30	0.80
Max.	175	85	0.60	2.30
Min.	80	50	0.09	0.30
Std Dev	35	10	0.12	0.60
TSS (mg/L)				
Average	120	125	2	3
Max.	160	160	10	4
Min.	80	85	0.6	1.3
Std Dev	25	20	2	0.9
SD (m)				
Average	0.18	0.30	5.3	3.2
Max.	0.22	0.33	6.7	3.7
Min.	0.16	0.26	3.8	2.3
Std Dev	0.013	0.02	0.7	0.5

The GP can decode system behaviours based on empirical data for symbolic regression [18]. In this work we developed a GP model or program that uses the inputs of surface reflectance bands to predict the outputs, which are [Chl-a], [TSS] and SD [12]. The ground-truth dataset was divided into two subsets, 2/3 of the total dataset for GP model training and 1/3 for the validation (1/3 of total data). Discipulus® was the software used in this study to perform GP.

Finally, the GP algorithms showing the best training-validation fits were selected. Statistical analyses of the comparison between algorithm predictions and field measurements were done following [19], including the root mean square error (RMSE), the mean absolute difference (MAD), the mean absolute percent difference (MAPD), and the biased estimator (BIAS).

Maps of Chlorophyll-a concentration, transparency and total suspended particles concentration were generated, and the evolution of the trophic state of the Albufera Lake and some other reservoirs was analyzed in terms of these parameters.

3. RESULTS

In terms of [Chl-a], results showed a correlation coefficient of 0.99 and 0.98 for the training and validation of the best GP algorithm, respectively. Table 2 shows the statistical values obtained with the data reserved for the validation of the algorithm. Values of RMSE = 6 µg/L and MADP = ±8% were obtained.

Similar results were observed in the case of the [TSS] and SD, with values of the correlation coefficient close to 1 for both the training and validation and an error lower than ±10%. In comparison with the other parameters, the [TSS] showed the lower statistical values in its estimation with an error of 10%, but the statistical values were acceptable in all cases, with a RMSE of 8 mg/L.

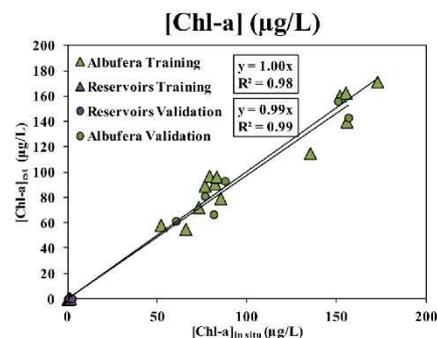


Figure 3. Estimated versus observed chlorophyll-a obtained with the GP model. Both, training and validation regressions are shown.

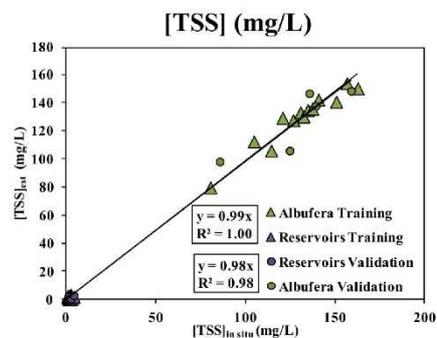


Figure 4. Estimated versus observed total suspended solids obtained with the GP model. Both, training and validation regressions are shown.

These results are in agreement with those obtained by [12]. Focusing on La Albufera de Valencia, results in the estimation of [Chl-a] are very similar to those reported in [12] (RMSE of 8 µg/L and a relative error of ±6%). In the present case, a RMSE of 9 µg/L and a MADP of ±7% are obtained. Errors are unacceptable for the reservoirs due to the low values of [Chl-a] in these water bodies combined with the limited number of data. In terms of transparency, results of RMSE=4 cm and relative error of 15% are obtained for La Albufera Lake, whereas for the reservoirs RMSE value is 20 cm, but in this case with a lower relative error of only 4%.

Finally, in terms of [TSS], results for La Albufera de Valencia Lake showed RMSE values of 12 mg/L and MADP of 8%, similar to the RMSE= 10 mg/L and MADP=15% reported by [14]. Again, dataset for the reservoirs is too short to provide acceptable statistics in terms of [TSS].

Table 2. Summary of the statistical assessment between predicted and observed [Chl-a], [TSS] and water transparency using the whole validation dataset.

Statistic	[Chl-a] ($\mu\text{g/L}$)	[TSS] (mg/L)	SD (m)
Observed Values	41	48	2.4
Std Dev _{obs}	56	64	2.4
Predicted Values	40	46	2.5
Std Dev _{pred}	56	64	2.4
BIAS	-1.05	-1.5	0.08
MAD	3	5	0.12
RMSE	6	8	0.17
MADP (%)	8	10	5

Note: Std Dev_{obs} and Std Dev_{pred} stand for the standard deviation of observed and predicted values, respectively.

As a demonstration of the potential of this tool, some examples of maps for the different parameters are showed in figures 6-8. Note the high values of [Chl-a] in La Albufera de Valencia, and its lower transparency, as a comparison to the reservoirs. Also, note the heterogeneous distribution of these variables.

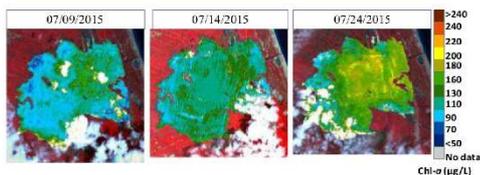


Figure 6. Examples of chlorophyll-a concentration maps of La Albufera de Valencia. The specific algorithm derived for the SPOT-5 images was applied.

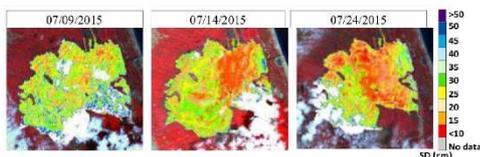


Figure 7. Examples of Transparency maps of La Albufera de Valencia. The specific algorithm derived for the SPOT-5 images was applied.

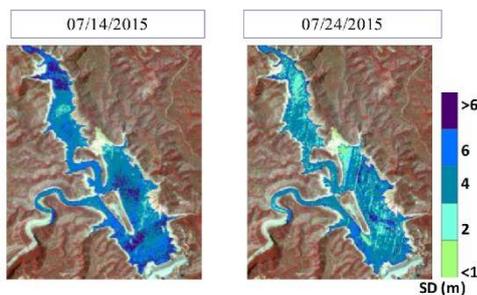


Figure 8. Examples of transparency maps for the Tous reservoir. The specific algorithm derived for the SPOT-5 images was applied.

4. CONCLUSIONS

This work shows some preliminary results of simulated Sentinel-2 water quality status products using images from the SPOT-5 take-5 experiment. The algorithms developed to estimate the different water quality parameters showed mean errors of $\pm 8\%$, $\pm 10\%$, $\pm 5\%$, for chlorophyll-a concentration, total suspended particles concentration and transparency, respectively. Focusing on La Albufera Lake, results are similar to those already reported in a previous works. However, the low values of chlorophyll-a and suspended solids concentrations observed in the two studied reservoirs make difficult the training of an accurate algorithm for these water bodies. Additional measurements and further analysis are required before extracting any firm conclusion about the performance of the algorithms in reservoirs. These results show the potential of Sentinel-2 to monitor and study the spatio-temporal trend of these water quality parameters.

5. ACKNOWLEDGEMENTS

This study was jointly supported by the Spanish Ministry of Economy and Competitiveness and by the European Union through the European Fund for Regional Development "One way to make Europe", through projects CGL2013-46862-C2-1/2-P; "ECOLAKE" (CGL2012-38909) and "CLIMAWET" (CGL2015-69557-R), the last two to AC. Conselleria d'Educació, Cultura i Esport of Generalitat Valenciana also supported this work with a PROMETEOII/2014/086 project. CR is a recipient of a VALi+d postdoctoral contract of Generalitat Valenciana APOSTD/2015/088.

6. REFERENCES

1. Mcullough, I. M.; Loftin, C. S.; Sader & S. A. (2013). High-frequency remote monitoring of large lakes with MODIS 500 m imagery. Remote Sens. Environ. **124**, 234–241.

2. Serrano M. L.; Camacho, A., Vicente, E. & Peña, R. (1997). Estudio por teledetección de la evolución del estado trófico de tres embalses del ámbito de la Confederación Hidrográfica del Júcar en el periodo estival de los años 1994 y 1995. *Limnetica*, **13**, 5-14.
3. Härmä, P., Vepsäläinen, J., Hannonen, T., Pyhäläthi, T., Kämäri, J., Kallio, K., Eloheimo, K. & Koponen, S. (2001). Detection of water quality using simulated satellite data and semi-empirical algorithms in Finland. *Sci. Tot. Environ.* **268**, 107-121.
4. Peña, R., Ruiz, A. & Domínguez J. A. (2004). Mapping of photosynthetic pigments in Spanish inland waters using MERIS imagery. *Envisat Symp. Salzburg, Austria*.
5. Duan, H., Zhang, Y., Zhang, B., Song, K. & Wang, Z. (2006). Assessment of chlorophyll-a concentration and trophic state for Lake Chagan using Landsat TM and field spectral data. *Environ. Monit. Assess.* **129**, 295-308.
6. Allan, M. G., Hamilton, D. P., Hicks, B. J. & Brabyn L. (2011). Landsat remote sensing of chlorophyll-a concentrations in central North Island lakes of New Zealand. *Int. J. Rem. Sens.* **32**, 2037-2055.
7. Olmanson, L. G., Bauer, M. E. & Brezonik, P. L. (2008). A 20-year Landsat water clarity census of Minnesota's 10,000 lakes. *Rem. Sens. Environm.* **112**, 4086-4097.
8. Mayo, M., Gitelson, A., Yacobi, Y. Z. & Ben-Avraham, Z. (1995). Chlorophyll distribution in Lake Kinneret determined from Landsat Thematic Mapper data. *Int. J. Remote Sens.* **16**, 175-182.
9. Allee, R. J. & Johnson, J. E. (1999). Use of satellite imagery to estimate surface chlorophyll-a and Secchi disc depth of Bull Shoals Reservoir, Arkansas, USA. *Int. J. Remote Sens.* **20**, 1057-1072.
10. McCullough, I. M., Loftin, C.S. & Sader, S.A. (2012). High-frequency remote monitoring of large lakes with MODIS 500 m imagery. *Rem. Sens. Environ.* **124**, 234-241.
11. Mancino, G., Nolè, A., Urbano, V., Amato, M. & Ferrara, A. (2009). Assessing water quality by remote sensing in small lakes: the case study of Monticchio lakes in southern Italy. *iForest*, **2**, 154-161.
12. Doña, C., Chang, N. B., Caselles, V., Sánchez, J. M., Camacho, A., Delegido, J. & Vannah, B. W. (2015). Integrated satellite data fusion and mining for monitoring lake water quality status of the Albufera de Valencia in Spain. *J. Environ. Manage.* **151**, 416-426.
13. Chang, N. B., Vannah, B. W. & Yang, Y. J. (2014). Comparative sensor fusion between hyperspectral and multispectral satellite sensors for monitoring microcystin distribution in Lake Erie. *IEEE J. Sel. Topics Appl. Earth Observ.* **7**, 2426-2442.
14. Doña, C., Sanchez, J. M., Caselles, V., Dominguez, J. A. & Camacho, A. (2014). Empirical Relationships for Monitoring Water Quality of Lakes and Reservoirs Through Multispectral Images. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **7**, 1632-1641.
15. Vicente, E. & Miracle, M. R. (1992). The coastal Lagoon Albufera de Valencia: an ecosystem under stress. *Limnetica*, **8**, 87-100.
16. Soria, J. M. & Vicente, E. (2002). Estudios de los aportes hídricos al parque natural de la Albufera de Valencia. *Limnetica*, **2**, 105-115.
17. Strickland, J. D. H. & Parsons, T. P. (1968). A practical handbook for seawater analysis, Fisheries Research Board, Ottawa, Canada.
18. Seifert, J. W. Data Mining: An Overview. *CRS Rep. Congr.* 2004.
19. Willmott, C.J. (1982). Some comments on the evaluation of model performance, *Bulletin Amer. Met. Society*, **11**, 1309-1313.

Anexo D

Artículo “Monitoring hydrological patterns of temporary lakes using remote sensing and machine learning models. Case study of La Mancha Húmeda Biosphere Reserve in Central Spain”

Este trabajo se realizó en colaboración con los doctores M. Bisquert, V. García-Santos y S. Imen y se publicó en Remote Sensing, en julio de 2016. El objetivo de este trabajo es el de desarrollar una metodología para evaluar las variaciones en la cubierta de agua de lagunas temporales a través de imágenes de satélite, con el fin de monitorizar la tendencia hidrológica de dichas masas de agua. Se aplicaron diferentes métodos de discriminación de píxeles de agua y tierra. Los resultados se compararon con medidas in situ coincidentes en fecha con el paso del sensor, y la bondad de los distintos métodos se evaluó mediante el coeficiente kappa. Los resultados mostraron el potencial de las técnicas probadas para monitorear los patrones hidrológicos de lagunas temporales en zonas semiáridas.

La revista Remote Sensing posee actualmente un factor de impacto de **3,04** y ocupa la posición relativa **5/28**, dentro de la categoría de Remote Sensing de la base de datos ISI Web of Knowledge.

Article

Monitoring Hydrological Patterns of Temporary Lakes Using Remote Sensing and Machine Learning Models: Case Study of La Mancha Húmeda Biosphere Reserve in Central Spain

Carolina Doña ¹, Ni-Bin Chang ², Vicente Caselles ¹, Juan Manuel Sánchez ³,
Lluís Pérez-Planells ¹, María del Mar Bisquert ³, Vicente García-Santos ¹, Sanaz Imen ²
and Antonio Camacho ^{4,*}

¹ Department of Earth Physics and Thermodynamics, University of Valencia, Burjassot, 46100 Valencia, Spain; Carolina.Dona@uv.es (C.D.); Vicente.Caselles@uv.es (V.C.); Lluís.Perez@uv.es (L.P.-P.); Vicente.García-Santos@uv.es (V.G.-S.)

² Department of Civil, Environmental, and Construction Engineering, University of Central Florida, Orlando, 32816 FL, USA; nchang@ucf.edu (N.-B.C.); imen.snz@gmail.com (S.I.)

³ Department of Applied Physics, University of Castilla-La Mancha, 16071 Cuenca, Spain; JuanManuel.Sanchez@uclm.es (J.M.S.); Mar.Bisquert@uclm.es (M.M.B.)

⁴ Cavanilles Institute of Biodiversity and Evolutionary Biology, University of Valencia, Burjassot, 46100 Valencia, Spain

* Correspondence: antonio.camacho@uv.es; Tel.: +34-9635-43935

Academic Editors: Ricardo Díaz-Delgado, Javier Bustamante, Alfredo R. Huete, Patricia Kandus, Magaly Koch and Prasad S. Thenkabail

Received: 29 February 2016; Accepted: 21 July 2016; Published: 26 July 2016

Abstract: The Biosphere Reserve of La Mancha Húmeda is a wetland-rich area located in central Spain. This reserve comprises a set of temporary lakes, often saline, where water level fluctuates seasonally. Water inflows come mainly from direct precipitation and runoff of small lake watersheds. Most of these lakes lack surface outlets and behave as endorheic systems, where water withdrawal is mainly due to evaporation, causing salt accumulation in the lake beds. Remote sensing was used to estimate the temporal variation of the flooded area in these lakes and their associated hydrological patterns related to the seasonality of precipitation and evapotranspiration. Landsat 7 ETM+ satellite images for the reference period 2013–2015 were jointly used with ground-truth datasets. Several inverse modeling methods, such as two-band and multispectral indices, single-band threshold, classification methods, artificial neural network, support vector machine and genetic programming, were applied to retrieve information on the variation of the flooded areas. Results were compared to ground-truth data, and the classification errors were evaluated by means of the kappa coefficient. Comparative analyses demonstrated that the genetic programming approach yielded the best results, with a kappa value of 0.98 and a total error of omission-commission of 2%. The dependence of the variations in the water-covered area on precipitation and evaporation was also investigated. The results show the potential of the tested techniques to monitor the hydrological patterns of temporary lakes in semiarid areas, which might be useful for management strategy-linked lake conservation and specifically to accomplish the goals of both the European Water Framework Directive and the Habitats Directive.

Keywords: temporary lakes; remote sensing; Landsat 7; genetic programming; multispectral indices; hydrological patterns; water cover detection

1. Introduction

Wetlands contribute a wealth of ecosystem services, including the regulation of the hydrological cycle for flood and drought control, and provide water supply, wildlife refuges, aesthetic enjoyment and recreational opportunities, among others. According to the Ramsar Convention, wetlands, in a broad sense, include all lakes and rivers, underground aquifers, swamps and marshes, wet grasslands, peatlands, oases, estuaries, deltas and tidal flats, mangroves and other coastal areas, coral reefs and all constructed sites, such as fish ponds, rice paddies, reservoirs and salt pans. Although these areas have a critical value to sustainable development, they are detrimentally impacted by urban growth, agricultural land reclamation and derived pollution [1].

The Biosphere Reserve of La Mancha Húmeda is currently the main wetland area in the Iberian Peninsula. This Lake District is one of the wetland complexes most threatened by anthropogenic activity, mainly by groundwater overexploitation due to excessive use for irrigated agriculture. This area is an important refuge for endangered waterfowl species, following the protection criteria for birds in Europe, and also holds endangered habitats [2]. Because of its natural and ethnographic values, it was designated a Biosphere Reserve.

The water balance and hydrological variations are intimately tied to potential changes in a lentic ecosystem. Understanding the dynamics of water in lakes helps the goal of conservation and recovery of these valuable ecosystems [3]. This is especially relevant given several environmental initiatives, such as the European Water Framework Directive (WFD), which came into force in 2000, and the Habitats Directive, delivered in 1992. These directives require each member state in the European Union to achieve a good ecological/conservation status for their water bodies and associated habitats and species, forcing the establishment of conservation actions. Previously, these critical areas in Spain were identified and studied using maps, field campaigns [4] and photo interpretation techniques. To enhance these systematic efforts, remote sensing technologies facilitate the content-based mapping over space and time, leading to multitemporal change detection of the hydrological variations in wetlands. Mapping surface water bodies allows for the investigation of water balance dynamics by providing information on the temporal and spatial variations of surface water coverage, this being especially relevant under the current climate change scenario. In the last few years, some published studies focused on finding a holistic approach to estimate the occurrence of water pixels with several satellite remote sensing imageries. These studies can be classified as based on: (1) single-band threshold; (2) image transformation; (3) two-band spectral indices (vegetation + water indices); (4) two-band spectral water indices; and (5) thematic classification methods.

The earliest related study, published in 1976 [5], identified ponds and lakes using single-band or combinations of bands from Landsat 1 and was followed by additional studies. To delineate open water body features, McFeeters [6] proposed the Normalized Difference Water Index (NDWI), developed using the green and Near Infrared (NIR) bands of the Thematic Mapper sensor (TM) onboard Landsat 4 and 5. Positive values of this index correspond to water pixels in the image. A modified NDWI with the aid of NIR and Short Wave Infrared (SWIR) bands was also proposed by Gao [7] to study liquid water in vegetation; later, Xu [8] proposed an additional modification to the NDWI including green and SWIR bands to study open water features. Several researchers used these indices in conjunction with Landsat, Quickbird or SPOT-4 imageries to estimate the water surface of lakes [9,10], deltas or river basins [11–13], the extents of seasonal and permanent waters in arid areas [14] and coast lines [15]. Extended studies covered different indices with the aid of several satellite imageries, such as Landsat 7-ETM+ (Enhanced Thematic Mapper Plus), SPOT (Satellites Pour l' Observation de la Terre or Earth-observing Satellites), ASTER (Advanced Spaceborne Thermal Emission and Reflection Radiometer) and MODIS (Moderate-resolution Imaging Spectroradiometer) [16].

Several comparisons among indices designed to estimate water pixels from remote sensing images have been recently published, including those comparing the NDWIs [14], other using density slicing of single bands and the Tasseled Cap Wetness (TCW) [17,18], as well as comparisons using a variants of the Principal Component Analysis (PCA) to estimate open water extent with TERRA/ASTER and

Landsat imagery [19]. A Linear Discriminant Analysis Water Index (LDAWI) composed of the green, red, NIR and SWIR bands of SPOT-5 images to map lake surface [20] has also been used to compare in water indices. Furthermore, classic thematic classification methods have been used, such as the maximum likelihood method used to map surface water in a wetland [21,22], the NIR single-band method, a pixel-based method, the object-based segmentation band [23] and the Automated Water Extraction Index (AWEI) [24]. Furthermore, several researchers applied single-band thresholds, such as the NIR band based on the Advanced Very High Resolution Radiometer (AVHRR) and MODIS sensors, to map seasonal inland waters in Central Asia [25] and to estimate the water surface of Sub-Saharan West-African wetlands [26], respectively.

Some methods applied to wetland monitoring combined the NIR band to estimate the water pixels and the NDVI to study wetland vegetation [26], as well as the spectral analytical process and Band 5 of Thematic Mapper [27]. The objectives of our study were to: (1) investigate the applicability of some of the previously-tested methods to estimate the water pixels in our study area and to compare them with the genetic algorithm approach; (2) develop a systematic methodology to monitor the changes in the water cover in several temporary lakes located in the Biosphere Reserve of La Mancha Húmeda (Spain) using Landsat 7-ETM+ imagery; and (3) investigate the interplay between precipitation, evaporation and the subsequent changes of water-covered area in these lakes to unveil the seasonal effect. Our findings elucidate the hydrological patterns of these lakes and provide tools to improve the management policies for lake protection, conservation and ecosystem recovery under the European Water Framework Directive and the Habitats Directive in semiarid areas heavily experiencing the effects of climate change.

2. Materials and Methods

2.1. Study Area

This study is focused on a host of lakes located in the Biosphere Reserve of La Mancha Húmeda, a wetland-rich region, the largest in the Iberian Peninsula, comprising up to 30,000 ha holding wetlands and lakes [28] distributed within the provinces of Albacete, Ciudad Real, Cuenca, and Toledo, in the Castilla-La Mancha region (Central Spain) (Figure 1). In 1981, UNESCO designated the area as the Biosphere Reserve of La Mancha Húmeda within the Man and Biosphere Programme (MAB), a scientific program to promote improved relationships between people and their environments. In 2014, this Biosphere Reserve was recognized as one of the largest wetland areas in Europe, and many of its wetlands are included in the Natura 2000 Network (European Habitats and Birds Directives) and the Convention on Wetlands of International Importance, called the RAMSAR Convention. Located in a very flat area (La Mancha), the Biosphere Reserve includes floodplain lagoons and, particularly a variety of saline lakes, mostly endorheic, and represents one of the main saline lake districts in Europe.

A set of 13 shallow saline lakes were selected for this study (Table 1), all located between 638 and 690 m above sea level. Most are temporary lakes located within the Záncara and Cigüela River Basins in agricultural landscapes. They have small watersheds, mostly covered by vineyards and cereal crops. Most of the lakes have a marginal belt of homophilous plants, and some also have helophytic vegetation in areas where the water is influenced by the discharges of treated wastewater, and thus, salinity is lower in these lakes. The main water inflows to these lakes come from direct precipitation, runoff of small basins, groundwater recharge in some cases and even include some cases of treated wastewater spills from nearby towns. Most of the lakes lack surface outlets and behave as endorheic systems with evaporation as the main water withdrawal process, causing salt accumulation in the lake beds; thus, these lakes range from mesosaline to hypersaline. Their hydroperiods fluctuate, though these lakes are mostly temporary as a result of the mixed Mediterranean-Continental semiarid climate patterns [28], which can be illustrated by the flooding pattern of Lake Alcahozo over a specific time horizon (Figure 2). Climate in this area shows a pronounced dry season with average annual rainfall

of 400–500 mm [29]. Sediments of the studied area are of continental nature; the terrain is flat; and the dominant lithology is limestone [2].

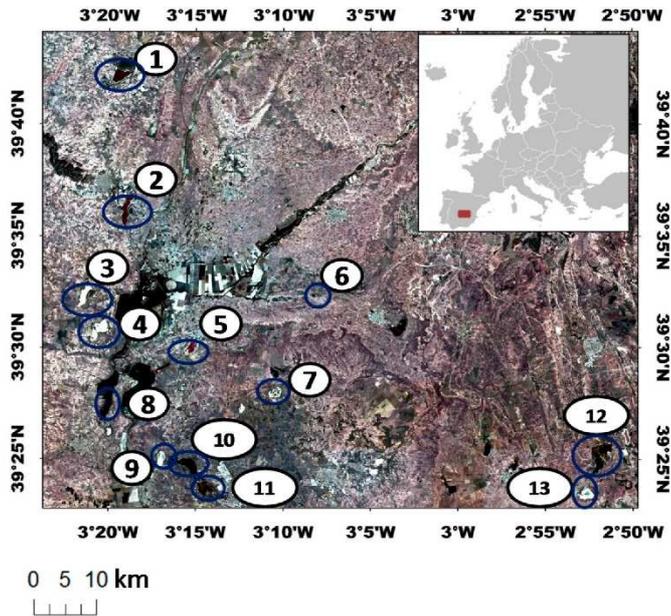


Figure 1. Landsat 5 true color image (RGB: 3, 2 and 1) of the study area. The studied lakes are: (1) El Longar; (2) Larga de Villacañas; (3) Tírez; (4) Peñahueca; (5) Grande de Quero; (6) Mermejuela; (7) Salicor; (8) Grande de Villafranca; (9) Las Yeguas; (10) Camino de Villafranca; (11) La Veguilla; (12) Manjavacas; and (13) Alcahozo.

Table 1. Main features of the study lakes.

Lake ^{*1}	Coordinates °N	Coordinates °W	Altitude (masl)	Max Area ^{*2} (km ²)	Seasonality
Alcahozo (13)	39.38	2.85	660	0.19	High
Camino de Villafranca (10)	39.41	3.26	638	1.4	Moderate
Grande de Quero (5)	39.50	3.24	664	0.9	High
Grande de Villafranca (8)	39.46	3.33	645	0.9	Low
La Veguilla (11)	39.39	3.24	638	0.19	Low+
Larga de Villacañas (2)	39.62	3.34	660	1.4	Low+
Las Yeguas (9)	39.41	3.28	638	0.6	High
El Longar (1)	39.70	3.32	690	0.3	High
Manjavacas (12)	39.41	2.86	670	1.4	Moderate+
Mermejuela (6)	39.54	3.13	660	0.09	High
Peñahueca (4)	39.51	3.34	650	1.6	High
Salicor (7)	39.47	3.17	668	0.6	High
Tírez (3)	39.54	3.36	650	1.3	High

^{*1} In brackets, the lake identification number (from Figure 1); ^{*2} source: Guadiana River Basin Administration (<http://www.chguadiana.es/>); +, these lakes receive treated wastewater inputs that artificially increase the water supply.

Lakes in La Mancha Húmeda have rapidly deteriorated because of the alteration of their hydrological patterns and pollution, the former mainly linked to the continuous overexploitation of the aquifers within the last 50 years due to increased agricultural demand [30]. This water extraction caused an unsustainable draft of the aquifers. Many of the lakes within the Biosphere Reserve were drained-out during the 20th century. Others, because of the salinity and the impossibility for agricultural use, were used for landfills or wastewater disposal by the nearby towns.



Figure 2. Time course of Landsat 7-ETM+ true color images (RGB: Bands 3, 2, 1) of Lake Alcahozo for the period June 2013–June 2014.

2.2. Reference Datasets

In addition to the Landsat imagery, three different datasets were used in this study (Table 2): (1) ground-measurements of the water depth in the 13 study lakes, used to test several methods to estimate water pixels coverage and to select the method showing the best performance; (2) the perimeter of two of these lakes (Alcahozo and Manjavacas) measured in situ by GPS; and (3) high spatial resolution Google Earth™ images. The second and third datasets were used to validate the area estimated by the selected method. Boundaries of the lakes obtained from high spatial resolution Google Earth™ images were digitized manually. The water column measurements were collected in the field during 2013–2014 as an integral part of the ECOLAKE project (ecological patterns in plateau lakes: the key for their conservation). Concurrent, or close in time, Landsat overpasses are listed in Table 2. Water depth was monthly measured by limnimeters installed in each lake during the sampling years. A major field campaign was conducted in July 2014 covering lakes Alcahozo, Camino de Villafranca, El Longar, La Veguilla, Las Yeguas and Manjavacas. Satellite images affected by clouds or with an excessive time delay between the acquisition date and field measurements date were discarded.

Table 2. Data sources for this study for each lake: A = ground-measurements of the water depth, B = GPS data and C = Earth Explorer™ ©Digital globe, respectively (lake abbreviations as follows: Alcahozo, AL; Camino de Villafranca, CA; Grande de Quero, QUE; Grande de Villafranca, GRAN; La Veguilla, VE; Larga de Villacañas, LAR; Las Yeguas, YE; Longar, LO; Manjavacas, MAN; Mermejuela, MER Peñahueca, PE; Salicor, SAL; Tírez, TI).

Landsat Image Date (Day/Month/Year)	Reference Data Date (Day/Month/Year)	Source	Lakes
3 February 2011	9 February 2011	C	LO/YE/CA/VE/LAR/PE/QUE/SAL
29 June 2012	29 June 2012	C	VE/LAR/PE
7 July 2012	15 July 2012	C	CA/SAL
24 February 2013	28 February 2013	A	AL/CA/MAN
31 May 2013	29 May 2013	A	AL/CA/LO/VE/MAN/GRAN/LAR/MER/PE/QUE/TI
25 July 2013	23 July 2013	A	CA/VE/YE/QUE/TI
29 October 2013	28 October 2013	A	PE/TI
23 November 2013	25 November 2013	A	AL/CA/LO/LAR/SAL
16 December 2013	16 December 2013	A	CA/LO/GRAN/LAR/PE/TI
26 January 2014	27 January 2014	A	CA/VE/MAN/YE/PE/TI
5 February 2014	29 April 2014	A	CA/MER
26 June 2014	24 June 2014	A	LO/VE/YE/GRAN/LAR/MER/QUE/SAL
21 July 2014	23 July 2014 *	A	AL/CA/LO/VE/MAN/YE/MER/PE/TI
2 March 2015	22 February 2015	B	AL/MAN
21 May 2015	21 May 2015	B	AL/MAN

* Field inspection work in AL/CA/LO/VE/MAN/YE.

2.3. Remote Sensing Data Collection and Pre-Processing

The ETM+ onboard Landsat 7 platform was used to conduct this study. All of the images are available free of charge at the United States Geological Survey website. The Landsat cloud-free images downloaded from the EarthExplorer visor (<http://earthexplorer.usgs.gov/>) correspond to the surface reflectance product corrected from the atmospheric contribution with the 6S radiative transfer code (CDR_sr). The downloaded scenes were synchronous or close in time to the reference data (Table 2). The path and row corresponding to our study area belong to 200/33 and 201/32–33, respectively.

Satellite image pre-processing was required prior the application of the different methods. All scenes were cut to fit our area of interest, and a normalization of the images was then conducted [31,32] using the image of July 2014 as a reference. The iteratively-reweighted multivariate alteration method (IRMAD) was applied [33] to minimize the spectral variability caused by seasonal sun-surface-sensor effects [31]. The IRMAD technique developed for automatic radiometric normalization of multi-spectral and hyper-spectral images allowed us to find linear combinations between the reference and target image bands used to generate a pair of new multispectral images by using canonical correlation analysis. The components of the new images were called canonical variates. This IRMAD technique considers that the reflectance values of some areas in the same scene acquired at different times would be changed, but not everywhere. With this assumption, pixels with the fewest differences between the canonical variates were labeled as the pseudo-invariant pixels, which were then used to normalize each image band-by-band to the reference image. The linear regression equations used to spectrally align each of the six bands of an image were obtained with regression coefficients (r^2) > 0.90 and root mean square errors (RMSE) <10%. After this pre-processing, all images were visually compared to ensure that they were co-registered correctly. If so, no further corrections or adjustments were necessary. A correction of the scan line corrector failure was performed following the methodology proposed by Scaramuzza et al. [34]. To isolate our study area (lakes), a water mask was manually digitalized from a Landsat 5-TM image (May 2010), corresponding to a wet year, to mark out the maximum flooding area of the lakes.

2.4. Water Mapping Methods

Several methods were tested to estimate the absence or presence of water in the lakebeds of the studied lakes (Figure 3). For a rigorous comparison, we focused on the data from the intensive field campaign in July 2014 and on five tested lakes (Alcahozo, Camino de Villafranca, El Longar, La Veguilla, Las Yeguas and Manjavacas) (Figure 4). The comparison was established between two-band vegetation indices, two-band water indices, single band threshold, classification methods, Artificial Neural Network (ANN), Support Vector Machine (SVM) and Genetic Programming (GP) algorithms. Note that the Landsat image used in this section (21 July 2014) will be later considered as the reference for the normalization procedure applied to the full image dataset. Regarding two-band spectral indices for vegetation, both the NDVI [35] and the SAVI [36] were tested. Within the two-band spectral indices for water, several variants of the NDWI were studied, including the Modification of the Normalized Difference Water Index (MNDWI) and the index proposed by Ángel-Martínez [37] (I_CEDEx) where CEDEx is the Centre for studies and experimentation on public works in Madrid, Spain. The approach proposed by Bustamante et al. [38] based on the MIR band and the single-band threshold with the NIR used by [39] were tested. Both supervised and unsupervised classification methods using different spectral bands were also tested. Unsupervised classification methods include k-means and the Iterative Self-Organizing Data Analysis Technique (ISODATA), whereas supervised classification included the parallelepiped method, minimum and Mahalanobis distance and the maximum likelihood method (Figure 3).

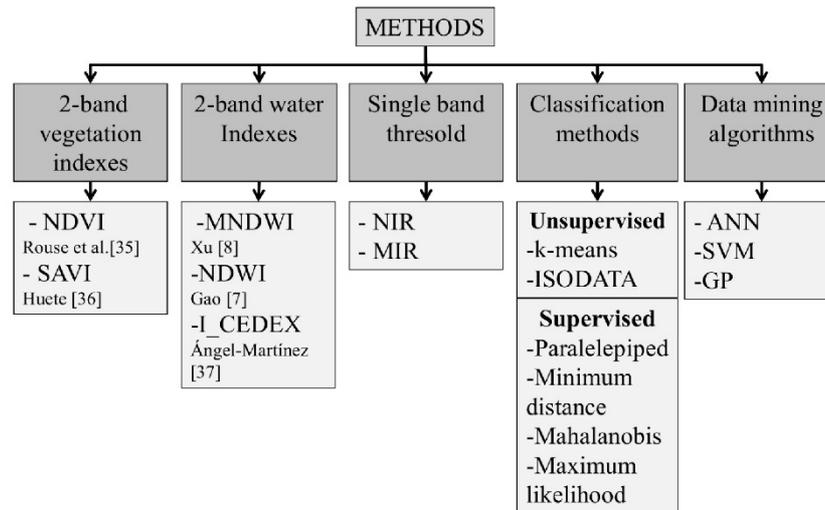


Figure 3. Collection of methods to estimate water pixels tested in this work. ISODATA, Iterative Self-Organizing Data Analysis Technique.

Machine learning algorithms were also considered, including SVM, GP and ANN, a machine-learning method inspired by the human brain function. These machine learning techniques are all nonparametric classification techniques that require no assumptions about the distribution of the data and thus need no a priori knowledge about the characteristics of feature data. An ANN model is based on three different layers: input layer (i.e., input data include reflectance bands and/or water-vegetation indices), one or more hidden layers and the output layer (i.e., dichotomous output includes either water or non-water class) [40]. The SVM is a supervised machine learning technique based on statistical learning theory [41] used to find boundary locations of different classes. In our study, different kinds of SVM were tested to carry out the classification, including linear, quadratic, cubic and Gaussian kernels. MATLAB[®] software was used to implement ANN and SVM. Finally, genetic programming (GP) is a subclass of evolutionary computation techniques designed to search for the best fit to perform a user-defined task. GP can decode system behaviors based on empirical data for symbolic regression, uncover relationships and make inferences using association path analysis, classification, clustering and forecasting [42]. A principal advantage of GP is that the solution methodology can learn the relationship between the inputs and outputs without any a priori knowledge or preconceptions, thus placing the burden of the discovery process primarily on the GP, reducing data contribution and preprocessing by the user [43]. In this study, the user-defined task was to develop a GP model that uses the inputs of surface reflectance data associated with common bands and different two-band indices to predict the outputs, including water and non-water categories. We used Discipulus[®] [44] software to run the GP algorithms.

The ground-truth dataset was divided in two subsets for ANN, SVM and GP model training (2/3 of the total data; 56 data points) and validation (1/3 of total data; 28 data points). Finally, the ANN, SVM and GP algorithms yielding the best results were compared to the methods above. A flowchart of the different steps generated in the methodology (Figure 4) shows that all of the pixel values used for the training/validation/testing of all methods were extracted before the correction of the Scan Line Corrector (SLC) failure of Landsat 7; thus, this correction did not cause any bias. A confusion matrix was obtained for each of the methods applied to rank all of the possible cases associated with these models in different categories, estimating whether or not the predicted value is consistent with the real value. To evaluate the consistency of our results, we used the confusion matrix to calculate the kappa coefficient (κ), an index ranging from -1 to $+1$; values higher than 0.4 are considered

acceptable [45]. The confusion matrix can also be used to generate the commission error and user accuracy. The commission error is the percentage of pixels wrongly assigned to a certain class by the classifier, whereas user accuracy is the probability that a pixel assigned to a class by the classifier correctly corresponds to that class. The omission error is the percentage of pixels that belong to the ground truth class, but were improperly classified, and the producer accuracy is the probability that the classifier has been correctly assigned to a class given by the ground truth data.

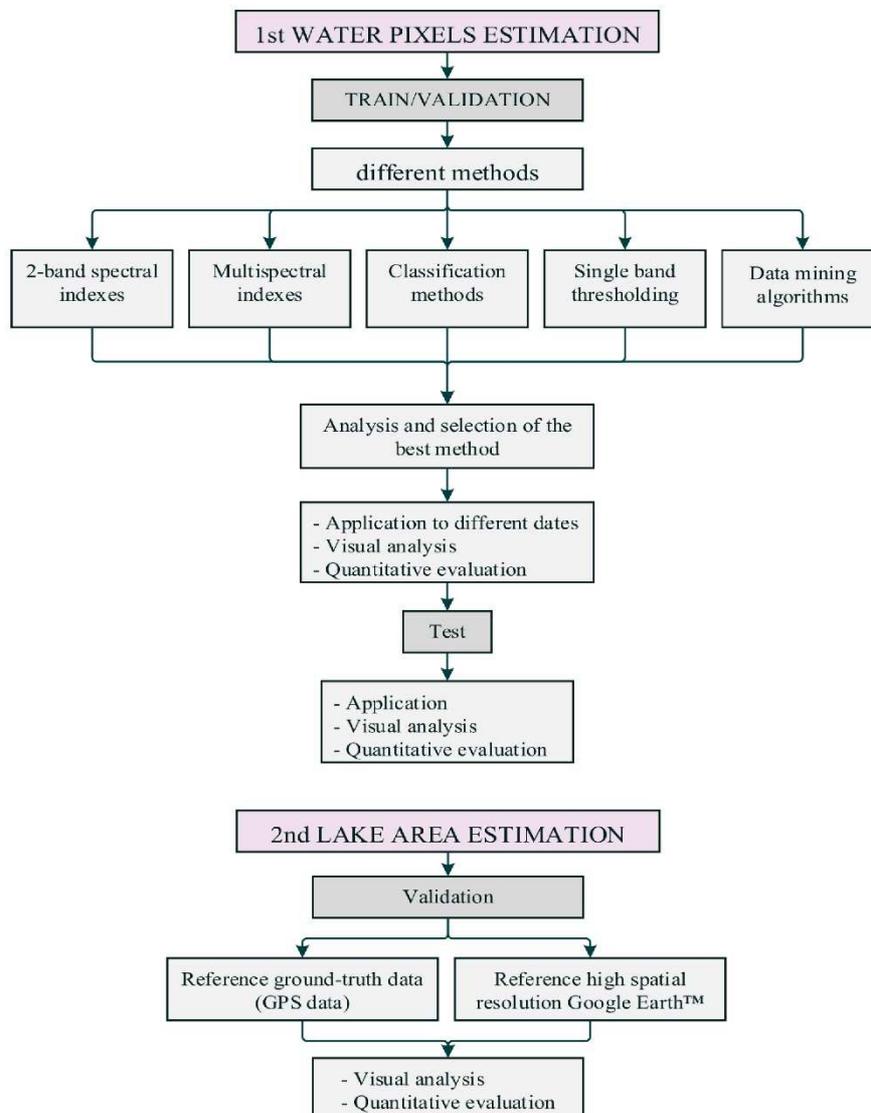


Figure 4. Flow chart showing the different steps followed in the methodology to determine the most appropriate water estimation method. (Lake abbreviations: Alcahozo, AL; Camino de Villafranca, CA; Grande de Quero, QUE; Grande de Villafranca, GRAN; La Veguilla, VE; Larga de Villacañas, LAR; Las Yeguas, YE; Longar, LO; Manjavacas, MAN; Mermejuela, MER Peñahueca, PE; Salicor, SAL; Tirez, TI).

The method showing the best performance in terms of discrimination of water/non-water pixels in our test area was selected and applied to the full set of images listed in Table 2 and to a variety of lakes different from those used in the algorithm training. Finally, the relationship between precipitation–evaporation and the lake water cover was analyzed for Lake Alcahozo, used as a model lake to explore the effect of meteorological variability characteristic of the Mediterranean climate during 2013–2015. Meteorological data were downloaded from the Servicio Integral de Asesoramiento al Regante de Castilla-La Mancha (SIAR) (<http://crea.uclm.es/siar/datmeteo/>), which provides daily information of mean, absolute maximum and minimum values of temperature, humidity, wind speed, cumulative global solar radiation, daylight hours, precipitation and reference evapotranspiration estimated with the Penman–Monteith equation [46]. In the present study, we focused on precipitation and reference evapotranspiration data to explore the hydrological cycle dynamics.

Finally, we include a brief discussion of climate elasticity as an application example. Climate elasticity of streamflow (e), an index commonly used to quantify the sensitivity of streamflow to meteorological pattern and climate change. It is defined as the proportional change in the streamflow (lake flooded area in our case) relative to the proportional change in a climatic variable, such as precipitation (p). The index was applied to our lake data, and the nonparametric estimator was used to calculate the climate elasticity [47]. In particular, we focused on the variation between the area of a lake and the precipitation Equation (1).

$$e_p = \text{median} \left(\frac{A_t - \bar{A}}{P_t - \bar{P}} \times \frac{\bar{P}}{\bar{A}} \right) \quad (1)$$

where e_p is the climate elasticity relative to precipitation, A and P represent area and precipitation, \bar{A} and \bar{P} are the corresponding yearly mean values and t is time.

3. Results and Discussion

3.1. Water Pixel Estimation

Results of the accuracy assessment of the different methods are summarized in Tables 3–5. Among the water indices, the NDWI based on the green and NIR bands [6] (Table 3) showed the best performance. This index showed the best results with the threshold proposed by the original author for to assign water pixels (values > 0). An overall accuracy of 90% and a kappa coefficient of 0.80 were obtained. In the case of the MNDWI and the NDWI [7], the best results were obtained with a threshold of 0.5 and -2 for the I_CEDDEX (Table 3).

Kappa values of 0.4 were obtained for the methods based on vegetation indices (NDVI and SAVI) (Table 3). Visualization analysis based on field conditions showed that both water and vegetation indices overestimated the water pixels when the lakes were completely dry in July 2014.

Table 3. Summary of accuracy assessment of the different water and vegetation indices and the single bands tested.

Method		Overall Accuracy (%)	Kappa	Commission Error (%)	Omission Error (%)	Producer Accuracy (%)	User Accuracy (%)	Total Error (%)
MNDWI	Xu [8]	90	0.80	30	3	95	70	30
NDWI	Gao [7]	90	0.60	40	15	85	60	60
NDWI	McFeeters [6]	90	0.80	25	3	95	75	30
I_CEDDEX	Ángel-Martínez [37]	90	0.70	40	10	90	60	50
NDVI	Rouse et al. [35]	70	0.40	60	30	70	40	90
SAVI	Huete [36]	70	0.40	60	30	70	40	90
ETM+4		99	0.97	2	2	98	98	4
ETM+5		90	0.70	40	0	100	60	40

Acceptable results were obtained for the single-band threshold, with kappa values of 0.97 and 0.70 for Bands 4 and 5 (Figure 5), although Band 5 overestimated the water pixels, possibly due to the presence of puddles or mud puddle areas on the lakes surfaces. The water index using near and middle infrared ETM+ bands (NDWI [7]) presented worse results than the other water indexes applied to our study area, possibly because these wavelengths are more sensitive to the presence water [38], thus causing mistakes in the water-cover pixel identification. These results agree with the conclusions of Campos et al. [14] reporting that indices using the middle infrared band can have contrasting performances in detecting seasonal and permanent water. The NDWI using near and middle infrared bands was more sensible for detecting the seasonality of the water cover despite significant overestimations. Although the kappa values were acceptable, the commission and omission errors were high, and the total error (percentage of omission and commission) was greater than 30% in all cases, except for ETM+4, where the total error was 4% (Table 3).

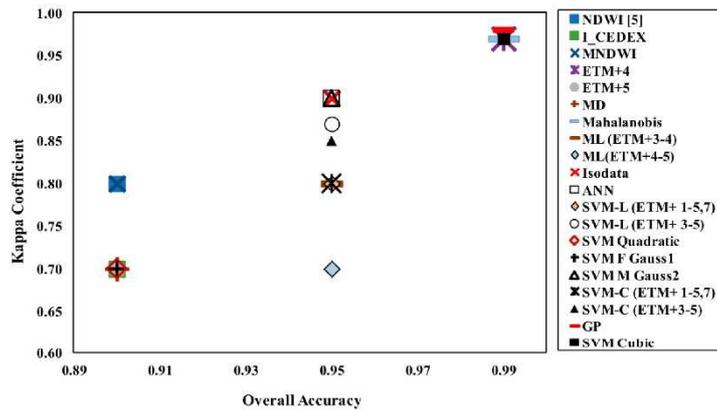


Figure 5. Kappa coefficient vs. overall accuracy for the methods showing acceptable kappa values: water-vegetation indices and single bands, classification and data mining methods.

The best results for the supervised classification based on the Mahalanobis distance and maximum likelihood methods (Table 4) were obtained for the combination of red and NIR bands, resulting in kappa values of 0.97 and 0.80, respectively. Overall, the classification methods showed acceptable kappa values (Figure 5), but with an underestimation of water pixels in the classification since one of the lakes with water in the ground-data test date (July 2014) was classified as a dry lake. Acceptable results were obtained from the ISODATA unsupervised classification method, with an overall accuracy of 95% and a kappa value of 0.90. A total error of 15% was obtained from the ISODATA method.

Table 4. Summary of the accuracy assessment for the different classification methods and band combinations tested in this study.

Method		Overall Accuracy (%)	Kappa	Commission Error (%)	Omission Error (%)	Producer Accuracy (%)	User Accuracy (%)	Total Error (%)
Parallelepiped	All combinations	<40	<0.20	>50	>50	<50	<50	>60
Minimum distance	ETM+ 4,5	90	0.67	40	0	100	60	60
Mahalanobis	ETM+ 3,4	99	0.97	5	0	100	95	5
Maximum Likelihood	ETM+ 3,4	95	0.80	0	30	70	100	30
Maximum Likelihood	ETM+ 4,5	95	0.70	0	40	60	100	40
K-means	ETM+ 4,5	80	0.60	10	30	90	75	40
K-means	ETM+ 4	80	0.60	10	30	90	75	40
ISODATA	ETM+ 3,4	95	0.90	10	5	100	80	15

Data mining methods applied in our study showed overall accuracy values greater than 90% and kappa values higher than 0.70 (Figure 5). Results from the combinations of Bands 3–5 as inputs in both ANN and SVM methods were acceptable (Table 5), as they were in some cases using all bands as inputs. An ANN model with four hidden layers presented the best results, with an overall accuracy of 95% and a kappa value of 0.9. Similar results were obtained when applying the SVM with all reflectance bands as inputs. The best results were obtained from the SVM method when using the linear and cubic kernel. In particular, the cubic kernel-based SVM showed the best results with an overall accuracy of 99% and a kappa value close to one.

In contrast to the ANN and SVM methods, the GP algorithm produces a complex relationship between the inputs and outputs via a nonlinear function. The best results were obtained by a relationship between the water pixels and the reflectance of the NIR band from the Landsat ETM+ imagery. The resulting nonlinear equation was:

$$GPRw = \cos ((0.003 \times ETM + 4)^{1/4}) \quad (2)$$

where GPRw is the GP outputs for water classification.

This GP algorithm presented a kappa value of 0.98 and an error of 2% in the classification of the water pixels (Table 5). With such a relationship between the reflectance values and the class given by Equation (2), a threshold could be established to determine whether there is water-cover or not in each pixel. In this study, values of GPRw higher than 0.1 imply that the classification as water pixels is favored.

Table 5. Summary of the accuracy assessment of the different data mining methods tested.

Method		Overall Accuracy (%)	Kappa	Commission Error (%)	Omission Error (%)	Producer Accuracy (%)	User Accuracy (%)	Total Error (%)
ANN	ETM+ 1–5,7	95	0.90	4	15	85	95	19
SVM linear	ETM+ 1–5,7	95	0.80	0	15	100	85	15
SVM linear	ETM+ 3–5	95	0.87	0	18	100	80	18
SVM Quadratic	ETM+ 3–5	90	0.70	0	40	100	65	40
SVM Cubic	ETM+ 1–5,7	99	0.97	0	6	100	94	6
SVM F Gauss ¹	ETM+ 3–5	90	0.70	0	40	100	59	40
SVM M Gauss ²	ETM+ 3–5	95	0.90	0	20	100	80	20
SVM Course	ETM+ 1–5,7	95	0.80	0	30	100	75	30
SVM Course	ETM+ 3–5	95	0.85	0	20	100	79	20
GP	ETM+ 4	99	0.98	2	0	100	97	2

Based on its accuracy assessment compared to other methods, the GP algorithm was selected as the most appropriate method (Table 5). As explained in the Methods section (Figure 4), this, as the best method, was selected to be applied to a set of images from different dates. First, it was applied for the same lakes for which the algorithm was performed and then applied to the rest of the lakes. When the results of the GP algorithm were applied to the full image dataset (Table 2), an overall accuracy of 95% with a kappa value of 0.8 was achieved. When applied to lakes different from those used in the algorithm training, an overall accuracy of 90% and a kappa value of 0.8 were obtained. We could further discern between two different thresholds to separate water and non-water pixels, one for each type of lake. The lakes in our study area are lakes with wide playas (Figure 6), lakes with smaller playa areas and lakes without playas. Playa lakes showed a different threshold than those for lakes without playas or with small playa areas (Figure 7). For the small playa area lakes, however, we observed that some have a threshold of -0.05 , similar to playa lakes, although most of the playa lakes show a 0.1 threshold value. The result of the GP algorithm after applying the test data can be seen in Figure 7. Lakes without playas presented GPRw values higher than 0.1 for pixels with the presence of water

(Class 1) and lower than 0.1 for the dry pixels (Figure 7). Instead, for playa lakes the threshold is -0.05 , so lower values presented pixels without water.

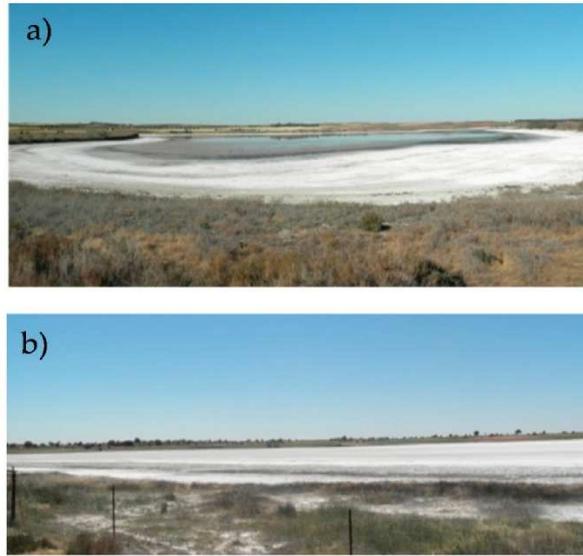


Figure 6. Playa lakes: (a) Camino de Villafranca and (b) Alcahozo.

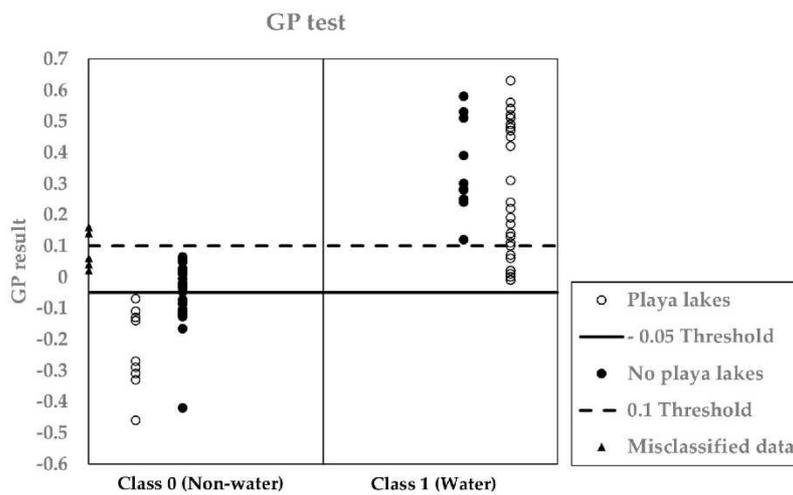


Figure 7. Different thresholds obtained after applying the GP algorithm to the test data. Class 0 belongs to dry pixels and Class 1 to water pixels. Data for playa lakes (white circles) are differentiated from those of lakes having small or no playas (black circles).

3.2. Sub-Pixel Extraction

The selected GP algorithm was applied to the synchronous images (or at least close to the date of the field campaigns) and to the high resolution Google Earth™ images (Table 2). The best results in the classification of water pixels were obtained for lakes Grande de Quero, Larga de Villacañas, Las Yeguas, Manjavacas and Peñahueca with kappa values higher than 0.60 and overall accuracies

above 80% (Table 6). We were unable to find close reference data to carry out this part of the study for lakes Grande de Villafranca and Mermejuela.

Table 6. Accuracy in the estimation of water area for the different lakes with reference data.

Lake ^{*1}	Threshold	Overall Accuracy (%)	Kappa
Alcahozo (13)	−0.05	70	0.45
Camino de Villafranca (10)	−0.05	94	0.62
Grande de Quero (5)	−0.05	99	0.85
* La Veguilla (11)	0.10	86	0.68
Larga de Villacañas (2)	0.10	90	0.61
* Larga de Villacañas (2)	0.10	80	0.80
Las Yeguas (9)	−0.05	97	0.76
El Longar (1)	−0.05	96	0.53
Manjavacas (12)	0.10	96	0.92
Manjavacas (12)	0.10	86	0.73
* Peñahueca (4)	−0.05	95	0.82
Salicor (7)	−0.05	94	0.15

* Synchronous reference and satellite data (same day); ^{*1} lake identification number listed in each bracket as in Figure 1.

Lakes Alcahozo, Salicor and Longar presented the lowest kappa values of 0.45, 0.15 and 0.53, respectively. The discrepancies in Lake Alcahozo were probably due to the one-week difference between the image acquisition date and the collection date of the reference data in the field. This lake experiences fast changes in water depth because of its low depth and the absence of regular water inflows, presenting challenges to the classification of these edge pixels. The classification output of the GP algorithm is, however, in agreement with the ground-truth data when the lake was completely dry for the 15 May 2011 campaign (Figure 8).

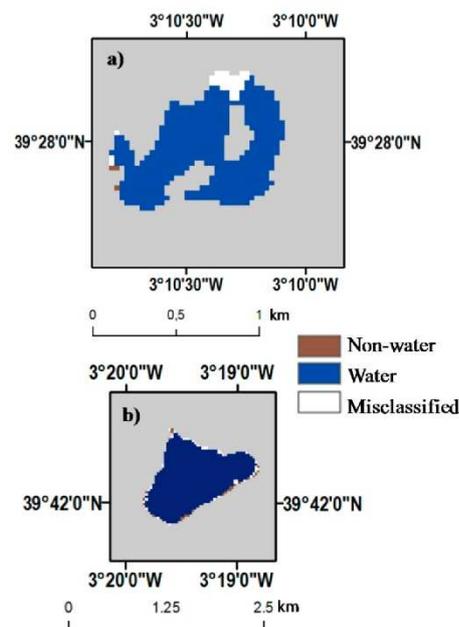


Figure 8. Results of the classification from the GP algorithm for (a) Salicor and (b) El Longar for the reference date 9 February 2011 and Landsat image of 3 February 2011.

It is noteworthy that the misclassified area of Lake Salicor is flagged by a kappa value of 0.15 (Figure 8), in part because SLC failure created strips in the original ETM+ images, which affected the observation of the lake. Most of the misclassified data belong to Lakes Salicor and El Longar (Figure 8). On the other hand, the proposed GP algorithm could not be validated for lakes Grande de Villafranca, Mermejuela and Tírez because no ground data were available, and the auxiliary Google Earth™ images were distant in time. However, the application of the GP algorithm to a Landsat image of the Lake Manjavacas obtained on 21 May 2015, against the GPS ground measurements of the water perimeter on the same date, confirmed that the threshold value of 0.1 is applicable (Figure 9). In this case, an overall accuracy of 95% and a kappa value of 0.90 were obtained.

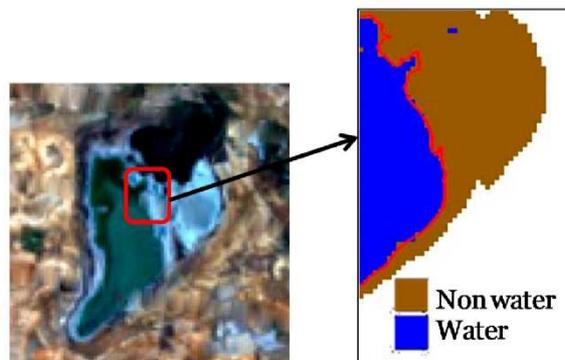


Figure 9. Results of the classification using the GP algorithm for Lake Manjavacas for 21 May 2015 (0.1 threshold). The red line represents the GPS track of the water limit, measured in situ.

3.3. Water Area Variation

The selected GP algorithm was applied to the available images to estimate the flooded area of all studied lakes from spring 2013 to spring 2015. Focusing on Lake Alcahozo, we further present a linkage between climate (precipitation and reference evapotranspiration (ET_0)) and the flooded area of the lake to elucidate the link between meteorology and lake hydrology at the local scale. The reference evapotranspiration (ET_0) was used as an indicator of the behavior and intensity of the evaporation in our study area. The results of the flooded area estimation for Lake Alcahozo, which is completely dry in summer periods following the seasonal trend typical for this type of lake, are shown in Figure 10. Figure 11 shows that the flooded area was mainly driven by the precipitation pattern, and larger flooded areas were obtained for wetter years. In May 2013, the flooded area presented a maximum value, because lake water remaining from winter was further increased by the March rainfall peak. In addition, this lake presents a dry season in summer, corresponding to the highest ET_0 values (Figure 10).

The proportional change in the flooded area with the precipitation was calculated for our study time frame for lakes that presented the best estimations of flooded area. After calculating the climate elasticity, two different values were obtained, around 0.2 and 1.3, in association with a low or high temporality, respectively. All lakes with a strong seasonal trend, such as Lakes Alcahozo, Camino de Villafranca, Grande de Quero and Las Yeguas, presented values of 1.3, indicating that in these two years, a 1% change in mean annual rainfall resulted in a 1.3% change in the mean annual flooded area of these lakes, whereas in the other cases (Lakes Larga de Villacañas, La Veguilla and Manjavacas), a similar change in annual rainfall only resulted in a variation of 0.2% of the mean annual flooded area. This finding is a consequence of the hydrological alteration driven by the wastewater disposal in the second set of lakes, which expanded the flooding period and buffered rainfall-dependent fluctuations in water level.

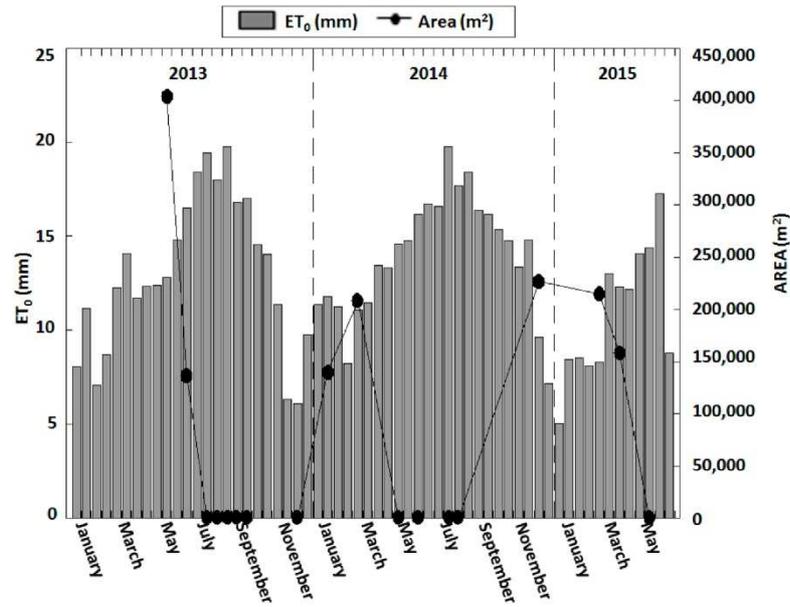


Figure 10. The flooded area in relation to reference evapotranspiration (ET_0) for the period 2013–2015 in Lake Alcahozo.

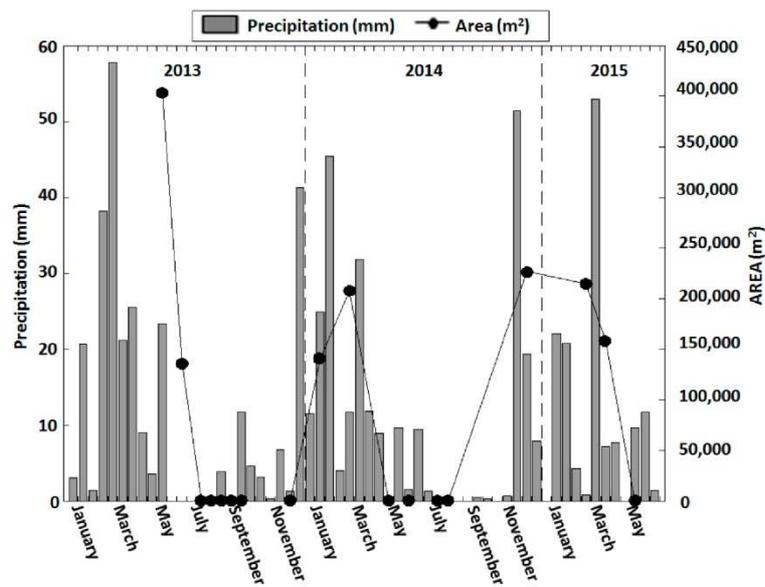


Figure 11. The flooded area in relation to precipitation for the period 2013–2015 in Lake Alcahozo.

Although the algorithm was tested for different lakes from our study area and the seasonal effects were minimized with the radiometric normalization, the chemical and biophysical changes in the water bodies could affect the results of the GP algorithm. We tried to minimize these factors by applying the GP algorithm to different seasons and different water levels of each lake. Although the

GP algorithm showed acceptable results, the robustness of the described methodology will need to be further assessed using additional datasets. In addition, the SLC failure of Landsat 7 imageries caused errors in the estimation of the total water area of some lakes. Water bodies presenting irregular shapes could be more vulnerable to this issue, such as the case of Salicor Lake in the present study.

The proposed GP algorithm can also be applied to images obtained from other sensors, such as Landsat 5, which will allow us to conduct long-term studies. Future works will deal with the applications to other sensors, such as SPOT-5 or the recent Sentinel 2. The future NASA satellite “Surface Water and Ocean Topographic (SWOT)”, which is expected to be launched in 2020 for a demonstration mission with an expected length of three years, is a good example of the importance of these types of studies. In addition, future prospects of our current work will be the long-term monitoring of the lakes to study their hydrological patterns in greater detail to help to develop adequate management plans for lake protection, conservation and recovery under the European Water Framework and the Habitats Directives, especially for semiarid regions, such as the Mediterranean basin.

4. Conclusions

Several supervised and unsupervised classification methods to estimate water pixels across a set of 13 shallow saline lakes in Spain were tested and validated in this study, using Landsat 7-ETM+ imageries. These methods tested include the two-band spectral indices, multispectral indices, classification methods, single threshold and the data mining or machine learning techniques. The GP algorithm showed the best performance with a kappa value of 0.98 and a low error of 2%, both in training and validation. The robustness of this GP method was reinforced with additional ground-truth data. This GP algorithm could serve as a tool to improve our understanding of the temporal trends in seasonal lakes and their dependence on meteorological patterns, especially on rainfall. Such an approach provides a cost-effective way to monitor the flooded area variations related to changes of precipitation and evaporation, as well as to resource exploitation, in order to establish sustainable water resources management plans that preserve the ecological health of wetlands and lakes in semiarid basins with a high water stress, such as the area of La Mancha Húmeda Biosphere Reserve.

Acknowledgments: This study was jointly supported by the Spanish Ministry of Economy and Competitiveness and the European Regional Development Fund (FEDER) “One way to make Europe”, through the projects: “CLIMAWET” CGL2015-69557-R (MINECO/FEDER,UE), “ECOLAKE” (CGL2012-38909) and “ESCLITEL” (CGL2013-46862-C2-1/2-P), the first two to A.C., by the Regional Government of Valencia through the project PROMETEOII/2014/086 and postdoctoral contract of V. García-Santos (“VALi+D” APOSTD /2015/033). The authors would like to thank García-Haro and S. Sánchez-Ruiz for their comments. Finally, the authors want to thank the valuable and detailed reviewers’ comments.

Author Contributions: C.D., N.-B.C. and A. C. conceived and designed the experiments; C.D. performed the experiments; C.D., N.-B.C. and A.C analyzed the data; V.C., J.M.S., L.P.-P., V.G.-S., M.M.B. and S.I. contributed reagents/materials/analysis tools and field campaigns. C.D., N.-B.C., J.M.S. and A. C. wrote the paper.

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

References

1. Florín, M.; Montes, C.; Rueda, F. Origin, hydrologic functioning, and morphometric characteristics of small, shallow, semiarid lakes (lagunas) in La Mancha, central Spain. *Wetlands* **1993**, *13*, 247–259. [[CrossRef](#)]
2. Gosálvez, R.; Gil-Delgado, J.A.; Vives-Ferrándiz, C.; Sánchez, G.; Florín, M. Seguimiento de aves acuáticas amenazadas en lagunas de la Reserva de la Biosfera de La Mancha Húmeda (España central). *Polígonos* **2012**, *22*, 89–122. [[CrossRef](#)]
3. Castañeda, C.; Herrero, J. Teledección de cambios en la Laguna de Gallocañta. *Memorias Real Soc. Española Hist. Nat.* **2009**, *7*, 103–126.
4. Camacho, A.; Miracle, M.R.; Vicente, E. Which factors determine the abundance and distribution of picocyanobacteria in inland waters? A comparison among different types of lakes and ponds. *Arch. Hydrobiol.* **2003**, *157*, 321–338. [[CrossRef](#)]

5. Work, E.A.; Gilmer, D.S. Utilization of satellite data for inventorying prairie ponds and lakes. *J. Photogramm. Eng. Remote Sens.* **1976**, *42*, 685–694.
6. McFeeters, S.K. The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. *Int. J. Remote Sens.* **1996**, *17*, 1425–1432. [[CrossRef](#)]
7. Gao, B.C. NDWI—A normalized difference water index for remote sensing of vegetation liquid water from space. *Remote Sens. Environ.* **1996**, *58*, 257–266. [[CrossRef](#)]
8. Xu, H. Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery. *Int. J. Remote Sens.* **2006**, *27*, 3025–3033. [[CrossRef](#)]
9. Bai, J.; Chen, X.; Li, J.; Yang, L.; Fang, H. Changes in the area of inland lakes in arid regions of central Asia during the past 30 years. *Environ. Monit. Assess.* **2011**, *178*, 247–256. [[CrossRef](#)] [[PubMed](#)]
10. 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. *Remote Sens.* **2013**, *5*, 3544–3561. [[CrossRef](#)]
11. Sakamoto, T.; Van Nguyen, N.; Kotera, A.; Ohno, H.; Ishitsuka, N.; Yokozawa, M. Detecting temporal changes in the extent of annual flooding within the Cambodia and the Vietnamese Mekong Delta from MODIS time-series imagery. *Remote Sens. Environ.* **2007**, *109*, 295–313. [[CrossRef](#)]
12. Li, W.; Du, Z.; Ling, F.; Zhou, D.; Wang, H.; Gui, Y.; Sun, B.; Zhang, X. A comparison of land surface water mapping using the normalized difference water index from TM, ETM+ and ALI. *Remote Sens.* **2013**, *5*, 5530–5549. [[CrossRef](#)]
13. El-Asmar, H.M.; Hereher, M.E. Change detection of the coastal zone east of the Nile Delta using remote sensing. *Environ. Earth Sci.* **2011**, *62*, 769–777. [[CrossRef](#)]
14. Campos, J.C.; Sillero, N.; Brito, J.C. Normalized difference water indexes have dissimilar performances in detecting seasonal and permanent water in the Sahara-Sahel transition zone. *J. Hydrol.* **2012**, *464*–*465*, 438–446. [[CrossRef](#)]
15. Maglione, P. Coastline extraction using high resolution WorldView-2 satellite imagery. *Eur. J. Remote Sens.* **2014**, 685–699. [[CrossRef](#)]
16. Ji, L.; Zhang, L.; Wylie, B. Analysis of dynamic thresholds for the normalized difference water index. *Photogramm. Eng. Remote Sens.* **2009**, *75*, 1307–1317. [[CrossRef](#)]
17. Jain, S.K.; Singh, R.D.; Jain, M.K.; Lohani, A.K. Delineation of flood-prone areas using remote sensing techniques. *Water Resour. Manag.* **2005**, *19*, 333–347. [[CrossRef](#)]
18. Ouma, Y.O.; Tateishi, R. A water index for rapid mapping of shoreline changes of five East African Rift Valley lakes: An empirical analysis using Landsat TM and ETM+ data. *Int. J. Remote Sens.* **2006**, *27*, 3153–3181. [[CrossRef](#)]
19. Lira, J. Segmentation and morphology of open water bodies from multispectral images. *Int. J. Remote Sens.* **2006**, *27*, 4015–4038. [[CrossRef](#)]
20. Fisher, A.; Danaher, T. A water index for SPOT5 HRG satellite imagery, New South Wales, Australia, determined by linear discriminant analysis. *Remote Sens.* **2013**, *5*, 5907–5925. [[CrossRef](#)]
21. Gardelle, J.; Hiernaux, P.; Kergoat, L.; Grippa, M. Less rain, more water in ponds: A remote sensing study of the dynamics of surface waters from 1950 to present in pastoral Sahel (Gourma region, Mali). *Hydrol. Earth Syst. Sci. Discuss.* **2009**, *6*, 5047–5083. [[CrossRef](#)]
22. Soliman, G.; Soussa, H. Wetland change detection in Nile swamps of southern Sudan using multitemporal satellite imagery. *J. Appl. Remote Sens.* **2011**, *5*, 053517. [[CrossRef](#)]
23. Sun, F.; Sun, W.; Chen, J.; Gong, P. Comparison and improvement of methods for identifying waterbodies in remotely sensed imagery. *Int. J. Remote Sens.* **2012**, *33*, 6854–6875. [[CrossRef](#)]
24. Feyisa, G.L.; Meilby, H.; Fensholt, R.; Proud, S.R. Automated water extraction index: A new technique for surface water mapping using Landsat imagery. *Remote Sens. Environ.* **2014**, *140*, 23–35. [[CrossRef](#)]
25. Klein, I.; Dietz, A.J.; Gessner, U.; Galayeva, A.; Myrzakhmetov, A.; Kuenzer, C. Evaluation of seasonal water body extents in Central Asia over the past 27 years derived from medium-resolution remote sensing data. *Int. J. Appl. Earth Obs. Geoinf.* **2014**, *26*, 335–349. [[CrossRef](#)]
26. Moser, L.; Voigt, S.; Schoepfer, E. Monitoring of critical water and vegetation anomalies of Sub-Saharan West-African wetlands. In Proceedings of the IEEE International Conference on Geoscience and Remote Sensing Symposium, Québec City, QC, Canada, 13–18 July 2014; pp. 3842–3845.

27. Sethre, P.; Rundquist, B.; Todhunter, P. Remote detection of prairie pothole ponds in the Devils Lake Basin, North Dakota. *GISci. Remote Sens.* **2005**, *42*, 277–296. [[CrossRef](#)]
28. Laguna, C.; Gosálvez, R.; Sánchez, G.; Falomir, J.; Velasco, A.; Florín, M.; Gil-Delgado, J.; Chicote, A. Climate change footprint in the Mancha húmeda biosphere reserve. In Proceedings of the Energy and Environment Knowledge Week, Toledo, Spain, 28–29 October 2013; pp. 183–185.
29. Vidal, D.; Anza, I.; Taggart, M.A.; Pérez-ramírez, E.; Crespo, E.; Hofle, U.; Mateo, R. Environmental factors influencing the prevalence of a Clostridium botulinum type C/D mosaic strain in nonpermanent Mediterranean wetlands. *Appl. Environ. Microbiol.* **2013**, *79*, 4264–4271. [[CrossRef](#)] [[PubMed](#)]
30. Florín, M.; Montes, C. Functional analysis and restoration of Mediterranean lagunas in the Mancha Húmeda Biosphere Reserve (Central Spain). *Plant Ecol.* **1999**, *109*, 97–109.
31. Schroeder, T.A.; Wulder, M.A.; Healey, S.P.; Moisen, G.G. Mapping wildfire and clearcut harvest disturbances in boreal forests with Landsat time series data. *Remote Sens. Environ.* **2011**, *115*, 1421–1433. [[CrossRef](#)]
32. Yang, J.; Weisberg, P.J.; Bristow, N.A. Landsat remote sensing approaches for monitoring long-term tree cover dynamics in semi-arid woodlands: Comparison of vegetation indices and spectral mixture analysis. *Remote Sens. Environ.* **2012**, *119*, 62–71. [[CrossRef](#)]
33. Canty, M.J.; Nielsen, A.A. Automatic radiometric normalization of multitemporal satellite imagery with the iteratively re-weighted MAD transformation. *Remote Sens. Environ.* **2008**, *112*, 1025–1036. [[CrossRef](#)]
34. Scaramuzza, P.; Micijevic, E.; Chander, G. *SCL Gap-Filled Products. Phase One Methodology*; USGS-United States Geology Survey, USA, 2004. Available online: https://landsat.usgs.gov/documents/SLC_Gap_Fill_Methodology.pdf (accessed on 17 September 2014).
35. Rouse, J.W.; Haas, R.H.; Deering, D.W.; Sehell, J.A. *Monitoring the Vernal Advancement and Retrogradation (Green Wave Effect) of Natural Vegetation*; Remote Sensing Center. Report RSC 1978-4; Texas A&M University: College Station, TX, USA, 1974.
36. Huete, A.R. A soil-adjusted vegetation index (SAVI). *Remote Sens. Environ.* **1988**, *25*, 295–309. [[CrossRef](#)]
37. Ángel-Martínez, M.C. *Aplicación de la Teledetección en la Localización de Superficies de Agua*; CEDEX: Madrid, Spain, 1994.
38. Bustamante, J.; Díaz-Delgado, R.; Aragonés, D.; Pacios, F. Determining water body characteristics of Doñana shallow marshes through remote sensing. In Proceedings of the IEEE International Conference on Geoscience and Remote Sensing Symposium, Denver, CO, USA, 31 July–4 August 2006; pp. 3662–3663.
39. Doña, C.; Sanchez, J.M.; Caselles, V.; Dominguez, J.A.; Camacho, A. Empirical Relationships for Monitoring Water Quality of Lakes and Reservoirs Through Multispectral Images. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2014**, *7*, 1632–1641. [[CrossRef](#)]
40. Srivastava, P.K.; Han, D.; Rico-Ramirez, M.A.; Bray, M.; Islam, T. Selection of classification techniques for land use/land cover change investigation. *Adv. Sp. Res.* **2012**, *50*, 1250–1265. [[CrossRef](#)]
41. Smola, A.J.; Schölkopf, B. A tutorial on support vector regression. *Stat. Comput.* **2004**, *14*, 199–222. [[CrossRef](#)]
42. Seifert, J.W. Data mining: An overview. In *National Security Issues*; Pegarkov, D.D., Ed.; Nova Science Publishers, Inc.: New York, NY, USA, 2006; pp. 201–217.
43. Doña, C.; Chang, N.B.; Caselles, V.; Sánchez, J.M.; Camacho, A.; Delegido, J.; Vannah, B.W. Integrated satellite data fusion and mining for monitoring lake water quality status of the Albufera de Valencia in Spain. *J. Environ. Manag.* **2015**, *151*, 416–426.
44. Francone, D. *Discipulus Software Owner's Manual, Version 3.0 DRAFT*; Machine Learning Technologies, Inc.: Littleton, CO, USA, 1998.
45. Landis, J.R.; Koch, G.G. The measurement of observer agreement for categorical data. *Biometrics* **1977**, *33*, 159–174. [[CrossRef](#)] [[PubMed](#)]
46. Allen, R.G.; Pereira, L.S.; Raes, D.; Smith, M.; Ab, W. *Crop Evapotranspiration—Guidelines for Computing Crop Water Requirements—FAO Irrigation and Drainage Paper 56*; FAO: Rome, Italy, 1998; pp. 1–15.
47. Sankarasubramanian, A.; Vogel, R.M.; Limbrunner, J.F. Climate elasticity of stream flow in the United States. *Water Resour.* **2001**, *37*, 1771–1781. [[CrossRef](#)]



© 2016 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).