

**CO₂ UPTAKE BY
MEDITERRANEAN
ECOSYSTEMS**

RETRIEVAL AND ASSESSMENT OF
CO₂ UPTAKE BY MEDITERRANEAN
ECOSYSTEMS USING REMOTE SENSING
AND METEOROLOGICAL DATA

2014

Álvaro Moreno Martínez
Doctoral Dissertation

Doctoral Dissertation
Doctoral studies in Remote Sensing
Universitat de València
February 2014

**RETRIEVAL AND ASSESSMENT OF
CO₂ UPTAKE BY MEDITERRANEAN
ECOSYSTEMS USING REMOTE SENSING
AND METEOROLOGICAL DATA**

Álvaro Moreno Martínez

ADVISORS

M^a Amparo Gilabert Navarro
Beatriz Martínez Díaz



VNIVERSITAT DE VALÈNCIA

Las Dras. M^a Amparo Gilabert Navarro y Beatriz Martínez Díaz

CERTIFICAN:

Que la presente memoria "*Retrieval and assessment of CO₂ uptake by Mediterranean ecosystems using remote sensing and meteorological data*" recoge el trabajo realizado, bajo su dirección, por Álvaro Moreno Martínez para optar al grado de Doctor en Teledetección (con Mención Internacional) por la Universitat de València.

Y para que así conste firman la presente a 17 de febrero de 2014.

M^a Amparo Gilabert

Beatriz Martínez

Este trabajo se ha realizado en el marco de los siguientes proyectos de investigación: **ÁRTEMIS** (CGL2008-00381) Desarrollo de técnicas de teledetección para estimar indicadores biofísicos y analizar flujos de CO₂ en bosques de España, subvencionado por el Ministerio de Ciencia e Innovación; **RESET CLIMATE** (CGL2012-35831) Teledetección de variables climáticas esenciales: efectos del estrés hídrico en la estimación de los flujos de carbono, subvencionado por el Ministerio de Economía y Competitividad; **LSA-SAF** (*Satellite Application Facilities for Land Surface Analysis*) subvencionado por EUMETSAT.

Los datos meteorológicos empleados han sido cedidos por la AEMet (Agencia Estatal de Meteorología) y el SMC (Servei Meteorològic de Catalunya). Los datos de GPP *in situ*, subvencionados por los proyectos CarboEuropeIP (EU-FP6), IMECC (EU-FP6) y GHG-Europe (EU-FP7), están accesibles a la comunidad científica gracias a la iniciativa *European Fluxes Database Cluster*.

Los productos del satélite MODIS son proporcionados por el LP DAAC de la NASA, los de MSG por el LSA-SAF siendo propiedad de EUMETSAT y el producto DMP ha sido desarrollado por VITO dentro del proyecto MARS-OP3.

El autor de este trabajo ha disfrutado de una ayuda Predoctoral de Formación de Personal Investigador (ref. BES-2009-018638) asociada al proyecto ÁRTEMIS.

Agradecimientos

Siempre que se aproxima el fin de una tarea de largo recorrido, como es el caso de la realización de una tesis, parece saludable e incluso necesario hacer un balance de los factores que nos han llevado a conseguir el objetivo. A pesar de que muchos han sido los cambios en mi vida en estos años, con mejor o peor acierto mis acciones han ido dirigidas a avanzar descubriendo, analizando y aprendiendo con el fin de culminar mi tarea. Esta, más que un trabajo, ha sido un paseo agradable, ya que a pesar de las pendientes pronunciadas y los valles, los esfuerzos, las frustraciones y los logros, lo más importante es lo aprendido en el transcurso del camino. Además, estoy convencido de que acabo con algo más que ciencia, técnicas de procesado de datos y algoritmos en mi cabeza, lo cual es mucho más importante y, sobre todo, gratificante.

Para que esto haya sido así, gran parte de la responsabilidad recae sin duda sobre mi directora María Amparo Gilabert, quien ha sabido aunar sabiamente presión, ánimo, paciencia, experiencia y conocimiento para que se pudieran alcanzar los objetivos propuestos centrándome hacia la consecución de los mismos. También tengo que agradecer a mi otra directora, Beatriz Martínez, por sus correcciones y aportaciones al trabajo que han permitido redondearlo y mejorarlo.

A Javier García su paciencia y lo mucho que aprendí y aprendo de él día tras día desde que me inicié en esto de la teledetección dentro de la UIT. A Joaquín Meliá por sus sabios consejos y proveerme de fondos para

permitirme terminar mi trabajo. No me puedo olvidar de Fabio y Marta, quienes hicieron tan agradables y fructíferas científicamente mis estancias en el CNR de Florencia. Ni de mis “compis” cuando aterricé en el departamento, Ana, Gloria y Jordi, con los que he pasado grandes momentos entre trabajo y risas dentro y fuera de la universidad.

A Emilio Soria y José David Martín de la ETSE, ya que mi viaje en esto de la ciencia vino motivado por ellos y les debo gran parte de mis pocos conocimientos, al resto de gente del departamento de *Física de la Terra i Termodinàmica* y la gente de EOLAB. Aunque ya se ha citado la procedencia de los datos utilizados en la memoria, agradezco especialmente a J. Tamayo, G. Seufert, A. Carrara, O. Pérez y P. Serrano por facilitar el acceso a ellos.

A mi familia, a la que está y sobre todo a la que se fue, por la oportunidad que me han brindado en la vida de alcanzar esta meta con sus consejos, apoyo y amor incondicional. Estoy seguro de que están orgullosos de mí, de igual forma que yo estoy infinitamente agradecido y orgulloso de ellos. Tampoco me puedo olvidar de mi otra familia, mis más que amigos de siempre: Emma, Jordi, Andrés, Amparo, Juanmi, Cristóbal, Javi Lambies, Miguel Ángel, Raúl, Jose Lambies, Pitu y otros más recientes como: Goyi, Amparito, Rojillez, Migueler, Virgil, Ángela, Carlitos, Jorge, Rubén, Paco, Rafa, Gess... porque son con ellos con los que recargo las pilas todos los fines de semana y comparto penas y alegrías siendo, por suerte, muchas más las segundas.

Álvaro Moreno Martínez

Índice general

Summary	xvii
1. Introducción	1
1.1. El ciclo del carbono	1
1.2. Flujos de carbono en los ecosistemas terrestres	4
1.2.1. Determinación <i>in situ</i> de flujos de C mediante la técnica <i>eddy covariance</i> (EC)	9
1.2.2. Cálculo de flujos de C a partir de datos de satélite	10
1.3. Objetivos del trabajo	13
2. Modelo optimizado para el cálculo de la GPP	17
2.1. Imágenes de f_{APAR}	17
2.1.1. Cálculo de la f_{APAR}	19
2.1.2. Postprocesado de las series obtenidas de f_{APAR}	21
2.2. Imágenes de PAR	30
2.2.1. Obtención a partir de datos de estaciones meteorológicas	33
2.2.2. Obtención a partir de modelos con datos de satélite	37
2.3. Imágenes de ε	41
2.3.1. Imagen de ε_{max} a partir de un mapa híbrido de tipos de cubierta	42
2.3.2. Estimación de la ε actual a partir de datos meteorológicos	45
2.3.3. Estimación de la ε actual a partir de datos de satélite	48

3. Cálculo y validación del producto de GPP	55
3.1. Cálculo de la GPP	55
3.2. Validación de resultados	63
3.2.1. Validación directa utilizando datos de torres EC . . .	63
3.2.2. Validación indirecta comparando con otros productos operacionales	71
3.3. Discusión	82
3.3.1. Análisis del potencial explicativo del modelo utilizando datos <i>in situ</i>	83
3.3.2. Análisis a nivel de imagen de la contribución de las variables de entrada al modelo optimizado	86
4. Conclusions	99
Apéndice A.	117
<i>Publication I: Mapping daily global solar irradiation over Spain: A comparative study of selected approaches</i>	117
Apéndice B.	133
<i>Publication II: Validation of daily global solar irradiation images from MSG over Spain</i>	133
Apéndice C.	147
<i>Publication III: Assessment of MODIS imagery to track light-use efficiency in a water-limited Mediterranean pine forest</i>	147
Apéndice D.	159
<i>Publication IV: Monitoring water stress in Mediterranean semi-natural vegetation with satellite and meteorological data</i> . .	159
Apéndice E. Error estimation	171

Índice de figuras

1.1.	Esquema del ciclo del carbono. Fuente: http://mediateca.cl .	2
1.2.	Procesos y tránsito de los flujos de carbono en los ecosistemas terrestres. En sombreado se muestran todas las componentes de la NEP en el ecosistema. Fuente: <i>Lovett et al.</i> [21].	6
2.1.	Perfiles de f_{APAR} para cuatro tipos de cubierta vegetal y cuatro productos operacionales: el algoritmo Roujean y Bréon de LSA-SAF (negro), inversión mediante redes neuronales y datos MERIS (azul), índice de vegetación optimizado con datos MERIS (verde) y el producto oficial de MODIS (rojo). Fuente: <i>Martínez et al.</i> [49].	21
2.2.	Perfiles de f_{APAR} para cuatro tipos de cubierta vegetal. El perfil de color rojo muestra la serie temporal filtrada y reconstruida mediante el método <i>loess</i> adaptado y en color negro se muestran los perfiles originales de f_{APAR}	24
2.3.	Imagen de f_{APAR} original (a) y de f_{APAR} filtrada y reconstruida mediante el método <i>loess</i> adaptado (b). Corresponde con la fecha 1/1/2011.	25
2.4.	Mosaico de perfiles de f_{APAR} sin/con ruido para el área clasificada como mezcla de bosques y los resultados del filtrado con los diferentes métodos considerados (a). RMSE para el mismo área considerando todos los valores de los parámetros del ruido considerado (b).	28

2.5. Variación del MBE en función del porcentaje de datos afectado por el ruido	29
2.6. Relación entre la tasa fotosintética neta y la radiancia fotosintéticamente activa mostrando los diferentes procesos que limitan la fotosíntesis a diferentes niveles de irradiancia.	31
2.7. Imágenes diarias de PAR para diferentes fechas (una de cada mes) a lo largo del año 2008.	36
2.8. Diagrama de cajas con los sesgos entre los datos medidos <i>in situ</i> y los estimados mediante satélite (<i>a</i>), diagrama de dispersión entre todos los datos medidos <i>in situ</i> por la AEMet y las estimaciones de satélite (<i>b</i>).	40
2.9. Imagen diaria de irradiación para el 1/7/2008 corregido de efectos topográficos y de línea de costa.	40
2.10. Mapa híbrido combinado de las principales cartografías de tipos de cubierta. Fuente: <i>A. Pérez-Hoyos [87]</i>	44
2.11. Ejemplo de la recta ajustada para la obtención del valor del factor de estrés por bajas temperaturas. Fuente: <i>Heinsch et al. [6]</i>	46
2.12. PRI calculado mediante un modelo exclusivamente dependiente de la información geométrica de adquisición (negro) y PRI estimado a partir de datos de reflectividad MODIS (rojo).	50
2.13. ε y NPRI calculado con datos MODIS.	51
2.14. Serie temporal original de C_{ws} y tres modificaciones empleando tres índices espectrales a partir de datos MODIS para la torre de San Rossore en Italia.	53
2.15. Cálculo de la GPP a partir de datos <i>eddy covariance</i> y a partir del modelo optimizado con el factor original C_{ws} y con el C_{ws}^* modificado empleando el índice espectral SIWSI.	54
3.1. Diagrama de flujo del modelo optimizado empleado para el cálculo de la GPP. Los pasos 1, 2, 3 se han descrito en el capítulo 2. El paso 4 se desarrolla en este capítulo.	57

3.2. GPP diaria calculada empleando el modelo optimizado a una resolución espacial de 1 km. Se muestra un día de cada mes para el año 2011.	59
3.3. GPP anual calculada para los años 2008 (a) y 2011 (b).	61
3.4. Mapa de diferencias anuales de GPP (a) y de precipitación anual (b) entre los años 2008 y 2011.	62
3.5. Evolución temporal de la GPP calculada con el modelo optimizado y la GPP procedente de las torres EC para el año 2008.	68
3.6. Diagramas de dispersión entre la GPP calculada con el modelo optimizado y la GPP procedente de todas las torres EC consideradas en el año 2008.	69
3.7. Evolución temporal de las estimaciones de la GPP calculada con el modelo optimizado y la GPP procedente de las torres EC para el año 2011.	71
3.8. Diagramas de dispersión entre la GPP calculada con el modelo optimizado y la GPP procedente de todas las torres EC consideradas en el año 2011.	72
3.9. Evolución temporal de la GPP de los tres productos operacionales y las estimaciones con datos de torres EC para el año 2011.	76
3.10. Cálculo de los sesgos (MBD) entre el modelo optimizado y DMP (a) y el modelo optimizado y MODIS (b) para la GPP anual en 2011 ³	78
3.11. Consistencia temporal entre las estimaciones obtenidas por el modelo optimizado y los modelos operacionales DMP (a) y MODIS (b).	80
3.12. Perfiles temporales de GPP para el año 2008 (Llano de los Juanes) y 2011 (Majadas del Tiétar) estimados con datos de torres EC (EC), con el modelo optimizado (OPT) y con el el modelo optimizado eliminado el efecto del factor de estrés hídrico (OPT w/o C_{ws}). En las figuras inferiores se muestran los correspondientes factores de estrés hídrico para cada una de las zonas de estudio.	85

3.13. Correlaciones temporales respecto al modelo optimizado para cada una de las variables involucradas.	88
3.14. Parámetros fenológicos de la vegetación ΔNDVI y t_{max} (amplitud del ciclo fenológico y mes del máximo de la vegetación, respectivamente) obtenidos mediante la técnica <i>wavelets</i> en series temporales del índice de vegetación NDVI. Fuente: <i>Martínez & Gilabert [101]</i>	91
3.15. Precipitación acumulada mensual y temperatura mínima media mensual para los dos años en los que se ha calculado la GPP.	91
3.16. Zonas seleccionadas para la discusión de resultados.	94
3.17. Perfil mensual del factor de estrés hídrico (a) y de f_{APAR} (b) para la zona de estudio “B” (Almería) y los años 2008 y 2011.	95
3.18. Perfil mensual de las temperaturas mínimas mensuales (a) y de f_{APAR} (b) para la zona de estudio “E” (Valladolid) y los años 2008 y 2011.	96

Índice de tablas

1.1. Modelos de productividad a partir de datos de satélite. La tabla muestra tanto sus variables de entrada como los flujos obtenidos. Tabla adaptada de [32].	12
3.1. Datos más relevantes de las torres <i>eddy covariance</i> . OSH hace referencia a matorrales dispersos, CSH a matorrales densos, CRP a cultivos (arroz) y EBF a bosques perennes de hoja ancha.	66
3.2. Resultados de la validación directa para los píxeles donde se ubican cada una de las torres EC disponibles en el año 2008 ¹	67
3.3. Resultados de la validación directa para cada una de las torres EC disponibles para el año 2011.	70
3.4. Características de los productos de GPP utilizados en la intercomparación.	73
3.5. Resultados de la comprobación de la calidad para los píxeles donde se ubican cada una de las torres EC disponibles en el año 2011. Se muestran los estadísticos obtenidos de los tres productos considerados GPP_{MODIS} , GPP_{DMP} y GPP_{OPT}	74
3.6. Valores de la correlación (R) en función de las variables incluidas en el modelo para las seis torres en 2008. AGU hace referencia a la torre de Aguamarga, LAG a Lagunaseca, JUA a Llano de los Juanes, LAN a Lanjarón, MAJ a las Majadas del Tiétar y SUE a la torre de Sueca.	83

3.7. Valores de la correlación (R) en función de las variables incluidas en el modelo para las cuatro torres en 2011. AGU hace referencia a la torre de Aguamarga, COR a Cortes de Pallás, JUA a Llano de los Juanes y MAJ a las Majadas del Tiétar.	84
--	----

Summary

According to the Doctorate Regulation^a, the Thesis Dissertation could be a compendium of the publications (three at least) containing original results of the work. The results of this Thesis have been reported in four peer-reviewed publications where the methodology is developed. However, it was decided to follow a standard scheme for the Dissertation to describe in detail the potential of the proposed GPP product optimized for the study area. Thus, Chapter 1 provides an introduction and states the main objectives. Chapter 2 offers a brief description of each of the inputs in Monteith's approach. Chapter 3 shows the GPP product and evaluates its quality. As a requirement of the Ph.D. International Mention^b, Spanish and English languages have been used. In particular, Chapters 1, 2, and 3 and are written in Spanish, while this Summary, Conclusions, and the Appendices) are written in English.

Regretfully, symbols have not been used consistently in the publications and this Report, or even among different publications, due to the different journals' policy. In the present Report we have adapted the symbols and unit rules to the conventions established by NIST and IUPAC. Hence, they can differ from the standard usage in GPP studies (for example, the GPP is given in $\text{g m}^{-2} \text{ day}^{-1}$, and not in $\text{gC m}^{-2} \text{ day}^{-1}$).

^aReal Decreto 99/2011, de 28 de enero y el artículo 136 de los Estatutos de la Universitat de València.

^bArticle 9, Reglament sobre Dipòsit, avaluació i defensa de la Tesi Doctoral.

Photosynthesis is a process by which carbon and energy enter ecosystems. The knowledge of where, when, and how carbon dioxide (CO_2) is exchanged between terrestrial ecosystems and atmosphere is crucial to close the Earth's carbon budget and predict feedbacks in a likely warming climate. *Gross photosynthesis* (uptake of CO_2) by vegetation is responsible for the *gross primary production* (GPP) of the ecosystem. Normally GPP refers to the sum of the photosynthesis by all leaves measured at the ecosystem scale. Part of the carbon gained is respired during daytime and in the dark by foliage, tree boles, and living roots, forming the autotrophic respiration (R_a). The *net primary production* is $\text{NPP} = \text{GPP} - R_a$. An additional flux of respired CO_2 , *heterotrophic respiration* (R_h) arises from the decay and consumption of dead carbon-containing material (e.g. dead roots, leaf litter, soil carbon) by fungi, micro-organisms, and other living components of the ecosystem, so that the *net ecosystem productivity* is $\text{NEP} = \text{NPP} - R_h = \text{GPP} - R_a - R_h$.

John Monteith proposed in 1972 [1] a simple approach that has become the paradigm for understanding GPP. It considers GPP as proportional to the incident short wave radiation (PAR), the fractional absorption of that flux (f_{APAR}) and the radiation use conversion efficiency, also known as light-use efficiency (ε):

$$\text{GPP} = \varepsilon f_{\text{APAR}} \text{PAR}$$

This simple equation involves a great deal of biological and biophysical complexity. Photosynthesis requires that the plant replace the water that inevitably escapes from its leaves when CO_2 is taken up from the atmosphere. Plants also require a supply of nutrients. Physiological and developmental mechanisms operate to adjust the GPP to the availability of resources. Thus, different types of stresses can affect the efficiency.

The different terms in Monteith's equation are emphasized by different scientists. Crop physiologists focus on the PAR term, which explains the seasonal growth of crops and year-to-year variation in yield. Early work within the remote sensing community focused on the f_{APAR} term, which is linked to canopy structure and condition (i.e. to green biomass). It has a clear season-

al evolution in deciduous species and shows limited variability in evergreen forest ecosystems. The f_{APAR} is a common biophysical product derived from different remote sensing missions through the inversion of radiative transfer models or from empirical relations with vegetation indices. More recently the strong influence of the ε term on productivity –particularly in strongly seasonal and nutrient-limited and/or water stressed vegetation canopies– has been recognized. Variation in ε is significant over shorter time scales when water or temperature stress develop.

The ε has been shown to vary spatially between biomes, ecosystems, and plant species, and to vary temporally during the growing season, due to environmental and physiological limitations [2], [3], [4]. ε responds more rapidly than f_{APAR} to different environmental factors related to the energy balance, water availability and nutrient levels [4],[5]. For operational applications, ε can be expressed as the product of a ε_{max} (maximum light-use efficiency), which depends on cover type, and different terms accounting for the reduction in efficiency due to different types of stress. The computation of these terms frequently requires meteorological data, which are seldom available at the needed spatial and temporal scales.

The Monteith’s approach provides the theoretical basis for most production efficiency models (PEMs), also known as light-use-efficiency (LUE) models: the MODIS-GPP model describes the global terrestrial photosynthesis at 1 km spatial scale and various time steps [6]; the parametric model C-Fix [7] has been applied to estimate forest GPP in several European countries [8], [9]; and the modified C-Fix [10] also takes into account the short-term water stress, a typical feature of the hot and dry Mediterranean summer. These models use remotely sensed data as well as meteorological data. In most PEMs, f_{APAR} is the only satellite-derived variable and, as such, it provides the link between ecosystem function and structure.

Validation of satellite-derived GPP products is problematic. The development of eddy covariance (EC) as a method for quantifying the carbon, water, and energy balance over so-called “flux sites” has provided observational data to test and calibrate models; but the EC towers measure net CO_2 exchange. GPP is obtained from these measurements after correcting

them for respiratory losses (about half) [6]. The density of sampling is never enough to get regional or continental scale GPP. This is the domain of models. The modeling approaches also have specific limitations concerning: (i) the uncertainties of vegetation indices due to the presence of soil background mainly in sparse areas, and due to cloud and aerosol contamination problems, (ii) errors in the re-analysis of meteorological data, and (iii) difficulty constraining the light-use-efficiency term. The quality assessment of GPP products is rather complicated by the fact that GPP cannot be measured directly on a geographically relevant scale.

In this Thesis, a model to estimate GPP for Mediterranean ecosystems at regional scale is proposed. The three terms in Monteith's equation have been obtained following procedures optimized for the study area, Spain (excluding Canary Islands). The "optimized model" is driven by meteorological and satellite data (MODIS/TERRA and SEVIRI/MSG).

Considering the peculiarities of the study area, i.e., the diversity of the vegetation type dynamics and its spatial heterogeneity, the algorithm has been developed to run at a daily time step (to capture the dynamics even in agro-ecosystems) and 1 km spatial resolution (to assure that the spatial resolution of the remote sensing estimates is comparable to the footprint of ground estimates). Thus, the inputs of the model have been retrieved at these temporal and spatial resolutions.

Daily PAR can be obtained as the 46 % of the total flux density of solar radiation that reaches the surface during a day [11], also known as daily irradiation. *Publication I* (Appendix A) shows a methodology to obtain irradiation images by applying an artificial neural network to temperature and precipitation images generated by ordinary kriging from in situ measurements. *Publication II* (Appendix B) evaluates the daily irradiation images obtained from the down-welling surface short-wave radiation flux (DSSF) product derived from the SEVIRI/MSG images. A downscaling of the DSSF product was also carried out, which was based on the spatial disaggregation of the irradiance by means of a digital elevation model that also accounts for the coastline correction and a topographic correction. The resulting images satisfactorily map the surface solar radiation at 1 km spatial resolution even

in rugged terrains. The validation of the images shows that the mean absolute error of the daily irradiation is ca. $2.3 \text{ MJ m}^{-2} \text{ day}^{-1}$ (product developed in *Publication I*) and $1.0 \text{ MJ m}^{-2} \text{ day}^{-1}$ (improved product developed from DSSF in *Publication II*). The second product relies only in remotely sensed data and presents thus a major advantage from an operational point of view. However, it is only available from 2011. For retrospective studies, the first product (*Publication I*) can be used.

f_{APAR} is obtained by applying the algorithm proposed by *Roujean & Bréon* [12] to MODIS data (at 1 km spatial resolution and 8-days temporal resolution). This algorithm is actually used to derive the SEVIRI/MSG fAPAR product, delivered by the LSA SAF network (EUMETSAT) (<http://landsaf.meteo.pt>) at 3.1 km spatial resolution (sub-satellite point) and daily frequency over the geostationary MSG grid. Subsequently, as described in [13] (Publication under revision), the f_{APAR} time series has been filtered using the *loess* local method in order to remove spurious drops and undesirable day-to-day variability (noise) in the data resulting from cloud, ozone, dust, and other aerosols that generally decrease the near-infrared reflectance. This methodology has been shown to capture the upper envelope of the time-series, to interpolate the missing data and to remove most of the noise of the original unfiltered signal.

A hybrid land-cover map for Spain [14] was used to assign ε_{max} to the different cover types. This multi-classification map was obtained from the synergistic combination of a number of four land-cover classifications (CORINE, GLC2000, MODIS and GlobCover) with different legends and spatial resolutions. Values of ε_{max} were obtained from literature [4]. Regarding the reduction in efficiency as due to stress, two terms corresponding to thermal and water stress were considered. In particular, water stress was recognized to highly influence the inter-annual variation of ε in Mediterranean ecosystems. The term quantifying the thermal stress was similar to that used by *Heinsch et al.* [6] in the MODIS-GPP algorithm. Meteorological data were also used to characterize the ε inter-annual variability due to water stress. In particular, the water stress factor proposed by *Maselli et al.* [10], C_{ws} , was selected because it accounts for limited photosynthetic activity in case of short-term water stress. C_{ws} was obtained from a local water budget

based on actual and potential evapotranspiration. The scientific community is now searching for a procedure to account for this stress by means of a remotely sensed data. Different spectral indices derived from MODIS data have been analyzed: the PRI and several broad-band spectral indices combining the near-infrared band with short-wave infrared bands. *Publication III* (Appendix C) shows that most (ca. 70%) of the observed concurrent seasonal variability in ε and PRI is due to the effect of illumination and viewing conditions (first order effects). The chlorophyll and canopy structure introduce second order effects, and the deepoxidation of xanthophylls (third order effects) involve corrections that are, at MODIS pixel scale, of the same magnitude order than the sensor noise. *Publication IV* (Appendix D) shows that the water stress dependence of ε cannot be estimated only from such broad-band spectral indices (as derived from MODIS data). However, existing series of meteorological data characterizing a site can be useful to update C_{ws} for the actual conditions.

The daily GPP product obtained as explained above is difficult to validate due to the lack of ground GPP data. Nevertheless, GPP estimations from several eddy covariance (EC) towers have been used. These towers belong to the European Fluxes Database Cluster (<http://www.europe-fluxdata.eu>). By chance, these EC towers are mainly located in the semi-arid areas, which are more difficult to model due to their larger soil background effects. Thus, this direct validation of the GPP product serves to establish its upper uncertainty level. Moreover, an indirect validation, by means of an inter-comparison with two other operational products (from MODIS and Copernicus), is carried out. The results have been highly satisfactory and promising. A further analysis of the percentage of variance associated with each input of the Monteith's equation clearly evidences the role of the water stress in the inter-annual variation of GPP in Mediterranean ecosystems.

February 2014

Capítulo 1

Introducción

En este primer capítulo se presenta el marco conceptual en el que se desarrolla este trabajo y los objetivos planteados.

1.1. El ciclo del carbono

El ciclo del carbono consiste en un conjunto de procesos físico-químicos complejos por los que el carbono existente en la Tierra circula entre los diferentes compartimentos en los que se almacena, siendo estos: las reservas geológicas, los océanos, la atmósfera, los suelos y la biomasa vegetal. En la figura 1.1 se muestran algunos de los diferentes procesos así como la dirección del transporte que ocasionan dentro del ciclo del carbono.

Aunque existen grandes incertidumbres acerca de las masas involucradas en el ciclo del carbono dentro de dichos compartimentos [15], a continuación se muestran unas cifras orientativas. En la atmósfera se estima que existen entre 720 y 780 Pg, principalmente en forma de dióxido de carbono. En la superficie vegetal terrestre se encuentra almacenado de forma transitoria en materia viva 550 Pg y en materia vegetal muerta 300 Pg. Se estima que en los suelos se pueden albergar unos 2 300 Pg, mientras que los mares y océanos pueden almacenar en forma inorgánica (disuelta como bicarbona-

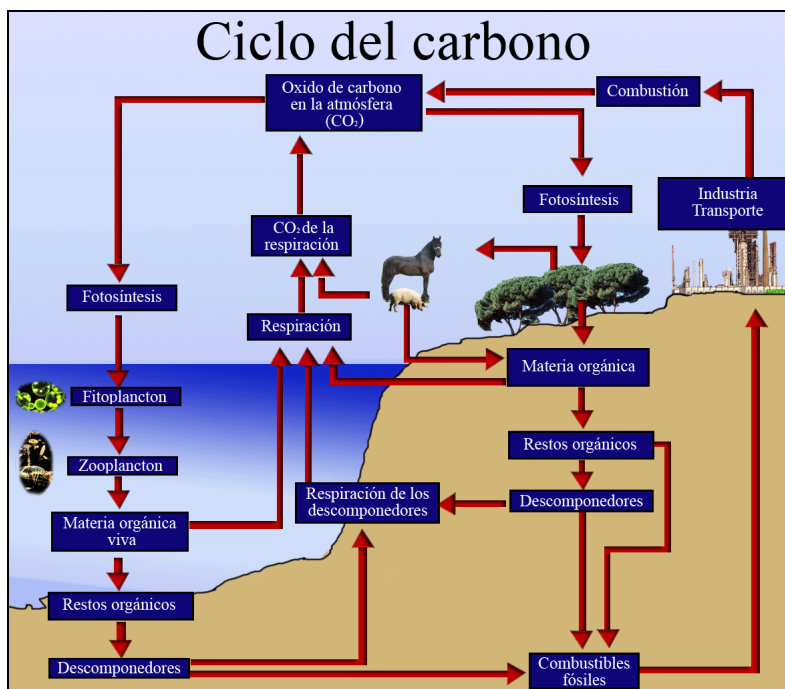


Figura 1.1. Esquema del ciclo del carbono. Fuente: <http://mediateca.cl>.

to cálcico y carbonato sódico) y de forma orgánica aproximadamente unos 3 800 Pg. Si comparamos las cifras anteriores con los datos de carbono geológico vemos que son notablemente inferiores, así, en forma activa (sedimentos precipitados, oceánicos y terrestres) tenemos unos 40 000 000 Pg, de los que una parte (10 000 Pg) está accesible para su extracción y posterior utilización como combustibles fósiles. Existe una gran demanda de combustibles fósiles para abastecer las necesidades energéticas de los países desarrollados y los países emergentes. Este problema se ve agravado por un constante aumento de la población mundial. Así pues, es evidente que las actividades humanas tienen una repercusión directa en los intercambios de C entre los diferentes compartimentos afectando su distribución en cada uno de ellos ya que, aparte de la problemática asociada con la extracción y uso de los combustibles fósiles, las acciones humanas también tienen un importante efecto sobre los ecosistemas terrestres transformando la vegetación natural en cultivos, ciudades o simplemente destruyéndola mediante incendios.

La temperatura media del planeta y, por tanto, las condiciones más favorables para el desarrollo de los procesos biológicos son consecuencia del denominado “efecto invernadero”, producido por algunos de los componentes minoritarios de la atmósfera como CO_2 , $\text{H}_2\text{O}(\text{v})$, CH_4 y N_2O , entre otros. Estos gases, fuertemente absorbentes y, por tanto, emisores de la radiación de onda larga, originan un incremento de la temperatura superficial del planeta en unos $33\text{ }^\circ\text{C}$ respecto a la que tendría en ausencia de una atmósfera absorbente, pasando así de $-18\text{ }^\circ\text{C}$ a $+15\text{ }^\circ\text{C}$.

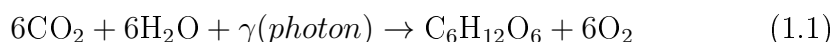
A partir del estudio de la composición de los hielos en la Antártida, se han inferido largas series históricas (400 000 años) de concentraciones de CO_2 atmosférico y temperatura superficial. El análisis de estos datos ha indicado la existencia de importantes correlaciones entre ambas variables. Los resultados también muestran que las concentraciones atmosféricas mundiales de CO_2 , CH_4 y N_2O han aumentado por efectos antropogénicos desde 1750, llegando en la actualidad a ser muy superiores a los valores preindustriales [15]. El IPCC (*Intergovernmental Panel on Climate Change*) apunta que, sin una reducción de las concentraciones de CO_2 , la temperatura superficial se incrementará en un futuro unos $0.3\text{ }^\circ\text{C}$ por década, causando la desaparición de los casquetes polares. Los modelos climáticos predicen cambios significativos en el sistema climático a lo largo del siglo XXI muy probablemente mayores a los ya observados en el siglo XX [15].

Los ecosistemas terrestres desarrollan un papel fundamental en el ciclo del carbono a través de la fotosíntesis, la respiración, la combustión de biomasa y la descomposición [16]. Existe una importante transferencia de masa de carbono entre estos y la atmósfera (60 Gt/año) [17], pero ambos sistemas están fuertemente acoplados. La cobertura vegetal afecta a la composición atmosférica y a los flujos de energía en la superficie terrestre y, al mismo tiempo, la composición atmosférica y las condiciones atmosféricas influyen directamente en la actividad de la vegetación a través de la irradiación solar, la disponibilidad de agua y la temperatura. Así pues, en un escenario como el que indican las predicciones de incremento de concentración de CO_2 atmosférico, la vegetación respondería aumentando su crecimiento y posterior aumento de contenido de materia orgánica edáfica, ya que se potencia la eficacia hídrica y la tasa fotosintética. Esta respuesta puede ser amortigua-

da bajo limitaciones de nutrientes, déficit de agua o temperaturas extremas de forma directa (condiciones ambientales no óptimas para la realización de la fotosíntesis) o indirecta (aumento del riesgo de incendios, plagas,...). En cualquier caso, la gestión estratégica de los ecosistemas terrestres es una herramienta susceptible de ser empleada para mitigar el aumento de las concentraciones de gases de efecto invernadero ya sea vía la conservación de las zonas existentes, incrementando su extensión o mediante el desarrollo de políticas diferentes respecto al consumo de la energía vegetal. El estudio, por tanto, de los ecosistemas terrestres es de gran importancia en el contexto actual del estudio de la variabilidad climática, tanto por la necesidad de un mayor conocimiento de sus interacciones con el resto de elementos del ciclo del carbono como por la importancia de entender mejor un sistema de gran dinamismo espacial y temporal.

1.2. Flujos de carbono en los ecosistemas terrestres

La fotosíntesis es el proceso biológico por el que la vegetación provee del carbono y la energía necesarias para realizar la mayoría del resto de procesos biológicos en los ecosistemas. Esta energía es fijada y es directamente empleada por la vegetación para su crecimiento produciendo materia orgánica, que será posteriormente consumida por microorganismos y resto de animales de manera directa o indirecta [18]. La reacción química 1.1 (ecuación de la fotosíntesis oxigénica) describe de forma simplificada el proceso de fotosíntesis por el cual las plantas capturan parte de la radiación solar y la transforman en energía química.



La reacción de fotosíntesis es activada por la radiación solar comprendida entre los 400 y 700 nm (radiación fotosintéticamente activa, *Photosynthetically Active Radiation*, PAR). En la parte derecha de la ecuación (productos

de la reacción) se muestran los carbohidratos en los cuales se almacena la energía. Existen diferentes “caminos” de fotosíntesis, si bien el más habitual es el de tipo C_3 realizado por el 85% de las especies vegetales vasculares (plantas superiores) [19]. En este proceso existen dos tipos de reacciones, en la primera se transforma la energía luminosa en energía química (conversión a trifosfato de adenosina) y en la segunda se transforma el CO_2 en azúcares para que pueda ser almacenada, transportada y metabolizada. Este proceso está regulado principalmente por la disponibilidad de agua y los nutrientes necesarios. Las plantas disponen de mecanismos de protección para evitar tasas dañinas de captura energética. Se pueden destacar dos principalmente: la fotorespiración y el ciclo de las xantofilas. El primero actúa como una válvula de seguridad para evitar la producción de radicales que destruirían los pigmentos fotosintéticos y el segundo es un mecanismo de fotoprotección cuyo objetivo es disipar el exceso de energía en forma de calor mediante un cadena de cambios entre pigmentos [19]. Aparte de estos métodos de protección para preservar las estructuras internas de las plantas, estas son capaces de controlar la fijación de CO_2 así como las pérdidas de agua modificando el tamaño de los estomas (conductancia estomática). Cuando las plantas reducen la conductancia estomática para conservar agua se reduce la adquisición de CO_2 , de manera que lo que este mecanismo consigue es un compromiso entre maximizar la fotosíntesis y minimizar las pérdidas de agua dependiendo de un buen abastecimiento de CO_2 , energía luminosa y nutrientes minerales [20].

Existe cierta ambigüedad en la literatura científica respecto a las definiciones de los términos utilizados en los balances y flujos de carbono [21]. En este trabajo se han empleado las definiciones que son más comunes en la bibliografía [22]. El carbono es incorporado en los ecosistemas como producción primaria bruta (*Gross Primary Production*, GPP) y es transportado a otros ecosistemas, devuelto a la atmósfera mediante respiración o perturbaciones de diferente índole al comportamiento rutinario del ecosistema [19]. Aproximadamente la mitad de la GPP [23] se emplea por las plantas en la respiración autótrofa (R_a), es decir, se dedica a la síntesis y mantenimiento de sus propias células. Otra variable de interés en los estudios de flujos viene definida por la diferencia $GPP-R_a$ a la que se denomina produc-

ción primaria neta (*Net Primary Production*, NPP). Esta da cuenta de la totalidad de biomasa producida por las plantas, los compuestos orgánicos exudados por las raíces en el suelo, las transferencias de carbono a los microorganismos simbióticamente unidos a las raíces y las emisiones volátiles perdidas por los hojas hacia la atmósfera. También se define la producción neta del ecosistema (*Net Ecosystem Production*, NEP) como la diferencia $GPP - R_e$ donde R_e es la respiración total del ecosistema y viene definida por la suma de la respiración autótrofa (R_a) y la respiración heterótrofa (R_h), que es la respiración debida a los animales y a los microorganismos presentes también en el ecosistema. El intercambio neto del ecosistema (*Net Ecosystem Exchange*, NEE) es, por definición, el flujo de carbono entre el ecosistema y la atmósfera y tendría en cuenta, a diferencia de la NEP, las fuentes y sumideros de CO_2 inorgánicos. Dado que en los ecosistemas terrestres estos suelen ser despreciables, se admite la aproximación $NEE \approx NEP$ [22].

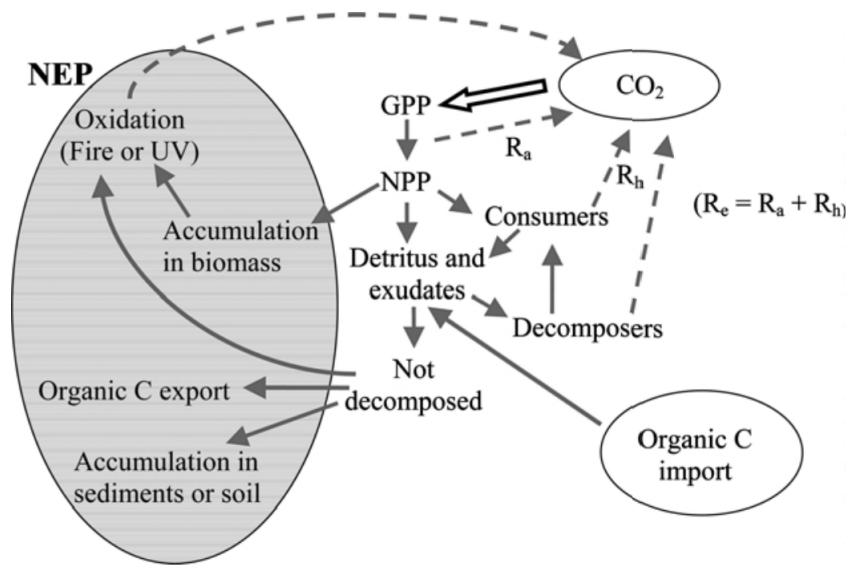


Figura 1.2. Procesos y tránsito de los flujos de carbono en los ecosistemas terrestres. En sombreado se muestran todas las componentes de la NEP en el ecosistema. Fuente: Lovett *et al.* [21].

Las definiciones y relaciones entre los diferentes flujos presentan dificultades para su correcta medida entre los diferentes ecosistemas a diferentes escalas espaciales y temporales. Los diferentes componentes del ecosistema

requieren de diferentes métodos de medida que han de ser autoconsistentes de acuerdo con la terminología mostrada [24]. Aún así, la fuente de errores más importante en la estimación de los flujos de carbono viene dada por la dificultad de determinar las pérdidas en los sumideros de carbono debido a variabilidad natural y a la transferencia lateral entre los ecosistemas terrestres [25]. A continuación se muestran diferentes aproximaciones complementarias para su estimación, ya que cada una permite estudiar diferentes aspectos del ciclo del carbono:

Estimaciones biométricas: Esta aproximación involucra campañas intensivas de medidas sobre el terreno, ya que se miden o estiman (dependiendo de la magnitud) incrementos anuales de la biomasa vegetal visible (ramas, hojas, tronco, etc) y la no visible (raíces, carbón orgánico, fauna del suelo, etc). Estas medidas suelen incorporarse a las series históricas de los inventarios forestales permitiendo estimar el carbono fijado. De acuerdo con los ecólogos, estas estimaciones son las más fiables a escalas temporales anuales y permiten validar otras metodologías [26]. Presentan incertidumbres asociadas con el uso de relaciones alométricas para estimar ciertos componentes, así como las propias del muestreo espacial para caracterizar un área determinada.

Estudios ecofisiológicos: Este tipo de trabajos proporcionan información acerca de los mecanismos potenciales que controlan los flujos de carbono dentro del ecosistema, ya que se centran en los procesos metabólicos o plantas individuales. Existen problemas cuando se desea extrapolar estos mecanismos a todo el ecosistema ya que existen procesos que no se pueden escalar convenientemente.

Modelos de ecosistemas: Este tipo de modelos, más que ser una vía para obtener estimaciones de los diferentes flujos de carbono, permiten establecer hipótesis y predicciones sobre diferentes escenarios. Generalmente son modelos complejos que describen mediante ecuaciones matemáticas los diferentes procesos que controlan los flujos de carbono en el ecosistema, requiriendo de métodos complejos de cálculo para la obtención de soluciones. Presentan el inconveniente de la dificultad para disponer de todos los datos de entrada que requieren, ya

sea por cuestiones económicas o técnicas.

Métodos micrometeorológicos: Existe una creciente línea de investigación en el uso de estos métodos para el estudio de las superficies vegetales. Se centran en el análisis de las interacciones entre la parte de la atmósfera más cercana a la cobertura vegetal. Existen aproximaciones a diferentes escalas temporales (desde milisegundos a horas) y espaciales (desde milímetros al kilómetro), si bien todos estos métodos presentan una alta resolución temporal, medidas continuas y automáticas, medidas directas del flujo y emplean instrumentos sofisticados que requieren un mantenimiento complicado [27]. Destacan los métodos de *eddy covariance* (EC) ¹, acumulación eddy, acumulación eddy relajada, aerodinámico, de resistencia y balance de masas (para áreas más pequeñas) como los más relevantes.

Modelos mediante datos de teledetección: Aunque pertenecerían al grupo de modelos de ecosistemas, se han incluido en otro apartado porque su filosofía es sustancialmente diferente. Los modelos que emplean datos de teledetección no pretenden caracterizar procesos complejos del ecosistema, sino que tratan de modelizar los procesos más significativos para poder obtener estimaciones a grandes escalas con datos principalmente de sensores remotos. El desarrollo de modelos para el uso de datos de satélite ha dado lugar a los llamados modelos de producción de ecosistemas (*Production Ecosystem Models*, PEM). Se basan fundamentalmente en la existencia de una relación aproximadamente constante entre la radiación solar incidente absorbida por la vegetación entre 400-700 nm (*Absorbed Photosynthetically Active Radiation*, APAR) y el carbono que es fijado por la misma (GPP) a través de fotosíntesis.

¹En este trabajo se utiliza la nomenclatura anglosajona “*eddy covariance*” ampliamente extendida en vez de su traducción (covarianza eddy, covarianza de torbellinos, correlación de remolinos,...).

Esta tesis está centrada en la optimización de un PEM para la estimación de la GPP en el área de estudio (España, exceptuando a Canarias). Parte de su validación ha sido llevada a cabo a partir de estimaciones de GPP mediante métodos micrometeorológicos (torres *eddy covariance* para ser más exactos). Es por esto, que ambas metodologías se explican con más detalle en los apartados siguientes.

1.2.1. Determinación *in situ* de flujos de C mediante la técnica *eddy covariance* (EC)

El uso de esta técnica micrometeorológica de medida de intercambio neto de gases entre los ecosistemas y la atmósfera ha sufrido un incremento espectacular en los últimos años. El número de publicaciones ha pasado de unas pocas decenas en el periodo comprendido entre los años 1980-1995 a unas 3000 entre el periodo 1996-2013. Las medidas de EC [28] se han de realizar por encima de la superficie del ecosistema, más precisamente en la capa superficial, donde la turbulencia es más constante. Estas parcelas de aire en movimiento transportan $\text{H}_2\text{O}(\text{v})$ y otros gases (CO_2 , O_3 , NO_x , CH_4 , ...) a una cierta temperatura, cuya heterogeneidad dentro del ecosistema no afecta a la metodología EC. El método requiere de medidas de las componentes tridimensionales de la velocidad del viento (anemómetro sónico), la concentración del gas que se quiera estimar (*Infrared Gas Analyzer, IRGA*) y la temperatura del aire a una alta frecuencia temporal (10 Hz). Los datos son procesados empleando metodologías estandarizadas para asegurar la intercomparabilidad de las medidas [29], [30]. A pesar de la gran utilidad de la técnica EC, presenta problemas para la estimación de los flujos en superficies con pendientes significativas y superficies vegetales no homogéneas (incrementos significativos de vegetaciones leñosas), ya que pueden aumentar de forma apreciable la rugosidad de la superficie.

Con la técnica EC se estima el intercambio neto de CO_2 (NEE). La dinámica de la NEE es controlada a su vez por dos procesos antagónicos no separables directamente a partir de las medidas tomadas con EC: la fotosíntesis (GPP) y la respiración del ecosistema (R_e). La R_e se suele estimar

a partir de la relación $GPP \approx NEE - R_e$. Primero se obtiene la R_e nocturna directamente de la NEE (la GPP durante la noche es nula), posteriormente esta se corrige para estimarla durante el día y se calcula la GPP [29]. Así pues, las estimaciones de GPP mediante EC están sujetas a una doble fuente de error, la correcta estimación de la NEE y de la R_e .

1.2.2. Cálculo de flujos de C a partir de datos de satélite

Los datos de satélite proveen de observaciones espacialmente coherentes a escala global que pueden ser empleadas para la correcta monitorización de los flujos de carbono en los ecosistemas terrestres. Monteith¹ [1] fue el primer autor en proponer que existía una relación lineal entre la productividad de la vegetación (GPP y NPP) y la radiación solar absorbida fotosintéticamente activa (APAR). A la relación entre estas dos magnitudes se le denominó eficiencia de conversión de la radiación absorbida en materia seca (*Light Use Efficiency*, ε) y fue empleada en muchos modelos sencillos de crecimiento cultivos. Hay tres razones fundamentales que muestran el gran potencial que tienen los datos de satélite para su uso en la estimación de flujos de carbono: 1) La GPP está directamente relacionada con la energía solar que es absorbida por la planta (APAR). 2) La fracción de energía fotosintéticamente activa interceptada por la planta (*Fraction of Absorbed Photosynthetically Active Radiation*, f_{APAR}) se calcula a escala global y rutinariamente mediante datos de teledetección. 3) La relación de conversión entre APAR y masa de carbono interceptado por fotosíntesis disminuye su valor potencial debido a razones biofísicas y a forzamientos del entorno. Para caracterizar este hecho, en los modelos se calculan factores a partir de datos meteorológicos o de teledetección [31]. Así pues, el aspecto común en los PEMs utilizados en teledetección es que estiman la productividad como una función de la ε , la APAR y factores ambientales que inhiben la fotosíntesis ([31]). Los modelos actuales suelen estar basados tanto en

¹Dado que Monteith estableció el fundamento teórico para el desarrollo de los PEM, en la bibliografía se habla también, de forma genérica, del modelo de Monteith para referirse a los mismos.

datos meteorológicos como de satélite y las estimaciones son obtenidas a altas resoluciones espaciales y temporales. A continuación se muestran las ecuaciones fundamentales:

$$\text{GPP} = \text{APAR} \varepsilon \quad (1.2)$$

$$\text{APAR} = f_{\text{APAR}} \text{PAR} \quad (1.3)$$

$$\varepsilon = \varepsilon_{\text{max}} f(w) g(T) h(\text{VPD})\dots \quad (1.4)$$

donde la GPP en la unidad de tiempo t se suele proporcionar en [$\text{g m}^{-2} \text{t}^{-1}$], la PAR en [$\text{MJ m}^{-2} \text{t}^{-1}$], la f_{APAR} es la fracción de PAR absorbida (adimensional), ε_{max} es la eficiencia máxima potencial en el uso de la radiación [g MJ^{-1}]. Los términos $f(w)$, $g(T)$ y $h(\text{VPD})$ son factores adimensionales correctores (del contenido en agua de la vegetación, temperatura del aire y déficit de presión de vapor del aire, respectivamente, por citar algunos ejemplos) de la ε_{max} , comprendidos entre 0 (inhibición total de la fotosíntesis) y 1 (no inhibición).

Entre los factores que ejercen un mayor control sobre la GPP en diferentes ecosistemas cabe destacar principalmente dos por orden de importancia: 1) el índice de área foliar y la duración del periodo fotosintéticamente activo, y 2) el efecto de los factores ambientales que ejercen un control sobre la fotosíntesis. Mientras los primeros poseen una respuesta en el ecosistema más lenta y afectan directamente a la cantidad potencial de luz que puede absorber la cubierta, los segundos producen cambios más a corto plazo en la GPP actuando sobre la capacidad fotosintética a nivel de las hojas, ya que reducen su eficiencia en el uso de la radiación. En los PEM el primer factor viene representado por la variable f_{APAR} y el segundo por los diferentes factores adimensionales que modifican la ε_{max} de la cubierta vegetal. A continuación se muestra una tabla comparativa con las diferentes características de algunos de los modelos más utilizados mediante datos de satélite [32].

Tabla 1.1. Modelos de productividad a partir de datos de satélite. La tabla muestra tanto sus variables de entrada como los flujos obtenidos. Tabla adaptada de [32].

Modelos	Abreviatura	Resolución (espacial/temporal)	ϵ_{\max} (g MJ ⁻¹)	Escalares	Variables Estimadas
Carnegie Ames Stanford Approach Mode	CASA	0.5° /mes	0.389	T , AET, PET	NPP [33]
Global Production Efficiency Model	GLO-PEM	8 km/10 días	55.2 ^a , 2.76 ^b	T , S_M , VPD	NPP=GPP- R_a [34]
Vegetation Photosynthesis Model	VPM	1 km/8 días MODIS 1 km/8 días SPOT VGT	2.208 ^c , 2.484 ^d	T_s , W_s , P_s	GPP [35]
Parametric Production Efficiency Model	C-Fix	1 km/10-30 días	1.10	T , S_M , EF, CO ₂	NPP=GPP- R_a NEP=GPP- R_h [7]
Terrestrial Uptake and Release of Carbon Model	TURC	1° /mes	1.10	-	NPP=GPP- R_a [36]
Eddy Covariance Light Use Efficiency Model	EC-LUE	1 km/8 días	2.14 ^e	T_s , S_M	GPP [37]
Vegetation Index Model	VI	1 km/8 días	-	-	GPP [38]
Temperature and Greenness Model	TG	1 km/8 días	-	T_{EVI} , T_{LS}	GPP [39]
3-Physiological Principles Predicting Growth using Satellites	3-PGS	8 km /mes	1.8	VPD, Y , S_M , F	GPP NPP=0.45 GPP [40]
MODIS primary productivity products	MOD17	1 km/8-365 días	0.997-1.382 ^f	T_{\min} , VPD	NPP=GPP- R_a [6]

T temp. aire, S_M humedad del suelo, T_s temperatura del suelo, T_{LS} temperatura superficial de emisión, T_{EVI} índice de vegetación mejorado, EF fracción evaporativa, CO₂ factor de fertilización, W_s factor de agua, P_s factor fenológico de la hoja, T_{MIN} temperatura mínima del aire, F factor de congelación, a eficiencia cuántica de la fotosíntesis para plantas C3, b para plantas C4, c y d para bosques de coníferas de hoja perenne o tropicales, respectivamente, e para bosques, pastos y sabanas, f para los 11 tipos de cubierta MODIS, AET evapotranspiración actual y PET evapotranspiración potencial.

Otras aproximaciones emplean datos de teledetección y consideran únicamente relaciones con índices espectrales sensibles al contenido en clorofila de la cubierta vegetal y la radiación PAR en el caso de cultivos [41]. O también se plantea, cuando existan sensores capaces de estimar la débil señal de fluorescencia de la vegetación, el uso de esta como un indicador del estado fotosintético de la planta [42].

1.3. Objetivos del trabajo

En esta tesis se aborda la estimación de la GPP mediante la optimización del modelo de Monteith a través de la obtención de sus *inputs* especificándolos para las características de nuestro territorio. Aunque este tipo de modelo es conceptualmente sencillo, la forma en que se incorporan las diferentes capas de información que lo componen es una línea activa de investigación en la estimación de la producción de la vegetación. Así, mientras que algunos autores abogan por el uso exclusivo de datos de satélite, otros combinan estos datos con variables meteorológicas. Este trabajo explora ambas posibilidades y pretende integrar ambas fuentes de información con el fin de obtener imágenes de GPP lo más precisas y a la mayor resolución espacial y temporal posibles. Se va a probar y evaluar el potencial de diferentes alternativas a las variables que pueden componer el modelo. Las diferentes opciones van a ser validadas de forma directa con estimaciones de GPP obtenidas mediante la técnica (EC) y de forma indirecta mediante la intercomparación con otros productos de GPP.

A continuación se enumeran los objetivos fundamentales que se desean alcanzar en este trabajo:

1. *Cálculo de la GPP en España a partir de un PEM optimizado.* Se va a estudiar la mejora de un modelo PEM mediante la optimización de los parámetros que lo componen, siendo estos la f_{APAR} , la PAR y la ε . Los requerimientos espaciales y temporales deseados son, en nuestro caso, imágenes de GPP diarias a una resolución espacial de 1 km.

- a) f_{APAR} . Para la obtención de imágenes de f_{APAR} se va a aplicar la aproximación de Roujean y Bréon [12]. Puesto que algunos productos de satélite presentan huecos y unas resoluciones espaciales y temporales no coincidentes con nuestras necesidades, se desarrollarán técnicas de filtrado e interpolado que permitan disponer de series diarias refinadas de este parámetro biofísico.
- b) PAR. Se van a probar y analizar diferentes técnicas para la obtención de imágenes de PAR para, de esta forma, garantizar su disponibilidad. Esto incluirá su estimación a partir de la construcción de modelos empleando datos meteorológicos medidos en tierra y el uso de datos de satélite.
- c) ε . La eficiencia en el uso de la radiación aplicada en el PEM optimizado requiere fundamentalmente la obtención de dos grupos de variables. El primero es un mapa de eficiencia máxima en el uso de la radiación (ε_{\max}) que se desea construir a partir de mapa de coberturas del suelo optimizado para el área de estudio. Y el segundo grupo consiste en el cálculo de factores de estrés que incorporan al modelo el estado fotosintético actual de la vegetación. En estos factores se va a ensayar tanto la viabilidad del uso de parámetros dependientes de variables meteorológicas, como de otros indicadores estimados a partir de la respuesta espectral de la cubierta vegetal (datos de satélite).

2. Validación del producto de GPP optimizado.

- a) *Validación directa.* Se llevará a cabo a partir de la red de torres *eddy covariance* disponible en el área de estudio. Estas torres proporcionan un método alternativo a los PEM para calcular la GPP y son la referencia habitual para la validación de productos de satélite.

-
- b) *Validación indirecta.* Se realizará por comparación con productos de referencia de flujos de carbono a escala global. Los productos que se van a emplear para la intercomparación son el desarrollado por la NASA (MOD17A2) y el de Copernicus (*Dry Matter Productivity*, DMP).
- c) *Análisis del potencial explicativo de cada uno de los inputs empleados.* La disponibilidad de datos de torres EC permitirá evaluar el potencial explicativo de los *inputs* del modelo sobre datos *in situ*. Posteriormente se analizarán en base a las propias imágenes de GPP calculadas con el fin de estudiar la variabilidad espacial de la influencia de los parámetros de entrada.

Capítulo 2

Modelo optimizado para el cálculo de la GPP

Si bien, como se ha comentado, el modelo de Monteith constituye una base conceptual general, la aportación de esta tesis radica fundamentalmente en la mejora y refinado de las variables fundamentales que lo componen (f_{APAR} , PAR y ϵ) y que dan lugar al denominado, en el contexto de esta memoria, modelo optimizado para el cálculo de la GPP. A continuación se detallan las características de cada una de las variables del mismo.

2.1. Imágenes de f_{APAR}

La f_{APAR} (*Fraction of Absorbed Photosynthetically Active Radiation*) es una variable biofísica que está íntimamente relacionada con la productividad de las cubiertas vegetales a través de la fotosíntesis. Caracteriza la fracción de radiación solar que es absorbida por la vegetación y que es susceptible de ser empleada para la realización de la fotosíntesis, siendo además, un indicador del vigor de la misma.

A nivel experimental se calcula con la siguiente expresión [43]:

$$f_{\text{APAR}_{\text{inst}}} = \frac{(\text{PAR}_{\downarrow 1} - \text{PAR}_{\uparrow 1}) - (\text{PAR}_{\downarrow 2} - \text{PAR}_{\uparrow 2})}{\text{PAR}_{\downarrow 1}} \quad (2.1)$$

donde $f_{\text{APAR}_{\text{inst}}}$ hace referencia a la f_{APAR} instantánea, \downarrow hace referencia a radiación incidente, \uparrow hace referencia a radiación ascendente, 1 por encima de la cubierta y 2 por debajo de la cubierta. Esta expresión tiene en cuenta tanto la radiación reflejada por el suelo de nuevo hacia la cubierta, como la que es reflejada directamente sin penetrar a través de ella. El valor instantáneo de este parámetro irá variando a lo largo del día, tanto por el ciclo de variación diaria de la propia PAR como por la forma en que esta interactúa con la cubierta vegetal.

El parámetro de interés con el que se suele trabajar no es tanto esta $f_{\text{APAR}_{\text{inst}}}$ sino su valor promedio a lo largo del día (que denominamos a lo largo del trabajo como f_{APAR}). La viabilidad de su estimación mediante satélite queda patente por el amplio número de iniciativas y aproximaciones diferentes llevadas a cabo para los diferentes sensores disponibles: MODIS (*Moderate Resolution Imaging Spectroradiometer*) a bordo de Terra y Aqua [44], MERIS (*Medium Resolution Imaging Spectrometer*) en el satélite ENVISAT [45], SeaWiFS (*Sea-viewing Wide Field-of-view Sensor*) en Seastar/OrbView 2 [46] o SEVIRI (*Spinning Enhanced Visible and Infrared Imager*) en MSG [47]. Aún así, los productos de f_{APAR} de satélite presentan las dificultades asociadas a la estimación de parámetros biofísicos de forma remota, pudiéndose citar por ejemplo incertidumbres en las observaciones, cambios debidos a la geometría de observación e iluminación, influencia del suelo de fondo, limitaciones en las parametrizaciones empleadas y calibraciones de los algoritmos empleados.

El proceso de obtención de f_{APAR} mediante datos de satélite pasa por la inversión (calibrado) de un modelo, de forma que nos relacione la información que proporciona el satélite con la variable que se desea estimar. Básicamente, todas las aproximaciones que aparecen en la bibliografía pueden ser agrupadas según la forma en que se emplea la información proporcionada por el sensor: 1) si se emplean índices de vegetación o 2) si se emplea la

información espectral (reflectividad) directamente. El proceso de inversión consiste en la construcción de un modelo más o menos complejo (valores tabulados, lineal, cuadrático, redes neuronales, métodos kernel,...) a partir de medidas *in situ* o simulaciones mediante un modelo físico complejo de transferencia radiativa. Después de su calibrado, se aplica a datos de satélite para obtener las estimaciones de f_{APAR} con la mayor precisión posible.

2.1.1. Cálculo de la f_{APAR}

La f_{APAR} se ha calculado siguiendo la aproximación propuesta por *Roujean & Bréon* [12]. Esta metodología propone el uso del índice de vegetación RDVI (*Renormalized Difference Vegetation Index*) con una geometría de observación e iluminación específica óptima, que minimiza el efecto de las incertidumbres asociadas con la geometría y con el suelo de fondo.

$$\text{RDVI} = \frac{(\rho_{\text{IRC}} - \rho_{\text{R}})}{\sqrt{(\rho_{\text{IRC}} + \rho_{\text{R}})}} \quad (2.2)$$

donde ρ_{IRC} es la reflectividad en el infrarrojo cercano (841 – 876 nm) y ρ_{R} en el rojo (620 – 670 nm), ambas referidas a la geometría óptima de observación e iluminación como se ha indicado. Como muestran los autores, la geometría óptima se corresponde con el plano principal y ángulos cenitales $\theta_{\text{i}} = 60^\circ$ (iluminación) y $\theta_{\text{o}} = 45^\circ$ (observación) en la dirección de retrodispersión ($\Phi = 0$).

La función de distribución de reflectividad bidireccional (*Bidirectional Reflectance Distribution Function*, BRDF) es una magnitud no medible y que describe la distribución angular de la radiancia reflejada por una superficie natural [48]. Con el fin de obtener una descripción de dicha función mediante datos de teledetección, en la literatura aparecen diferentes aproximaciones para la caracterización de la misma con un número reducido de parámetros ajustables. A estas aproximaciones se les denominan modelos empíricos o semiempíricos de la BRDF. El conocimiento de los parámetros del modelo para un píxel determinado permite obtener, por tanto, estima-

ciones de los valores de la reflectividad para una geometría de observación e iluminación deseada. En este trabajo se ha empleado el modelo semiempírico de la BRDF del producto MODIS (MCD43A1), compuesto por los satélites Terra y Aqua para una mejor inversión de los parámetros que los describen. Depende de tres parámetros para cada banda espectral (k_0 , k_1 y k_2) y cada uno de ellos describe características diferentes (geométricas, volumétricas,...) de la reflectividad que compone la escena. Posteriormente se ha calculado el RDVI a partir de las reflectividades de acuerdo con la expresión anterior (2.2)

La f_{APAR} se calcula a partir del RDVI empleando la ecuación siguiente que maximiza la correlación entre ambos [12]:

$$f_{\text{APAR}} = 1.81 \text{ RDVI} - 0.21 \quad (2.3)$$

Esta relación se ha obtenido considerando una superficie situada a diferentes latitudes norte (0° , 15°N , 30°N , 45°N , 60°N , 75°N) y estaciones (los dos solsticios y equinoccios) con el fin de obtener unas estimaciones más generales para un uso a escala global.

Esta metodología forma parte de la cadena de procesado de los productos operacionales de vegetación de *Satellite Applications Facility on Land Surface Analysis* (LSA-SAF). Ha sido validada de forma directa a partir de medidas *in situ* y de forma indirecta comparando con otros productos operacionales [49]. Los resultados han mostrado unos valores intermedios frente a estimaciones de diferentes productos (figura 2.1).

En este trabajo se obtienen estimaciones de f_{APAR} con una resolución espacial superior a la del producto oficial disponible operacionalmente en LSA-SAF, dado que se aplica a datos MODIS de 1 km.

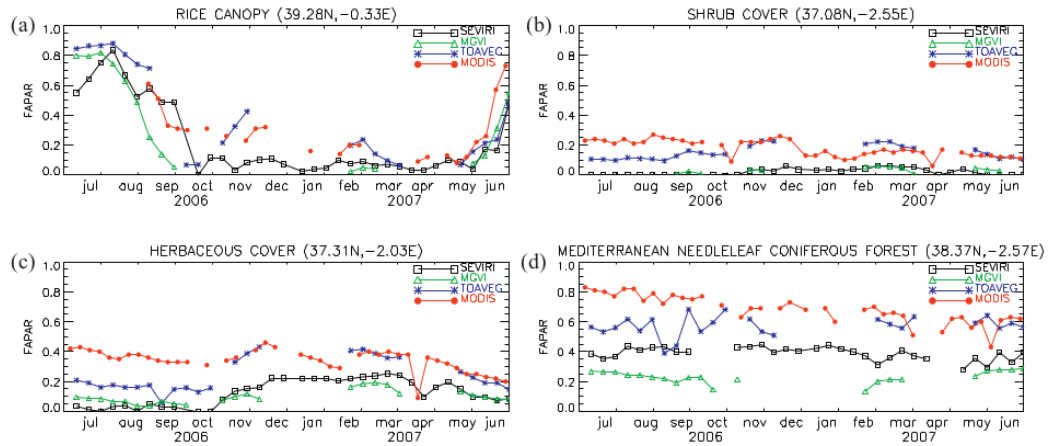


Figura 2.1. Perfiles de f_{APAR} para cuatro tipos de cubierta vegetal y cuatro productos operacionales: el algoritmo Roujean y Bréon de LSA-SAF (negro), inversión mediante redes neuronales y datos MERIS (azul), índice de vegetación optimizado con datos MERIS (verde) y el producto oficial de MODIS (rojo). Fuente: *Martínez et al.* [49].

2.1.2. Postprocesado de las series obtenidas de f_{APAR}

Las series temporales de f_{APAR} presentan variaciones no deseadas que pueden ser atribuidas, entre otras razones, a interferencias atmosféricas (contaminación de nubes, ozono, polvo, sombras...), variaciones radiométricas (asociadas con cambios en la geometría de iluminación y observación) o ruido intrínsecamente relacionado con las características técnicas del propio sensor de adquisición [50], [51]. Los efectos atmosféricos y radiométricos tienden a afectar de forma diferente a las distintas bandas espectrales así, por ejemplo, suelen producir un descenso de la reflectividad en el infrarrojo cercano en mayor medida que en el rojo, provocando caídas drásticas en índices de vegetación asociados con estas bandas espectrales, como es el caso del RDVI empleado en el cálculo de la f_{APAR} [52], [53]. En el caso de que estos eventos persistan en el tiempo, pueden impedir la obtención de estimaciones produciendo la aparición de huecos en las series temporales. La metodología habitual soluciona este problema sustituyendo los huecos mediante técnicas de interpolación, ya sea mediante información de píxeles vecinos (espacial) o mediante datos de fechas próximas (temporal). Las metodologías de filtrado de ruido de series temporales que aparecen en la

bibliografía pueden ser agrupadas como:

- Métodos locales basados en el rango y ajustes lineales o polinómicos (filtros de mediana, métodos de regresión local pesada *loess*, etc).
- Métodos globales mediante el ajuste de modelos o descomposición de la señal (funciones gaussianas asimétricas, transformadas de Fourier, etc).

Los métodos globales presuponen una cierta forma *a priori* de la evolución temporal y no son eficaces cuando la serie es irregular o asimétrica [54]. En esta tesis las series temporales de f_{APAR} han sido filtradas mediante el método polinómico *loess* [55]. Los métodos locales se caracterizan por poseer una mayor versatilidad y permiten incluir, en el caso del *loess* utilizado, información auxiliar acerca de la calidad de los datos. En este trabajo se ha empleado un polinomio de quinto grado que es consistente con el número de observaciones en la ventana de trabajo (17), correspondiente con un periodo temporal de 128 días. Valores más altos podrían sobreajustar la serie temporal llegando a producir resultados incluso más ruidosos que los originales [56]. La información auxiliar (*Quality Flags*, QF) empleada ha sido extraída del producto MCD43A2. Estos datos los suministra la NASA sobre los parámetros de la BRDF (MCD43A1) involucrados en el cálculo de la f_{APAR} . La QF consiste en un indicador de calidad para cada una de las bandas involucradas y su rango de variación va de 0 a 6: 0 indica la mejor calidad (inversión de parámetros con todas las fechas y la mayoría de valores de RMSE bajos), 1 indica buena calidad (inversión de parámetros con todas las fechas), 2 indica una calidad baja (inversión de parámetros con un número de observaciones mayor de 6) y 4 (el peor caso) indica que para la inversión de parámetros se han empleado un número de observaciones comprendidas entre 3 y 6 (poco fiable).

Las ecuaciones para el cálculo de los pesos introducidos en el método de filtrado son las siguientes:

$$w_i^* = \frac{1}{0.5 \text{ QF}_i + 1} \quad (2.4)$$

$$w_j = w_j^* \left(\frac{1 - |t_i - t_j|}{\Delta t} \right) \quad (2.5)$$

$$w'_j = w_j^* \left(\frac{1}{1 + \frac{|\Delta y_j|}{S \sigma}} \right) \quad \text{si } \Delta y_j < 0 \quad (2.6)$$

donde w_i^* es el peso asociado a la estimación atendiendo a la información auxiliar, QF_i es la información auxiliar acerca de la f_{APAR} , t_i es el valor de tiempo asociada al valor que se quiere suavizar, t_j son los valores de tiempo de los vecinos de la fecha a suavizar dentro de la ventana temporal de suavizado, $d(t)$ es la semi-anchura temporal de la ventana de suavizado, Δy_j es la diferencia entre el dato de la curva de ajuste (y_j^*) y el dato original (y_j), σ es la desviación estándar de los valores negativos de las diferencias Δy y S es un parámetro (0.1 en este trabajo) para el reajuste de los pesos. En este caso QF varía entre 0 y 6 ya que se calcula como la suma de dos bandas. El proceso de asignación de pesos se realiza en dos iteraciones. Primeramente se realiza el filtrado únicamente con la información de 2.5 y, en segundo lugar, se reajustan los pesos obtenidos de la ecuación 2.4 decreciendo sus valores de acuerdo con 2.6, ya que los datos originales están por debajo de la recta ajustada. Estos nuevos pesos se emplean finalmente en 2.5 para la segunda iteración dentro de la ventana temporal de ajuste para cada dato. Esto favorece que la curva ajustada considere como mejores los valores más altos frente a los más bajos, combinando así información auxiliar junto con una tendencia a capturar la envolvente superior, ya que el ruido tiende a decrecer el valor de la f_{APAR} [52], [53].

La principal ventaja de esta metodología de filtrado es la inclusión de información auxiliar mediante el QF en los pesos del *loess*, así como su corrección posterior para favorecer la adaptación de la curva ajustada a la envolvente. La figura 2.2 muestra algunos ejemplos que ponen de manifiesto la idoneidad del método de filtrado aplicado, así como su capacidad para eliminar los huecos en las imágenes (figura 2.3).

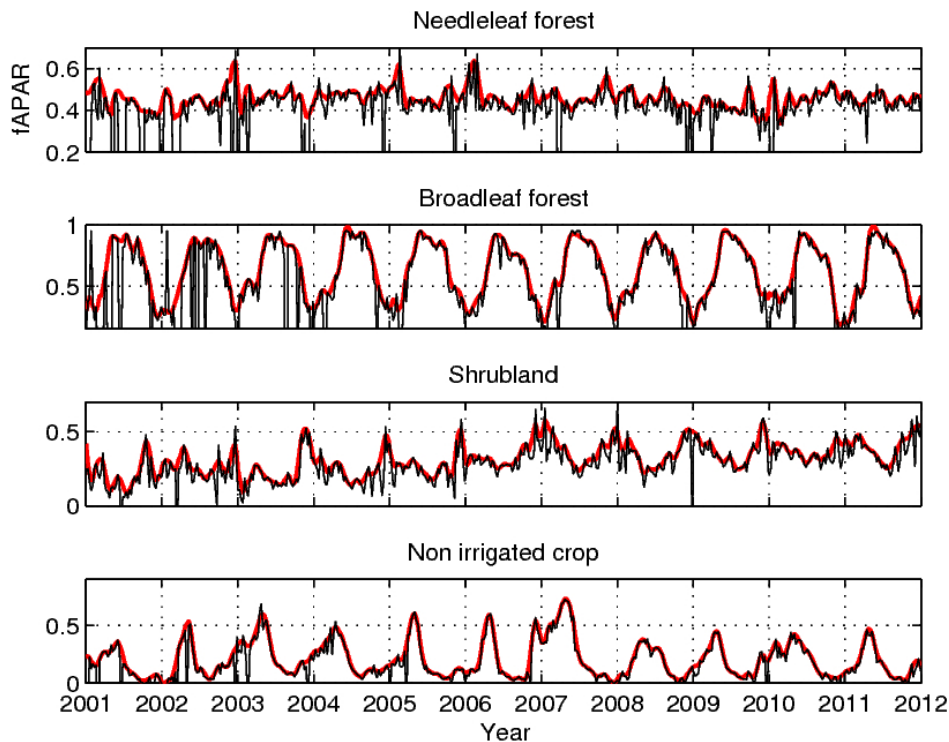
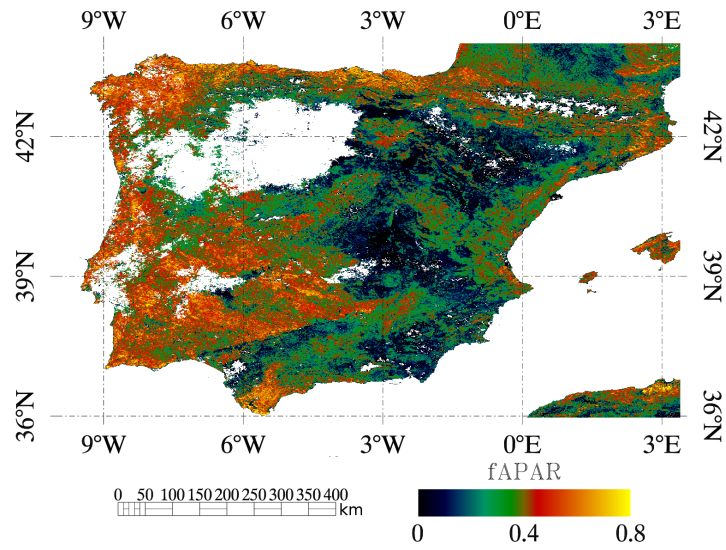
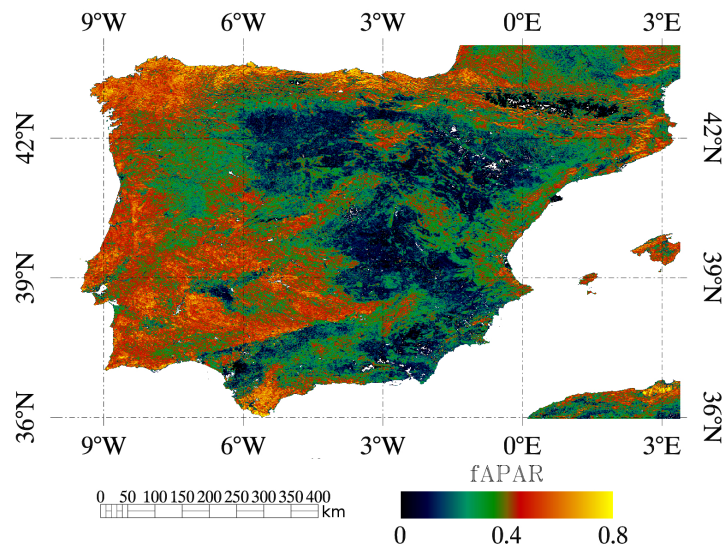


Figura 2.2. Perfiles de f_{APAR} para cuatro tipos de cubierta vegetal. El perfil de color rojo muestra la serie temporal filtrada y reconstruida mediante el método *loess* adaptado y en color negro se muestran los perfiles originales de f_{APAR} .



(a)



(b)

Figura 2.3. Imagen de f_{APAR} original (a) y de f_{APAR} filtrada y reconstruida mediante el método *loess* adaptado (b). Corresponde con la fecha 1/1/2011.

Para evaluar la eficacia del método *loess* adaptado, se ha intercomparado con otros que habitualmente se utilizan en el procesado de datos de satélite [13]: el método de composición de máximos (*maximum value composite*, MVC), un método lineal (media móvil, MA) y otro método polinómico (Savitzky-Golay, SG). De forma resumida se indican algunas características de dichos métodos:

- Compuesto de máximo valor (MVC) [53]: Consiste en tomar el máximo valor dentro de una ventana temporal de anchura prefijada (en nuestro caso 48 días). Su fundamento recae en el hecho de que los efectos atmosféricos siempre influyen negativamente a los índices de vegetación.
- Media móvil (MA): Considera como valor óptimo el promedio de los valores dentro de una ventana temporal de anchura prefijada (en nuestro caso 128 días). Este método preserva el área de la curva de la fenología pero puede llegar a alterar su altura y su anchura [57].
- Savitzky-Golay [58]: Es un método de regresión local (como el *loess*) donde los coeficientes del filtro son obtenidos mediante mínimos cuadrados. Este método es muy eficiente computacionalmente porque sus coeficientes están calculados previamente para un grado dado (un quinto grado en este trabajo). Esto reduce el coste computacional ya que únicamente se realiza una convolución con los coeficientes prefijados en una ventana temporal seleccionada (en nuestro caso 128 días).

Para la estimación de los errores de las distintas metodologías y crear series suavizadas e interpoladas de datos de f_{APAR} , se han simulado escenarios similares a los reales considerando múltiples parámetros de ruido, diferentes porcentajes de datos afectados por ruido y huecos. Se han extraído perfiles típicos partiendo de datos MODIS, a partir de los cuáles se han seleccionado nueve zonas cubriendo diferentes tipos de cubiertas vegetales (según el mapa híbrido mencionado anteriormente) existentes en el área de estudio. Para asegurar la homogeneidad y minimizar problemas asociados al mal solapamiento y corregistro de estos perfiles, las nueve zonas presentan mínima varianza temporal de las series de f_{APAR} dentro de una ventana de

3 x 3 píxeles. Siendo la curva típica de cada zona un promedio de los valores diarios de máxima calidad (QF=0) en una serie de 12 años de datos. Por último, se han replicado las series típicas para disponer de 12 años de 9 curvas de diferentes clases de cubierta y poder extraer estadísticos con alta significación.

La premisa básica de este trabajo de evaluación de calidad de los métodos de filtrado es que la única fuente de variación indeseada en las series es debida al ruido, y por esto, se ha sumado a los datos un porcentaje variable de ruido gaussiano (entre el 10 % y el 100 % de los datos) con diferentes valores de la media (entre -0.01 y -0.05) y de desviación estándar (entre 0.01 y 0.05).

En la figura 2.4 (a) se muestran los perfiles reconstruidos por los diferentes métodos de filtrado con respecto a una serie típica (mezcla de bosques) calculada a partir de datos MODIS. Los valores de los parámetros del ruido gaussiano aplicados sobre un 25 % de los datos sobre el perfil típico son: media -0.05 y desviación estándar 0.05. En la figura 2.4 (b) se muestra el RMSE sobre esta misma área pero con todos los valores de los parámetros del ruido para poder analizar su evolución. El porcentaje de datos afectados por ruido permanece constante en el 25 %. Estos resultados ponen de manifiesto la superioridad del método *loess* frente al resto, especialmente frente al MVC y MA.

En la figura 2.5 se muestra la evolución de los sesgos en función del porcentaje de datos afectados por el ruido. Se han fijado la media (-0.025) y la desviación estándar (0.025) del ruido, variando la proporción de datos afectados por error. Esta figura indica que el método MVC tiende siempre a sobreestimar el valor real de la f_{APAR} y la media móvil (MA) y el SG subestiman de forma proporcional al porcentaje de errores en la serie, indicando su incapacidad para capturar adecuadamente la envolvente. El método *loess* obtiene los mejores resultados, pero cuando los errores afectan a más del 50 % de los datos comienza a manifestar un sesgo creciente y negativo.

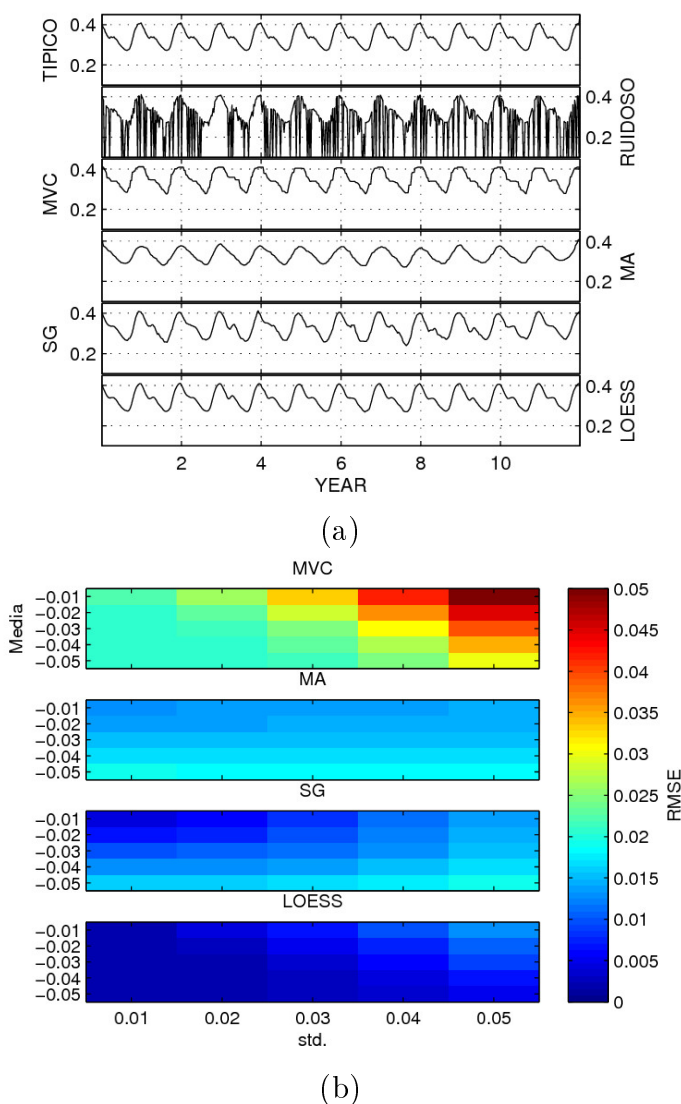


Figura 2.4. Mosaico de perfiles de f_{APAR} sin/con ruido para el área clasificada como mezcla de bosques y los resultados del filtrado con los diferentes métodos considerados (a). RMSE para el mismo área considerando todos los valores de los parámetros del ruido considerado (b).

Los resultados de la intercomparación entre los diferentes métodos de filtrado ha puesto de manifiesto que:

- Tanto el método MVC como el de media móvil pueden llegar a empeorar las estimaciones en presencia esporádica (frecuencias bajas de aparición) de ruido en la serie.

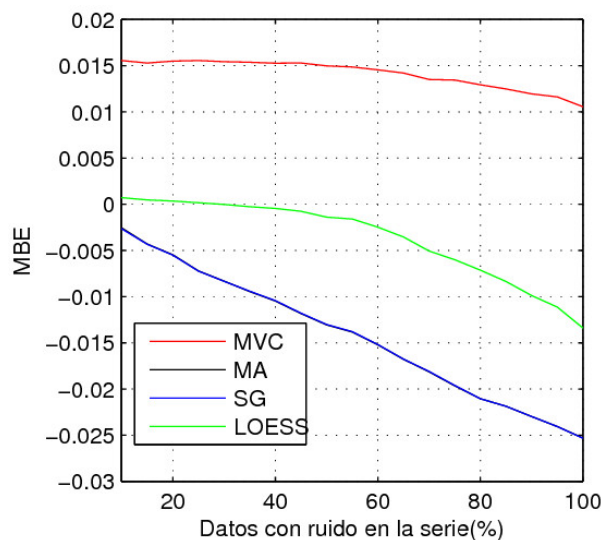


Figura 2.5. Variación del MBE en función del porcentaje de datos afectado por el ruido

- Es importante remarcar que MVC introduce cierto sesgo en las estimaciones, asociado principalmente con la presencia de errores positivos en la serie que son considerados como óptimos y con valores sobrestimados en los periodos de bajos valores de la f_{APAR} que no son capturados porque dentro de la ventana temporal de trabajo del filtro hay valores superiores.
- Los métodos SG y *loess* obtienen en general mejores resultados ya que permiten una mayor versatilidad (son no lineales) para adaptarse a perfiles fenológicos más complejos.
- Cuando se estudian los métodos en series temporales con ruido y huecos se ha observado que en general el uso combinado de filtrado e interpolación lineal simple tiende a producir series excesivamente suavizadas indicando que se destruye información relevante.

- El método *loess* ha obtenido los mejores resultados y esto puede deberse a dos razones fundamentalmente: 1) que ha sido la mejor técnica de filtrado de acuerdo con los resultados únicamente en presencia de ruido y 2) que interpola de forma más conveniente.
- La razón del mejor interpolado es atribuible a que en los métodos MVC, MA y SG es lineal mientras que en el método *loess* es no lineal.
- La inclusión de información auxiliar mediante el QF en los pesos del *loess*, así como su corrección posterior para favorecer la adaptación de la curva ajustada a la envolvente, permite obtener los mejores resultados. Esto indica que para el filtrado de series temporales de f_{APAR} , donde el ruido tiene unas ciertas características determinadas, conviene el diseño de herramientas de filtrado adaptadas como ha sido el caso del *loess*.

2.2. Imágenes de PAR

La radiación solar es la fuente de energía que permite que se desarrollen la mayor parte de los procesos físicos, químicos y biológicos de la superficie terrestre. Esta no solo controla la temperatura y el perfil de humedades del suelo, sino que proporciona la energía necesaria para la realización de la fotosíntesis. Observando la ecuación fundamental de los PEM (ecuación 1.2) se aprecia que es de gran importancia la obtención de imágenes precisas de PAR, ya que esta variable juega un papel fundamental en el cálculo de la GPP. La fotosíntesis, y la fijación de carbono por tanto, está controlada por diversos factores entre los que se encuentra la distribución y características de la PAR que llega a las plantas [59]. Así, por ejemplo, la forma en que las plantas aprovechan la radiación incidente (convierten la PAR en energía química) es diferente en función de la cantidad de energía luminosa disponible [19]. Bajo condiciones de poca iluminación, la conversión de PAR es aproximadamente lineal, mientras que para irradiaciones altas, la fotosíntesis satura, es decir, no responde a cambios en la irradiación (figura 2.6).

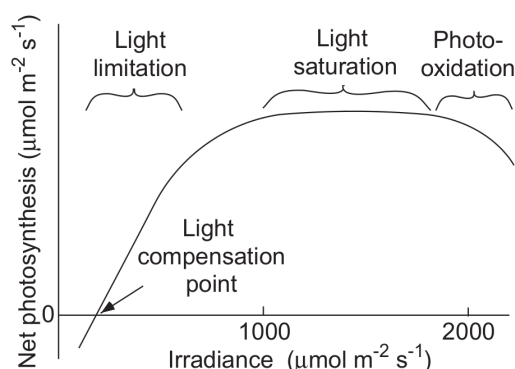


Figura 2.6. Relación entre la tasa fotosintética neta y la irradiancia fotosintéticamente activa mostrando los diferentes procesos que limitan la fotosíntesis a diferentes niveles de irradiancia.

La proporción entre irradiancia global y la PAR es alrededor de 0.5, aunque es sabido que varía con la turbiedad atmosférica, la elevación solar y el contenido en agua de la atmósfera. *Ross & Sulev* [60] apuntan que existe un rango de variación de 0.42 – 0.49, principalmente debido a las diferentes proporciones entre radiación directa y difusa. Esta proporción se mantiene en diferentes áreas de estudio. En este trabajo se va a considerar el valor intermedio de 0.46 [11].

Dado que existe una red de estaciones que miden de forma rutinaria la irradiancia solar diaria, se puede plantear la obtención de las imágenes de PAR directamente a partir de las medidas de estas estaciones meteorológicas mediante su interpolación espacial. El problema que presenta esta metodología es que únicamente una pequeña fracción de las estaciones meteorológicas disponibles miden la irradiancia global. Esta baja densidad espacial en las observaciones no permite una óptima interpolación espacial de los datos. Ante esta problemática, en la bibliografía aparecen diferentes alternativas para estimar la irradiancia solar a partir de otras variables más densamente muestreadas empleando diferentes modelos. Estos pueden ser clasificados en tres grupos:

Modelos deterministas (físicos). Se basan en leyes físicas que describen los procesos de interacción entre la radiación solar y los constituyentes atmosféricos que afectan a la radiación a su paso por la atmósfera

hasta la superficie terrestre [61]. Requieren información de la que no se dispone habitualmente en las estaciones de medida debido a la complejidad de los fenómenos que rigen la física atmosférica.

Modelos conceptuales. Incorporan los mecanismos físicos que describen las interacciones entre la radiación y los componentes atmosféricos, pero de una forma simplificada. Cada uno de los elementos que componen estos modelos da cuenta de los fenómenos importantes que intervienen en el sistema que se quiere describir.

Modelos paramétricos. Estos modelos, conocidos también como analíticos o empíricos, se construyen mediante ajustes de expresiones matemáticas entre valores de radiación y datos de diversas variables meteorológicas sin un modelo explícito de las interacciones entre la radiación solar y los componentes atmosféricos.

En esta tesis se han desarrollado y validado diversos modelos paramétricos. Así mismo, se ha validado y mejorado un producto operacional a partir de datos del satélite MSG. Fruto de este trabajo se adjuntan dos publicaciones abordando los dos métodos (apéndices A y B). La ventaja de disponer además de un modelo empírico calibrado y validado para el área de estudio es que, en ausencia de datos de satélite, podemos obtener imágenes de PAR. Hay que remarcar que las series de datos de satélite son más cortas y más recientes, impidiendo la posibilidad de realizar estudios retrospectivos. Como se ha mencionado, los modelos empíricos desarrollados sí permiten este tipo de estudios, ya que la irradiación se estima a partir de datos de estaciones meteorológicas de las que se disponen largas series históricas (más de 100 años).

2.2.1. Obtención a partir de datos de estaciones meteorológicas

El número de estaciones que miden la irradiación global diaria en el área de estudio es mucho más reducido que el de estaciones de medida de temperatura y precipitación (del orden de 1:10). Esto impide obtener imágenes de radiación espacializando directamente los datos medidos, ya que los estimadores de la calidad del interpolado indican errores demasiado grandes (las nubes presentan una alta variabilidad espacial en el área de estudio). Como solución a este problema, numerosos trabajos estiman la irradiación a partir de modelos que emplean otras variables más extensivamente muestreadas, como la precipitación, temperatura del aire, horas de insolación,... [62], [63]. Los métodos seleccionados son: el modelo clásico de Bristow-Campbell (BC) y dos modelos no lineales avanzados, el método de regresión sesgada mediante *kernels* (*Kernel Ridge Regression, KRR*)¹ y redes neuronales artificiales (RNAs). El método de Bristow-Campbell [64] consiste en calibrar una ecuación empírica en el área de estudio para estimar la irradiación solar a partir de la amplitud del ciclo diario de temperaturas (diferencia entre la temperatura máxima y mínima diaria). Por su parte, los métodos no lineales avanzados como el KRR y las RNAs son técnicas de aprendizaje máquina que presentan las siguientes características [65],[66]:

- No necesitan un conocimiento *a priori* del problema. Esto tiene una gran relevancia cuando se intenta modelizar sistemas no lineales, ya que no sabemos la forma en que se relacionan las diferentes variables consideradas en el problema.
- Son capaces de encontrar relaciones en problemas con una alta dimensionalidad.
- Presentan un cierto grado de inmunidad al ruido en los datos. Este es un escenario bastante habitual y puede producirse por diferentes causas: errores en la adquisición de los datos, errores en la transcripción de los datos, al cambiar de formato, etc.

¹En este trabajo se ha mantenido el acrónimo anglosajón KRR dado el poco uso de este término en castellano.

- Son capaces de capturar la naturaleza de fenómenos como la histéresis, que son difícilmente modelizables mediante una aproximación lineal.

El procedimiento llevado a cabo para la obtención de las estimaciones fue el siguiente:

1. *Intercomparación de los modelos seleccionados para estimar la irradiación solar global.* Se obtuvieron modelos que permitieran relacionar la irradiación global con variables extensivamente muestreadas permitiendo, por tanto, una espacialización adecuada. Más concretamente se consideraron las variables temperatura máxima, temperatura mínima e irradiación solar global extraterrestre para los tres modelos indicados. El uso de diferentes técnicas de modelizado permitió realizar, como objetivo específico de este apartado, una intercomparación de diferentes modelos paramétricos y semi-paramétricos de estimación de la irradiación global.
 - a) *Análisis de las variables de entrada mediante el test gamma.* Esta técnica [67] permitió evaluar el poder explicativo de diferentes variables susceptibles de ser empleadas para la construcción de modelos. La posibilidad de incluir nuevas variables de forma sencilla en los modelos de RNA y KRR motivó la aplicación de esta metodología para, de esta forma, considerar variables adicionales como la precipitación. Esta metodología permitió realizar un análisis de la importancia de las diferentes variables meteorológicas extensivamente muestreadas.
 - b) *Calibrado y evaluación de modelos.* Los parámetros de los modelos BC, RNA y KRR se ajustaron (calibrado) al área de estudio a partir de un conjunto de datos de entrenamiento. En el caso del modelo BC se empleó la técnica de mínimos cuadrados, y para los modelos RNA y KRR se utilizó un conjunto de datos extra (validación) para su obtención, ya que el método de calibrado de estos modelos requiere de un conjunto adicional de datos. Las tres metodologías se evaluaron y compararon en base a criterios objetivos sobre un conjunto de datos de test idéntico.

- c) *Espacialización de las variables de entrada.* Con el fin de obtener imágenes de irradiación global diaria, los parámetros de entrada de los modelos fueron espacializados. Este problema se abordó mediante la técnica geoestadística de krigado [68]. Las características exigidas a este procedimiento para la obtención de imágenes continuas de temperatura máxima, temperatura mínima y precipitación fueron una resolución espacial de 1 km con el mínimo error posible empleando la técnica de validación cruzada.
2. *Aplicación del mejor modelo para derivar imágenes de PAR.* El mejor modelo seleccionado se aplicó sobre datos espacializados para producir imágenes diarias de irradiación global y PAR para la zona de estudio en el periodo de 2005 a 2008 (ambos inclusive). Las imágenes de irradiación se derivaron de forma directa aplicando el modelo ajustado a las imágenes de las variables de entrada. A partir de las imágenes de irradiación se obtuvieron las imágenes de PAR. Se muestra un ejemplo de los resultados obtenidos en la figura 2.7.

Los resultados obtenidos (apéndice A) indican que el modelo BC muestra una tendencia a sub-estimar la irradiación sobre todo para sus valores más elevados. El modelo KRR proporciona resultados bastante buenos, aunque presenta una precisión ligeramente menor a la del modelo RNA. El comité de RNAs (promedio de las estimaciones de las 6 mejores RNAs) ha sido la técnica que mejores resultados ha obtenido. Ha conseguido modelizar la irradiación global diaria con un error MAE de $2.3 \text{ MJ m}^{-2} \text{ day}^{-1}$. Las RNAs y KRR son modelos no lineales con una alta capacidad para describir relaciones complejas entre las variables consideradas. El carácter marcadamente no lineal de las relaciones entre la irradiación global diaria y las variables meteorológicas consideradas, además de incluir adicionalmente la variable precipitación (como indica el test gamma), hace que ambas metodologías obtengan mejores resultados con respecto al modelo clásico BC.

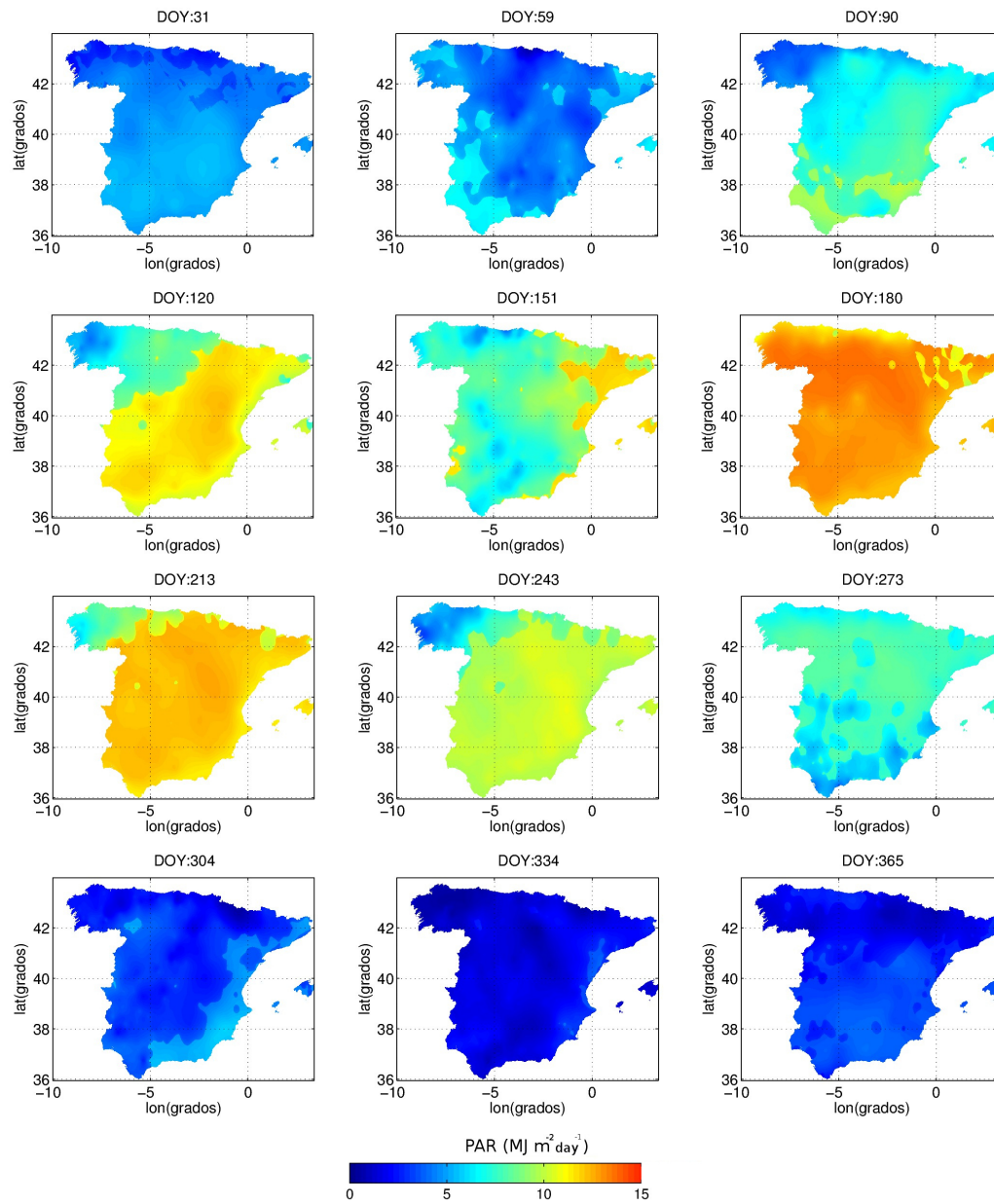


Figura 2.7. Imágenes diarias de PAR para diferentes fechas (una de cada mes) a lo largo del año 2008.

2.2.2. Obtención a partir de modelos con datos de satélite

Existen diversos métodos que permiten la estimación de la irradiación solar a partir de datos de satélite. La mayoría de estas metodologías son híbridas y combinan métodos físicos con empíricos. La cantidad de radiación que llega a la superficie terrestre está principalmente determinada por las condiciones atmosféricas, así que cuando son conocidos los componentes atmosféricos como el ozono, vapor, aerosoles y las nubes, la irradiancia instantánea puede ser calculada mediante un modelo de transferencia radiativa. Sin embargo, la información capturada por el sensor del satélite es una mezcla del estado de la atmósfera y la superficie terrestre y debe de ser desacoplada para la estimación de la irradiación.

El equipo científico de MODIS proporciona operacionalmente productos de componentes atmosféricos que son susceptibles de ser empleados para la estimación de la irradiación mediante algoritmos sencillos [69]. Estos productos contienen importantes incertidumbres dado que se estiman para una única escena adquirida por el sensor y presentan grandes huecos sin procesar, que requieren ser interpolados o estimados por otros métodos. Sin embargo, la mayoría de métodos para la estimación de la irradiación emplean imágenes de satélites geostacionarios, ya que presentan una alta resolución temporal (permitiendo obtener múltiples escenas en un mismo día) y una resolución espacial de hasta 1 km. Algunos algoritmos puramente empíricos, como el de Heliosat [70], estiman la irradiación mediante el cálculo de un parámetro normalizado como es el índice de nubosidad a partir de las medidas de radiancia registradas por el sensor del satélite. Este modelo ha evolucionado a Heliosat 2 [71], en el que se incorporan parámetros físicos de la atmósfera para mejorar la calidad de las estimaciones. Otros modelos, como el DLR-SOLEMI [72], estiman la irradiancia directa normal (*Direct Normal Irradiance*, DNI) a partir de datos de Meteosat y empleando el modelo de Bird [73] para cielos despejados. Los diferentes parámetros atmosféricos se obtienen a partir de otros satélites meteorológicos y otras fuentes como TOMS (*Total Ozone Mapping Spectrometer*), CDC-NOAA (*National Oceanic and Atmosphere Administration-Climate Diagnostic Center*) para

el vapor de agua y NASA-GACP (*Global Aerosol Climatology Project*). La precisión de las estimaciones horarias a partir de datos de satélite está comprendida entre el 17-25 % de RMSE y entre el 10-15 % para las estimaciones diarias.

En esta tesis se ha empleado el producto operacional de irradiancia de onda corta que suministra el LSA-SAF. Además se ha validado en el área de estudio y se ha propuesto una mejora de resolución espacial (publicación adjunta a la tesis, apéndice B) con respecto al producto original.

El algoritmo de LSA-SAF emplea datos de Meteosat Segunda Generación (MSG) y dos parametrizaciones de la atmósfera diferentes para cielos cubiertos y descubiertos. MSG es la generación actual de satélites meteorológicos geostacionarios europeos y para realizar sus observaciones emplea el sensor a bordo SEVIRI (*Spinning Enhanced Visible and Infrared Imager*). Las características técnicas de este sensor, especialmente la alta resolución temporal, permiten monitorizar su disco terrestre de observación (Europa, África y Sur América) con sus 12 canales espectrales cada 15 minutos con una resolución nominal de 3 km, excepto para el canal visible de banda ancha que es de 1 km. La alta resolución temporal es de especial interés para poder capturar el ciclo diario de la radiación con precisión. Su algoritmo de estimación parte de la premisa de que los factores más influyentes en la transmisividad atmosférica son la altura solar y la cobertura nubosa. Esta alta frecuencia de adquisición de imágenes permite un seguimiento muy preciso de la posición de las masas nubosas, así que la máscara de nubes que determina el modelo de transmisividad a emplear es una pieza clave en las estimaciones del producto de LSA-SAF [74]. Para condiciones de cielo descubierta, la transmisividad se calcula a partir de las concentraciones de la composición del aire (vapor de agua, O₃, CO₂, O₂ y aerosoles), el albedo de la superficie y el albedo esférico de la atmósfera. Por otro lado, en los píxeles correspondientes con cielos cubiertos, se emplea un modelo de transferencia radiativa sencillo entre las nubes, la atmósfera y la superficie. La información necesaria para conocer los componentes de la atmósfera se obtienen a partir del *European Centre for Medium Range Weather Forecasts* (ECMWF) y el TOMS.

El producto de irradiancia de onda corta (48 imágenes por día aproximadamente) se ha integrado para obtener finalmente la densidad de flujo de radiación que incide a lo largo del día en cada píxel de la escena. Esta metodología se ha aplicado a todos los días del año 2011. La validación de este producto operacional ha sido realizada en dos redes de estaciones meteorológicas de medida, una menos densa (45 estaciones) a escala de todo el país (Agencia Estatal de Meteorología, AEMet) y otra mucho más densa (140 estaciones) en Cataluña (*Servei Meteorologic de Catalunya*, SMC). Los estadísticos resultantes de la validación de las imágenes de irradiación con los datos de la AEMet dan como resultado un sesgo de $-0.12 \text{ MJ m}^{-2} \text{ day}^{-1}$ y una diferencia absoluta de $1.0 \text{ MJ m}^{-2} \text{ day}^{-1}$ respecto a los datos medidos *in situ*. Esto indica una buena calidad de las estimaciones del producto del LSA-SAF, que mejora sustancialmente las estimaciones a partir de métodos empíricos ($2.3 \text{ MJ m}^{-2} \text{ day}^{-1}$) con datos meteorológicos. A pesar de esto, el análisis de las estimaciones sobre zonas topográficamente más complejas (datos SMC), reveló sesgos significativos y crecientes en función de la elevación del terreno. En la figura 2.8 se muestra mediante un diagrama de cajas la dependencia de los residuos con la elevación y la dispersión entre las estimaciones de satélite y las medidas *in situ*.

La corrección adicional en función del espesor óptico de la atmósfera teniendo en cuenta la elevación redujo del 5 % al 0.5 % estos sesgos, indicando la necesidad de esta mejora en terrenos accidentados. Aparte de la corrección de elevación, se ha propuesto una metodología operacional para incrementar la resolución espacial del producto a 1 km. El método consiste en:

- 1) Corrección de la línea de costa e interpolado espacial del producto original.
- 2) Corrección de elevación y validación.
- 3) División de la irradiación en sus diferentes componentes (directa, difusa y reflejada).
- 4) Corrección topográfica por medio de un modelo digital de terreno de las diferentes componentes, según proceda.

Aunque los resultados son coherentes y reflejan cualitativamente los efectos topográficos esperados sobre la irradiación, esta corrección topográfica no ha podido ser validada ya que las redes de estaciones colocan sus sensores de medida horizontalmente y no de acuerdo con la pendiente del terreno donde se ubican. En la figura 2.9 se muestra un ejemplo de las estimaciones

de irradiación de satélite corregida de los efectos topográficos.

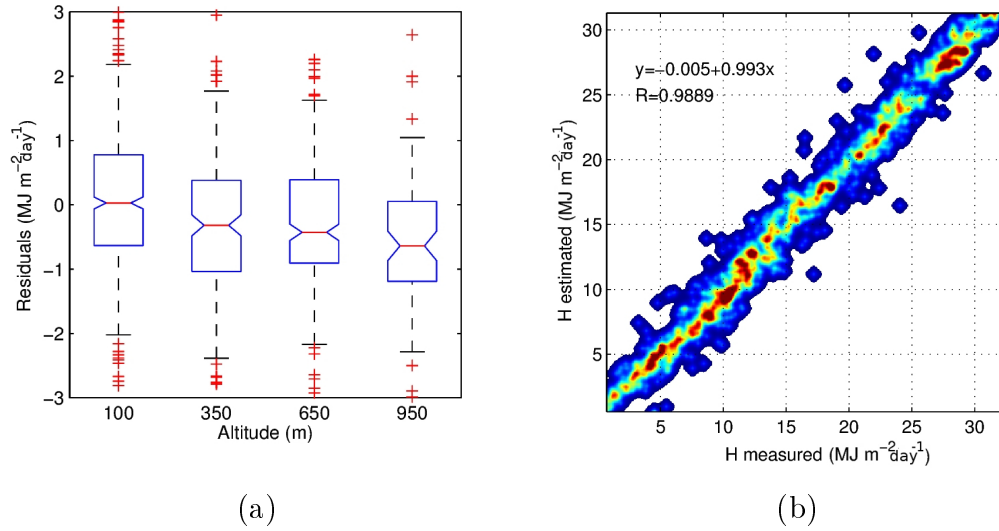


Figura 2.8. Diagrama de cajas con los sesgos entre los datos medidos *in situ* y los estimados mediante satélite (a), diagrama de dispersión entre todos los datos medidos *in situ* por la AEMet y las estimaciones de satélite (b).

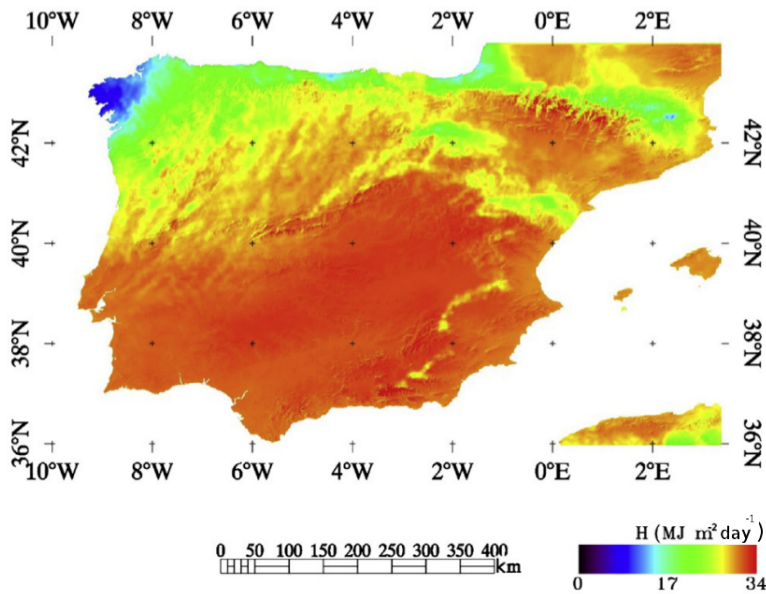


Figura 2.9. Imagen diaria de irradiación para el 1/7/2008 corregido de efectos topográficos y de línea de costa.

2.3. Imágenes de ε

Uno de los parámetros más difíciles de caracterizar y que más controversia causan en su estudio dentro de los PEM es la eficiencia en el uso de la radiación (ε). Esta presenta una gran variabilidad espacial y temporal, no puede ser medida directamente y se calcula de forma indirecta a partir de la GPP y la APAR, lo cual añade incertidumbres a su estimación. Aunque en la mayoría de modelos se asume que el factor más influyente en variabilidad de la ε es el tipo de cubierta [36], lo cierto es que se observa experimentalmente una variabilidad importante dentro de un mismo tipo. Estudios realizados sobre parámetros que controlan la ε indican que varía en función de la composición de especies, la ontogenia y las condiciones ambientales como la temperatura, disponibilidad de nutrientes, humedad del suelo, CO₂ atmosférico y el vapor de agua [75]. Estos parámetros serían los principales factores de variabilidad dentro de tipos de vegetación similares. La teoría de eficiencia funcional propuesta por Field [3] afirma que la vegetación tiende a no exceder la capacidad fotosintética, favoreciendo la adquisición del recurso más escaso, para ello limita la producción de más materia verde (e.g. hojas y cloroplastos). Así pues, el uso de cualquier recurso (nutrientes, agua o la PAR) está directamente ligado a la disponibilidad del resto, de la misma forma que cada característica de la planta está asociada a las otras de acuerdo con su tipo funcional [5].

En los modelos de teledetección, los forzamientos del entorno sobre el ecosistema se introducen en forma de factores adimensionales que impiden la óptima realización de la fotosíntesis. Estos factores reducen la eficiencia máxima (para cada tipo de cubierta) y se estiman a partir de los propios datos de satélite o a partir de datos meteorológicos espacializados. Esta aproximación permite, por tanto, estudios globales o regionales ya que captura los cambios relacionados con diferentes usos del suelo y con el clima [76]. Así, a escalas temporales cortas (de horas a días), los factores ambientales (agua y temperatura) parecen ser los más influyentes, mientras que en los cultivos anuales se observa que están correlacionados negativamente con el déficit de presión de vapor del aire (*Vapour Pressure Deficit*, VPD) y positivamente con su temperatura. Diferentes trabajos han empleado datos

medidos mediante la técnica de *eddy covariance* para buscar y cuantificar posibles relaciones entre carbono asimilado y factores de forzamiento ambiental. En este sentido *Garbulsky et al.* [4] apuntan que la distribución espacial de la ε_{\max} está más determinada por la precipitación típica anual que por el tipo de cubierta vegetal. Además sugieren que el cociente entre la evapotranspiración actual y la potencial es un buen indicador del balance de energía, presentando un alto potencial explicativo respecto a la variación estacional de la ε . Este factor corrector ya había sido considerado por otros PEM como por ejemplo el CASA (tabla 1.1). Existe cierta controversia a este respecto en la bibliografía, ya que otros estudios indican que la precipitación, evapotranspiración actual, VPD o la temperatura (esta última solo parece tener un peso importante en bosques donde se alcanzan bajas temperaturas) pueden ser empleados para cuantificar los forzamientos ambientales.

En esta tesis se ha empleado una aproximación para la estimación de la ε considerando un mapa de tipos de cubierta especialmente adaptado (mapa híbrido de diferentes cartografías) al área de estudio para la asignación de las eficiencias máximas (ε_{\max}). Como principales factores inhibidores se han considerado la temperatura mínima diaria y el estrés hídrico, siguiendo un formalismo como el mostrado en la ecuación 1.4. La inclusión de un factor de estrés hídrico viene motivada porque, de acuerdo con la bibliografía, es uno de los factores más influyentes en la mayoría de ecosistemas mediterráneos [10],[4],[77],[78]. La ε se cuantifica tanto utilizando datos de satélite como meteorológicos.

2.3.1. Imagen de ε_{\max} a partir de un mapa híbrido de tipos de cubierta

El desarrollo de productos de cobertura del suelo ha sido un activo campo de investigación en la última década, donde los avances en las características tecnológicas de los sensores junto con el incremento del potencial computacional han aumentado tanto la disponibilidad de datos como la capacidad de procesarlos, permitiendo producir mapas de cubiertas de mayor precisión,

con mayor rapidez y menor coste económico [79]. Dichos productos incluyen, entre los más utilizados, el International-Geosphere-Biosphere Programme Data and Information Systems (IGBP-DISCover) [80], University of Maryland Land Cover (UMD) [81], el Global Land Cover 2000 (GLC2000) [82], MODIS Land Cover [83], ECOCLIMAP-II [84], CORINE Land Use Land Cover Database [85] y GlobCover [86].

Sin embargo, existen importantes discrepancias entre estos productos de coberturas, puesto que han sido generados bajo la supervisión de diferentes iniciativas nacionales e internacionales, así como con diferentes objetivos y metodologías. Por tanto, existe la incógnita de determinar qué producto deber ser utilizado y cuál representa mejor la realidad de un determinado territorio. A este respecto, la combinación de las mejores propiedades de cada uno de los productos en un mapa híbrido o sinérgico mejora notablemente la disponibilidad de este tipo de información.

En este trabajo se ha empleado el mapa de cubiertas de *Perez-Hoyos et al.* [14]. Este producto consiste en el uso combinado de cuatro de los productos más utilizados en la actualidad (CGL2000, CORINE, MODIS y GlobCover) para obtener un mapa híbrido que representa con una mayor fiabilidad las cubiertas del suelo de la Península Ibérica. El método aplicado para derivar este mapa se basa esencialmente en un esquema de votación donde se introducen dos propiedades importantes: (1) el uso de la precisión de cada clase en cada uno de los productos considerados, lo que permite reducir las fuentes de error de cada producto individual, y (2) el uso del Land Cover Classification System (LCCS) como sistema de traducción entre leyendas, lo que permite reconciliar la ambigüedad entre las clases. La principal ventaja del mapa híbrido generado es que dispone de una estimación del error a nivel de cada píxel, es decir, un mapa de calidad. A partir de los datos del mapa de cubiertas se ha derivado la imagen de eficiencia máxima (ε_{\max}), tomando para ello como referencia los valores propuestos por *Garbulsky et al* [4]. En dicho trabajo los autores realizaron un análisis de valores y varianza de la ε_{\max} en diferentes tipos de cobertura vegetal a escala global.

Así pues, en el modelo optimizado se han calibrado los valores de la ε_{\max} en las clases vegetales en las que se disponía de datos de GPP *in*

situ, empleando los valores de la mediana obtenidos en dicho trabajo en los tipos de cubierta sin datos de validación. La decisión de utilizar este mapa híbrido, que mejora notablemente la caracterización de las cubiertas en la Península Ibérica, supone un valor añadido frente a otros productos operacionales de cálculo de la GPP ya que, por ejemplo, el producto oficial de MODIS (MOD17) emplea la clasificación IGBP cuya precisión en esta área de estudio es de tan solo el 36.1 % [87]. En la figura 2.10 se muestra el mapa híbrido empleado en esta tesis.

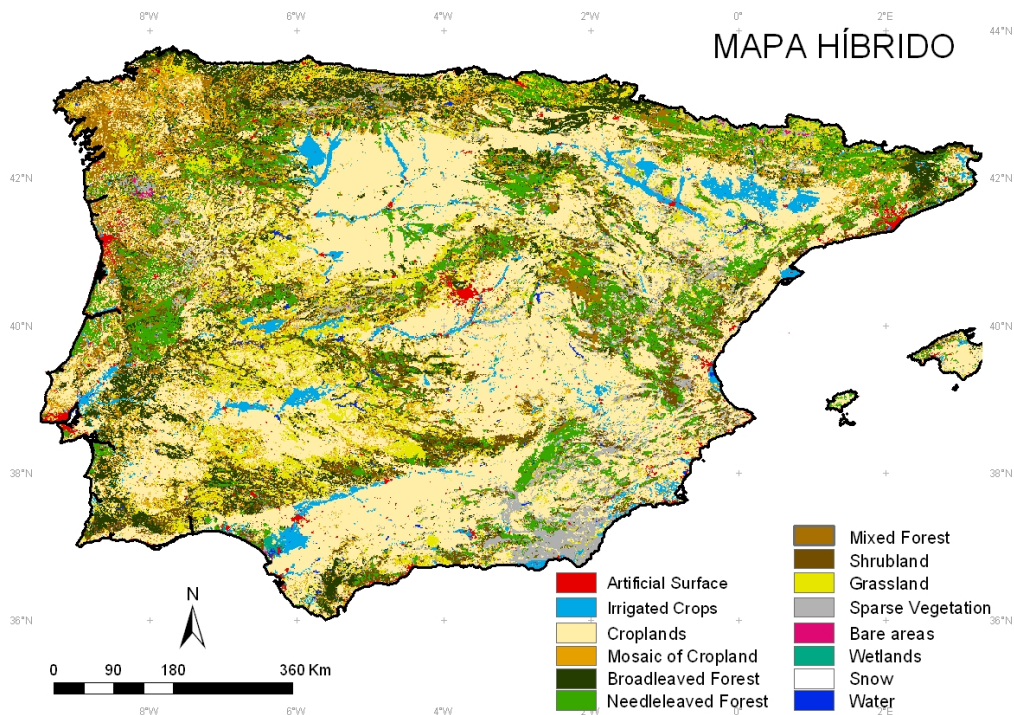


Figura 2.10. Mapa híbrido combinado de las principales cartografías de tipos de cubierta. Fuente: A. Pérez-Hoyos [87].

2.3.2. Estimación de la ε actual a partir de datos meteorológicos

A partir de datos exclusivamente meteorológicos, se han calculado dos factores: una corrección por bajas temperaturas basada en la metodología empleada por el producto oficial MODIS [6] y otra por estrés hídrico de forma similar a la desarrollada por *Maselli et al.* [10]. Al igual que ocurría con los datos meteorológicos empleados para la obtención de imágenes de irradiación solar, la obtención de imágenes de estos factores para utilizarlos como una capa adicional dentro del PEM requiere de un espacializado. Dicho espacializado, se ha llevado a cabo también mediante la técnica de krigado.

Factor corrector de la ε_{\max} por bajas temperaturas (T_{\min}). La tasa fotosintética suele ser típicamente más alta para temperaturas de hoja superiores (días soleados). Esta temperatura suele diferir de forma sustancial de la del aire debido a los efectos de enfriamiento que producen la transpiración, las propiedades de la hoja en la absorción de la luz y la influencia de las superficies cercanas. Los efectos de las bajas temperaturas en las plantas incluyen cambios en la bioquímica y biofísica de las membranas, en la síntesis proteica, modificaciones en enzimas de mitocondrias y cloroplastos y en los metabolismos fotosintético y respiratorio [20]. Las bajas temperaturas pueden producir separación de fases y, en caso de episodios prolongados en el tiempo, impiden mantener los gradientes iónicos y el metabolismo comienza a sufrir alteraciones llegando a producir la muerte de la célula en casos más extremos. Para considerar este tipo de efectos en el modelo optimizado se introduce un factor para la corrección de la disminución de eficiencia por efecto de bajas temperaturas diarias (T_{\min}). Su valor se obtiene a partir del ajuste lineal entre los puntos $(TMIN_{min}, 0)$ y $(TMIN_{max}, 1)$ correspondientes con nula y máxima eficiencia en el uso de la radiación respectivamente a unas temperaturas mínimas diarias dadas (figura 2.11). Para valores intermedios de $TMIN$ el valor de la reducción en la eficiencia se obtiene del ajuste, mientras que para valores menores o mayores se saturan a los valores extremos. Esta recta es diferente para cada tipo de cubierta (de acuerdo con los valores

suministrados por [6]), ya que se desea tener en cuenta las diferentes características de resistencia de las especies vegetales consideradas. Por lo tanto, las imágenes de T_{\min} se construyen utilizando también el mapa híbrido de vegetación mencionado anteriormente.

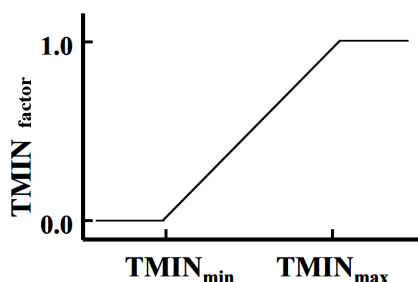


Figura 2.11. Ejemplo de la recta ajustada para la obtención del valor del factor de estrés por bajas temperaturas. Fuente: *Heinsch et al.* [6].

Factor corrector de la ε_{\max} por estrés hídrico (C_{ws}). La respuesta fotosintética de la vegetación varía a diferentes escalas temporales ante el déficit hídrico [10]. Hay una primera fase, a corto plazo, que afecta principalmente a las capas superficiales del suelo y la fotosíntesis. Este efecto se mitiga mediante mecanismos rápidos como el cierre estomático, que impide la pérdida rápida de agua y el subsiguiente daño en el aparato fotosintético. Estos cambios afectan principalmente a la ε dejando prácticamente inalterados los pigmentos y las características estructurales de la planta. Por otro lado, existe una segunda fase en presencia de periodos prolongados de sequía que afecta a las capas del suelo más profundas y altera la estructura interna del aparato fotosintético (estructura de las hojas y pigmentos).

El factor al que se hace referencia en este apartado captura los efectos más rápidos de la sequía y que afectan principalmente a la ε , mientras que los daños estructurales que alteran los pigmentos de la cubierta vegetal son capturados por la f_{APAR} (ya que se estima a partir del RDVI, sensible a estos cambios).

La expresión empleada para su cálculo es la siguiente:

$$C_{ws} = 0.5 + 0.5 \frac{\text{AET}}{\text{PET}} \quad (2.7)$$

donde AET y PET son la evapotranspiración actual y potencial respectivamente. La AET se identifica con la precipitación y la PET se calcula empleando la expresión de *Jensen & Haise* [88]:

$$\text{PET} = \left[\frac{R_s}{2450} ((0.025 \bar{T}) + 0.08) \right] \quad (2.8)$$

donde R_s (kJ m^{-2}) es la irradiación global diaria y \bar{T} ($^{\circ}\text{C}$) es la temperatura media diaria.

Este factor se calcula como el valor integrado para un periodo de 60 días. El periodo ha sido escogido para maximizar la correlación temporal entre las disminuciones de la ε debido al estrés hídrico y el factor C_{ws} de acuerdo con los resultados obtenidos por *Maselli et al.* [10]. Posteriormente, los valores son acotados de forma que C_{ws} varía entre $[0.5, 1]$, siendo 0.5 el máximo estrés hídrico y 1 la ausencia del mismo.

2.3.3. Estimación de la ε actual a partir de datos de satélite

La estimación de la ε mediante datos de satélite presenta un gran interés en la comunidad científica internacional, ya que su cálculo a escala global a partir de variables meteorológicas puede ser problemático. Entre las razones principales se puede destacar la no total disponibilidad de datos meteorológicos a esta escala y la dificultad de su obtención en tiempo real, ya que puede depender de los diferentes organismos locales que los gestionen. Además, estos datos pueden presentar una escasa densidad de muestreo espacial requiriendo de técnicas de interpolado espacial costosas computacionalmente que generan nuevas incertidumbres.

En esta tesis se ha explorado el uso de índices espectrales empleando datos de satélite para la estimación de la ε (apéndices C y D). Se han ensayado y analizado los índices espectrales PRI, NDWI, SIWSI, NDII7 que a continuación detallamos.

Estimación de la ε mediante el índice PRI. La búsqueda de indicadores precisos de estrés hídrico en tiempo real ha conducido el estudio de diversos parámetros fisiológicos o bioquímicos. En el segundo grupo han recibido gran atención los pigmentos fotoprotectores y compuestos antioxidantes que se inducen específicamente como respuesta a condiciones de estrés. El llamado “ciclo de las xantofilas” es un proceso de disipación térmica relacionado con la interconversión de tres xantofilas (carotenoides oxigenados): violaxantina, anteraxantina y zeaxantina. Este proceso ocasiona un cambio en la respuesta óptica de la cubierta vegetal provocando que, cuando es iluminada, uno de los procesos asociados sea el llamado *green-shift* o cambio de absorptividad en el verde ($\lambda \cong 500$ nm). Este cambio es debido fundamentalmente a dos componentes. Primero un cambio de absorptividad centrado en 505 nm debido a la aparición de zeaxantina y, segundo, a una disminución de la absorptividad centrado en 535 nm debido a la protonación del complejo recolector de luz en presencia de zeaxantina [89]. Estos cambios de absorptividad pueden ser utilizados para analizar de forma no des-

tructiva el estado del ciclo de las xantofilas a través de las respuestas (transmisión o reflexión) espectrales de la hoja. Como consecuencia de estos cambios en la respuesta espectral *Gamon et al.* [90] desarrollaron un índice espectral (*Photochemical Reflectance Index*, PRI) con el fin de relacionarlo con la ε :

$$\text{PRI} = \frac{R_{531} - R_{\text{ref}}}{R_{531} + R_{\text{ref}}} \quad (2.9)$$

donde R_{531} es la reflectividad en una banda espectral estrecha centrada en 531 nm y R_{ref} en una banda de referencia insensible a los procesos involucrados en el “ciclo de las xantofilas”. En presencia de estrés hídrico, la mayor absorción en 531 nm origina una disminución casi instantánea del PRI.

Aunque diferentes estudios demuestran que a nivel de hoja existen relaciones significativas entre el índice y la ε [91], [92], los resultados no son tan concluyentes cuando empleamos imágenes de satélite [93], [94], [2]. En esta tesis se adjunta una publicación sobre el análisis del potencial del PRI calculado empleando datos del satélite MODIS para monitorizar el estrés hídrico estival en un bosque Mediterráneo en la región de la Toscana en Italia (apéndice C). Las principales conclusiones de dicho análisis fueron las siguientes:

- Existen correlaciones significativas entre el índice PRI y la ε , siendo el coeficiente de determinación entre ambas de 0.69. Este resultado es coherente con los previamente publicados en la bibliografía.
- Se emplearon métodos estadísticos avanzados (*elastic net*) para evaluar cuantitativamente las posibles dependencias entre el PRI y características geométricas de la adquisición de las imágenes del satélite. El análisis indica que la mayoría (70 %) de la variabilidad del PRI se explica por las condiciones de observación e iluminación (efectos de primer orden), otra parte por el contenido en clorofila y la estructura de la cubierta (efectos de segundo or-

den), mientras que los efectos de las xantofilas (efectos de tercer orden) pueden ser enmascarados por el ruido del sensor a escala de píxel MODIS (1 km).

- Se propuso un índice normalizado de los factores geométricos (NPRI), cuya serie temporal manifestaba un retraso respecto a la serie temporal de ε . Esto significa que, al eliminar la influencia de los factores geométricos, también desaparece la dependencia con el ciclo de las xantofilas (cuya respuesta es instantánea al estrés pero muy poco significativa a escala de MODIS), quedando sólo la dependencia del PRI a cambios de los pigmentos y estructurales, que no se manifiestan a corto plazo.

La figura 2.12 muestra cómo gran parte de la varianza temporal de la señal del PRI puede ser explicada con un modelo lineal sencillo partir de la información geométrica de la adquisición de la escena, más concretamente, se han empleado como variables de entrada los ángulos zenitales de observación e iluminación. Cuando se construye un índice normalizado (NPRI) a partir de los residuos resultantes de eliminar del PRI la información relacionada con la geometría de observación (modelo lineal), se observa (figura 2.13) cómo el nuevo índice presenta un comportamiento similar a la ε con un cierto retardo.

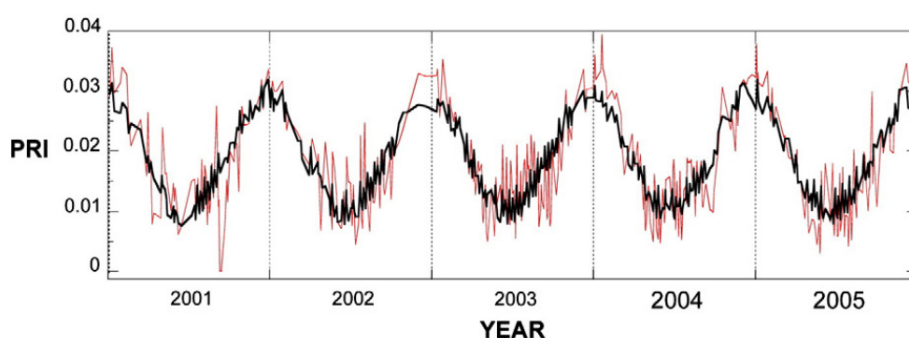


Figura 2.12. PRI calculado mediante un modelo exclusivamente dependiente de la información geométrica de adquisición (negro) y PRI estimado a partir de datos de reflectividad MODIS (rojo).

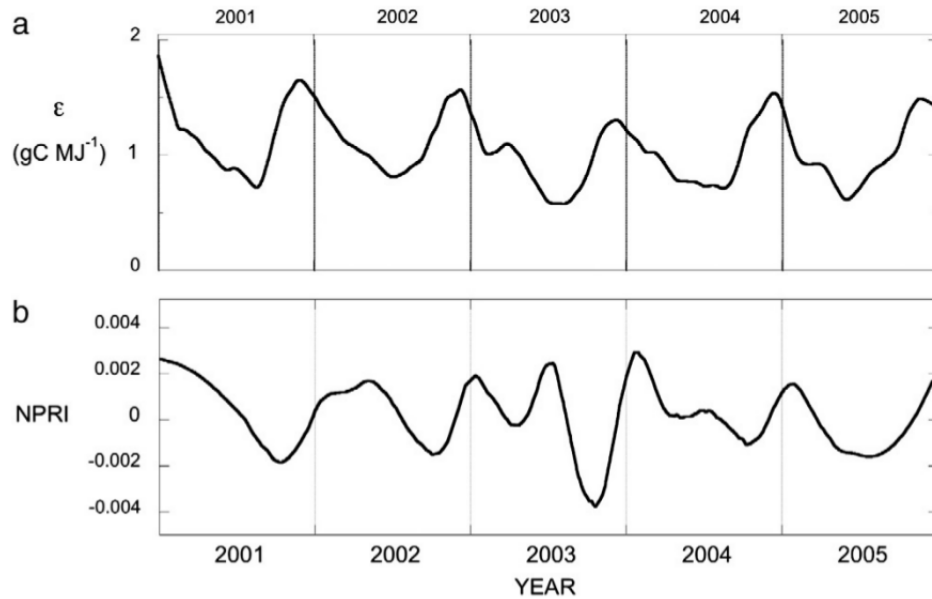


Figura 2.13. ε y NPRI calculado con datos MODIS.

A pesar de la gran ventaja que puede suponer caracterizar la ε utilizando datos de teledetección, el PRI (o, en su caso, el NPRI) no es adecuado para trabajar a escala diaria y cuando se pretende detectar el estrés hídrico en tiempo casi real.

Estimación de la ε mediante los índices NDWI, SIWSI, NDII7. La signatura espectral de la vegetación puede proporcionar información acerca del contenido en agua. Experimentos de laboratorio, simulaciones y diferentes trabajos empleando imágenes de satélite indican la viabilidad de estimar el espesor equivalente en agua (*Equivalent Water Thickness*, EWT) a partir de la reflectividad de las plantas [95],[96],[97]. El descenso del contenido en agua se manifiesta de forma relativamente temprana en la signatura espectral de la vegetación en bandas espectrales centradas en 1.45, 1.94 y 2.50 μm , mientras que los cambios en el contenido en clorofila, cuyas variaciones están más relacionados con la destrucción del aparato fotosintético debido a periodos más prolongados de estrés, se manifiestan modificando la respuesta espectral de reflectividad en la región del visible (0.4–0.7 μm). Esta característica ha impulsado el desarrollo de índices que incluyan alguna de las bandas

espectrales del infrarrojo de onda corta (*Short Wave Infrared, SWIR*) y otra de referencia del infrarrojo cercano (*Near Infrared, NIR*). En esta línea se han desarrollado tres índices susceptibles de ser empleados con datos del satélite MODIS, el NDWI (*Normalized Difference Water Index*) [98], el SIWSI (*Shortwave Infrared Water Stress Index*) [99] y el NDII7 (*Normalized Difference Infrared Index 7*) [100].

En esta tesis se ha evaluado el potencial de estos tres índices para monitorizar el estrés hídrico y sus impactos en la ε sobre siete tipos de cubierta vegetal típicos de la Toscana italiana empleando datos del satélite MODIS en 70 áreas de estudio (apéndice D). En particular, se ha evaluado el potencial de estos índices para capturar la variabilidad del factor de balance de agua C_{ws} , que permite capturar el descenso en la ε por efecto del estrés hídrico. Las conclusiones más destacadas de este trabajo fueron las siguientes:

- El estudio estadístico llevado a cabo indica que los tres índices espectrales son moderadamente informativos (bajas correlaciones) de las variaciones interanuales del factor C_{ws} . El índice SIWSI es el que mejores resultados proporciona en la mayoría de clases de cubierta consideradas.
- El reemplazo directo del factor C_{ws} por un índice espectral no es posible, imposibilitando el uso inmediato del índice en el PEM como un indicador de la ε . La razón es que el índice espectral no solo es informativo del posible estrés hídrico que sufre la vegetación, sino que también reflejan en su evolución temporal otros aspectos como las variaciones estacionales y fenológicas típicas de la cubierta.
- En el caso de que no existan datos meteorológicos recientes en la zona de estudio, se ha desarrollado un método operacional que combina la información espectral relacionada con el estrés hídrico e información meteorológica. Utiliza series climáticas de C_{ws} y las actualiza a las condiciones actuales de la vegetación empleando información de los índices de contenido en agua.

- La validación sobre estimaciones de la GPP del factor original C_{ws} y su versión modificada utilizando índices espectrales C_{ws}^* sobre tres áreas de estudio revela resultados similares. Esto indica que la utilización de información espectral de satélite puede ser una solución viable para monitorizar el efecto del estrés hídrico sobre la ε en ausencia de datos meteorológicos con el mismo grado de fiabilidad, pero siempre que se disponga de series “históricas” de C_{ws} .

En la figura 2.14 se muestra una comparativa de perfiles del factor de estrés hídrico empleando de forma combinada la información espectral de satélite y datos meteorológicos (C_{ws}^*) y el factor original (C_{ws}) que únicamente requiere de datos meteorológicos para su cálculo. Dada la capacidad de ambas aproximaciones para capturar convenientemente la información interanual del estrés en la cubierta pueden ser intercambiados de acuerdo con nuestros resultados (apéndice D). La figura 2.15 muestra que para el cálculo de la GPP en el PEM, el impacto del uso de una u otra metodología es poco relevante en términos de la calidad de las estimaciones. En esta tesis, no obstante, se utiliza el C_{ws} puesto que contamos con las series de datos meteorológicas necesarios para su cálculo.

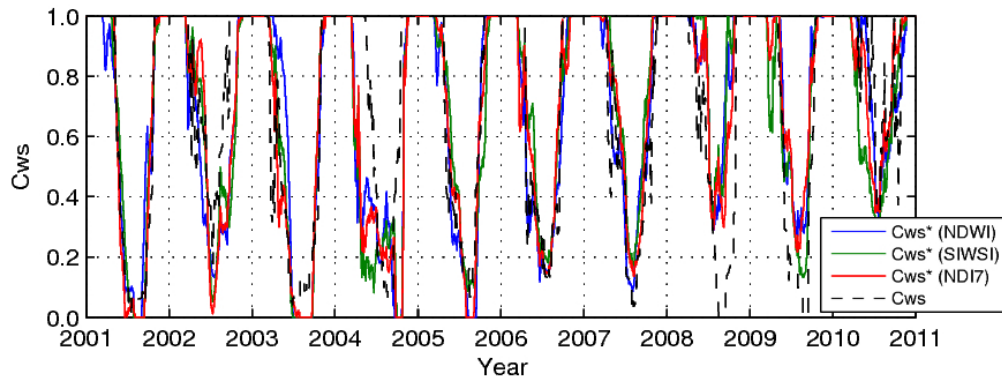


Figura 2.14. Serie temporal original de C_{ws} y tres modificaciones empleando tres índices espectrales a partir de datos MODIS para la torre de San Rossore en Italia.

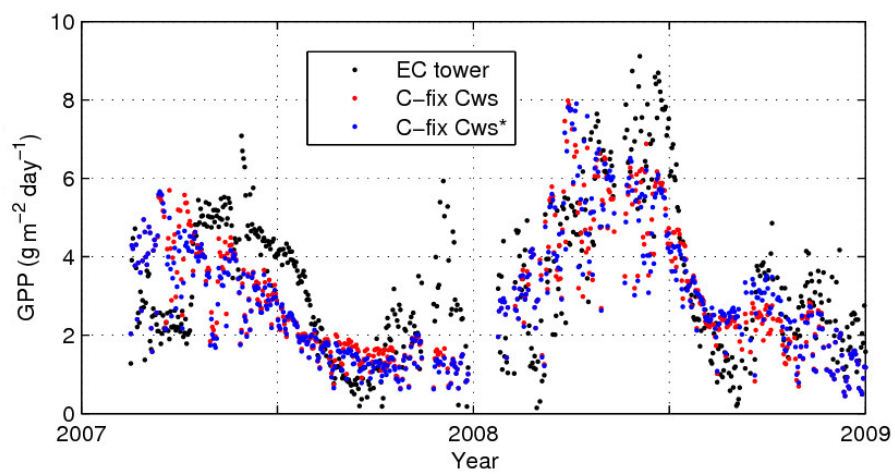


Figura 2.15. Cálculo de la GPP a partir de datos *eddy covariance* y a partir del modelo optimizado con el factor original C_{ws} y con el C_{ws}^* modificado empleando el índice espectral SIWSI.

Capítulo 3

Cálculo y validación del producto de GPP

En este capítulo se muestran las imágenes de GPP obtenidas a partir del modelo optimizado, se realiza la validación de las mismas y, por último, se discuten los factores que justifican su variabilidad. La validación de las imágenes es de gran importancia para poder cuantificar la calidad de nuestras estimaciones. Se ha realizado de forma directa utilizando datos de GPP estimados a partir de datos de flujos medidos en torres *eddy covariance* y de forma indirecta a través de comparación con otros productos operacionales disponibles. Más concretamente, se han empleado los productos que forman parte de la cadena de producción de MODIS (MOD17A2) y Copernicus (DMP).

3.1. Cálculo de la GPP

Como se ha comentado en el capítulo 2, en esta tesis se aborda el cálculo de la GPP empleando una aproximación de tipo Monteith. Para ello, se han investigado los diferentes *inputs* del modelo y se han obtenido con una baja incertidumbre como se ha mostrado en el capítulo anterior.

Así, en los resultados que se muestran y validan en el siguiente apartado, la GPP ha sido estimada a partir de una f_{APAR} que sigue la aproximación de *Roujean & Bréon* [12]. La PAR se ha calculado a partir de dos métodos: 1) empleando datos de satélite del producto operacional DSSF suministrado por el LSA-SAF que ha sido integrado, remuestreado y validado y, 2) construyendo modelos mediante métodos lineales avanzados con datos puntuales de estación espacializados (permite que, en caso de ausencia de datos de satélite, se disponga de ellos). En cuanto a la eficiencia en el uso de la radiación (ε), se ha empleado una cartografía adaptada al área de estudio (mapa híbrido) de tipos de cubierta vegetal para asignar los valores máximos de eficiencia (ε_{max}) y dos factores inhibidores de la misma. Un factor que tiene en cuenta los efectos de las bajas temperaturas (T_{min}) y otro que considera los efectos del estrés hídrico típico de los ecosistemas mediterráneos (C_{ws}). En la figura 3.1 se muestra esquemáticamente cómo se integran los diferentes parámetros empleados para el cálculo de la GPP diaria en el modelo optimizado.

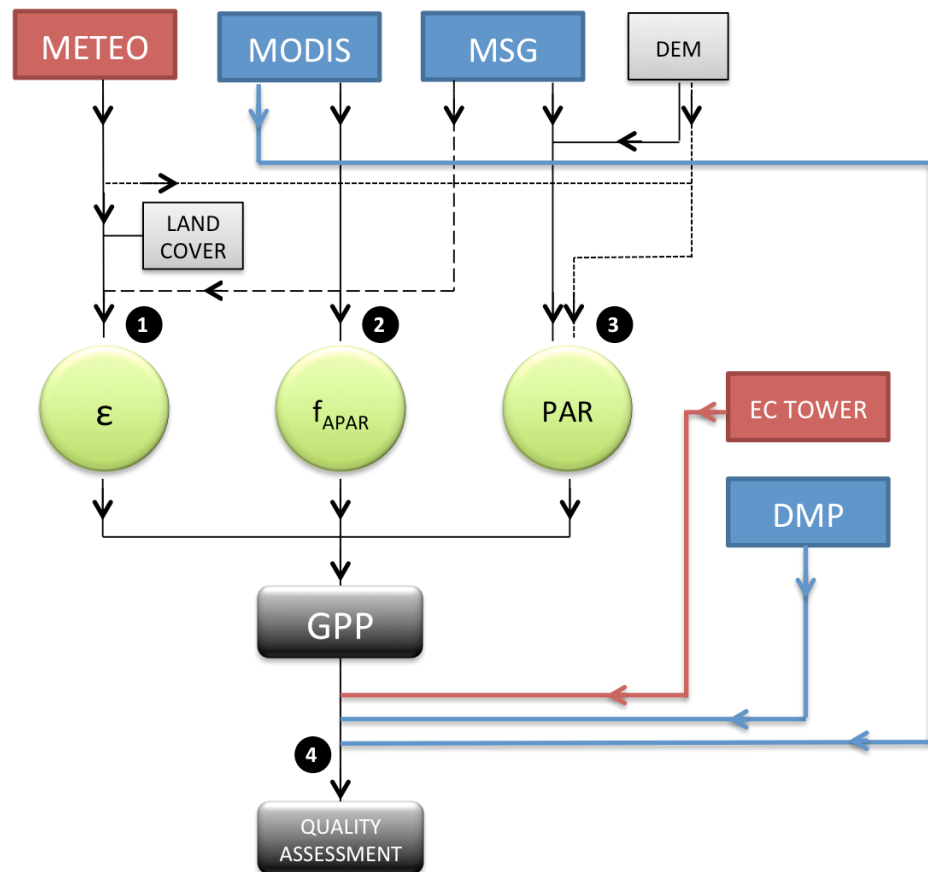
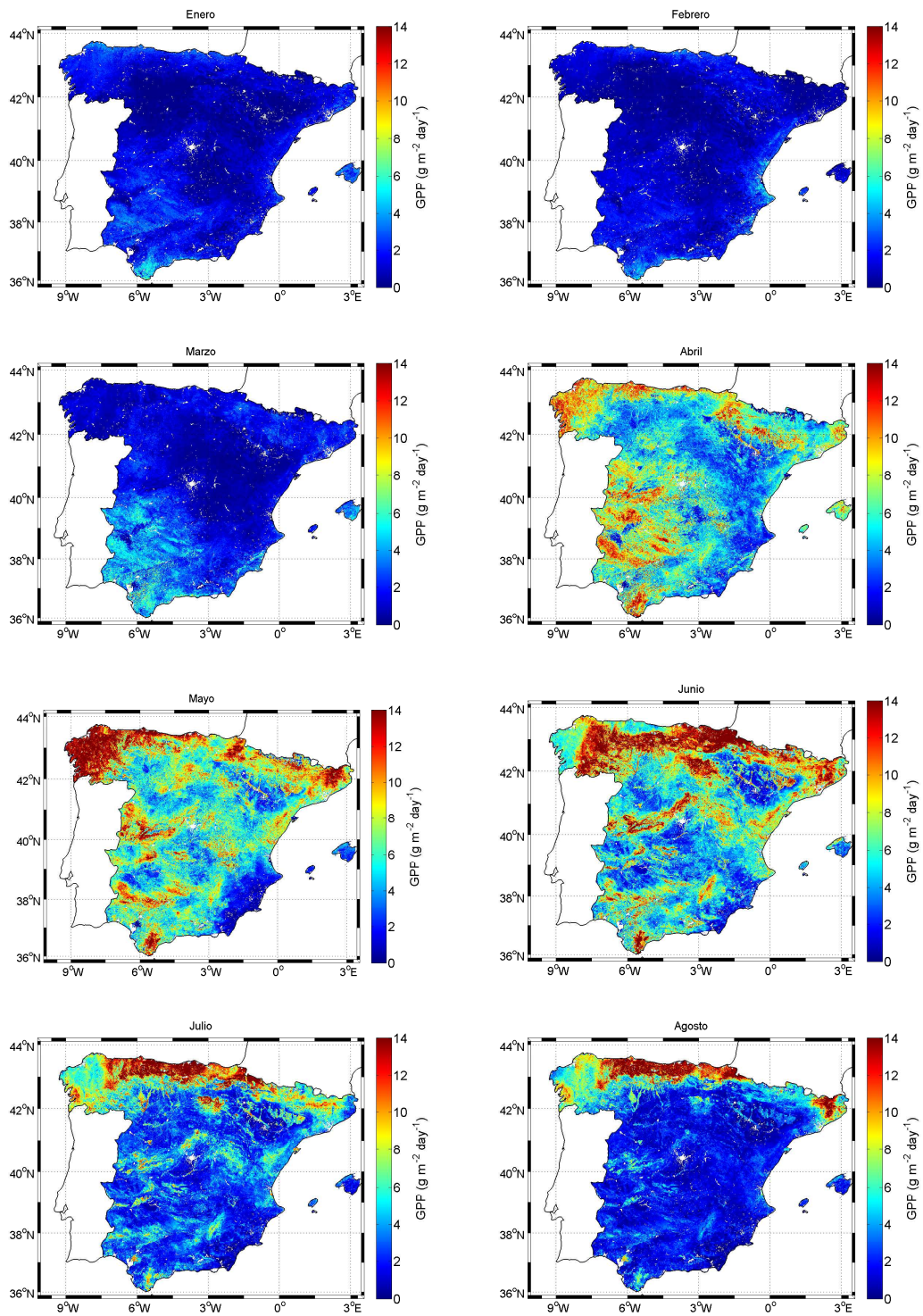


Figura 3.1. Diagrama de flujo del modelo optimizado empleado para el cálculo de la GPP. Los pasos 1, 2, 3 se han descrito en el capítulo 2. El paso 4 se desarrolla en este capítulo.

Siguiendo el esquema mostrado, en esta tesis se ha calculado la GPP diaria a una resolución de 1 km para el área de estudio y los años 2008 y 2011. En la figura 3.2 se muestra, a modo de ejemplo, la GPP de un día de cada mes del año 2011.



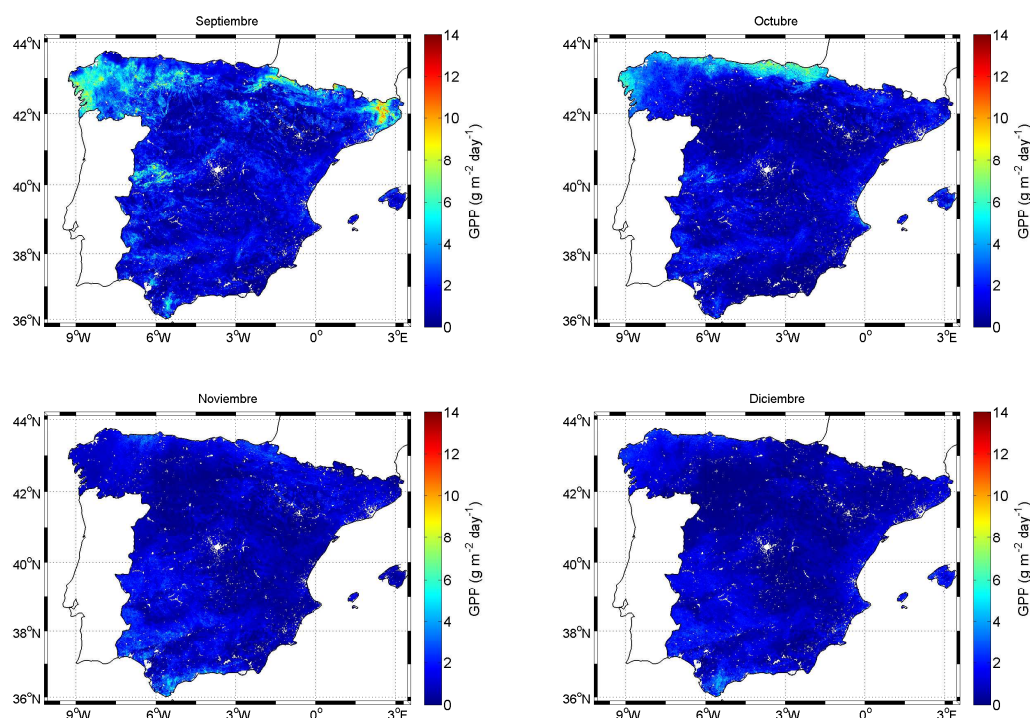


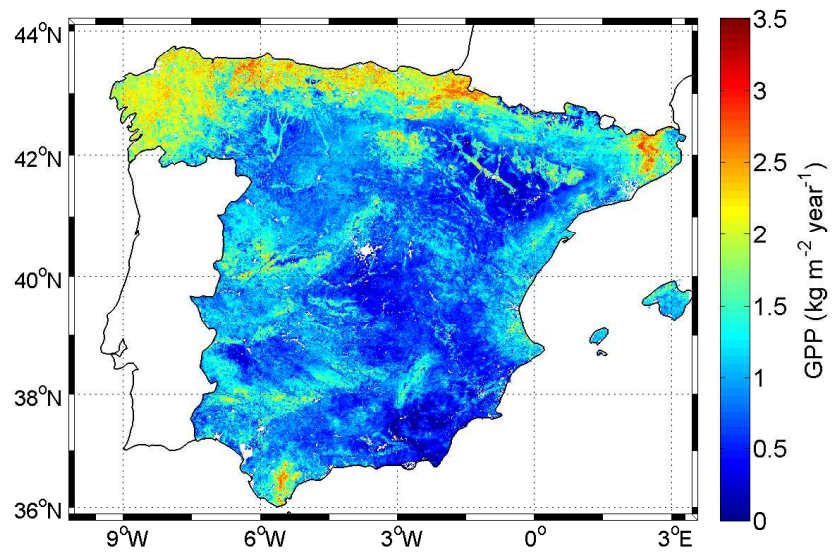
Figura 3.2. GPP diaria calculada empleando el modelo optimizado a una resolución espacial de 1 km. Se muestra un día de cada mes para el año 2011.

Como se observa en la figura 3.2, existe una gran variabilidad de la GPP, sobre todo en los meses de primavera y verano, muy ligada con el desarrollo fenológico de cada tipo de vegetación (a través de la f_{APAR}). Además, la variabilidad temporal viene influenciada por la PAR, que a medida que transcurre el año se incrementa hasta alcanzar su máximo anual en junio para volver a disminuir gradualmente. Los cambios bruscos en PAR entre fechas consecutivas son consecuencia de la presencia de nubes en la escena. Así, las imágenes diarias de GPP indican que esta es mínima hacia el final del otoño y el periodo invernal (de noviembre a marzo), cuando la radiación es menor y el efecto de las bajas temperaturas es más patente coincidiendo con un menor desarrollo de la vegetación. Por otro lado, la GPP alcanza su máximo en el periodo de primavera y a principios de verano, cuando la PAR diaria es máxima y la vegetación ha alcanzado su máximo desarrollo dentro de su ciclo fenológico anual. En los meses de julio y agosto se pro-

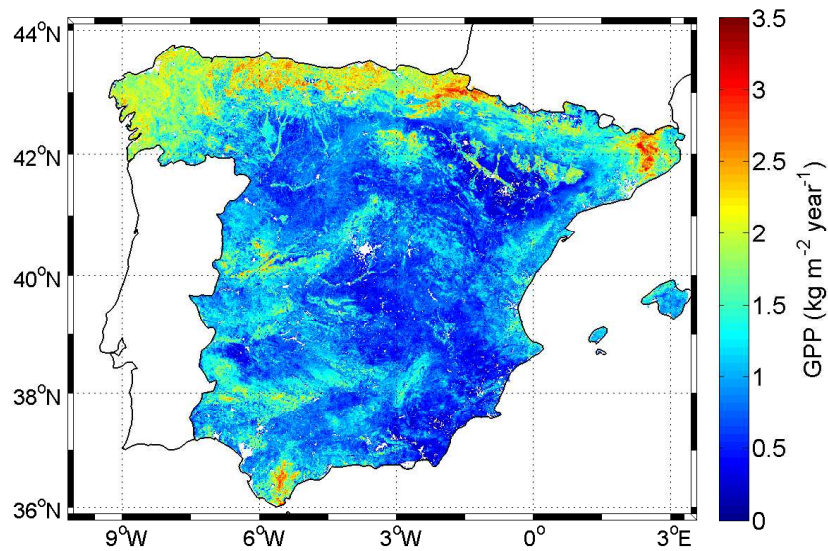
duce un descenso significativo de la GPP, a excepción de la zona norte, la más húmeda del área de estudio. Esto ocurre porque el resto del territorio está sometido a estrés hídrico que inhibe la actividad fotosintética de la vegetación. Este efecto se caracteriza adecuadamente a través del factor C_{ws} incluido en el modelo. En la época de otoño se produce un mayor declive de la actividad fotosintética debido a un descenso de la PAR. El vigor de la vegetación decrece mucho en esta época, llegando incluso a perderse la mayoría de las hojas en las especies caducifolias.

En la figura 3.3 se muestra la GPP anual para el año 2008 y 2011. La comparación de estas estimaciones con valores típicos del índice NDVI y precipitación [101] indica que la distribución espacial de los valores más altos de GPP están relacionados con ambas variables de manera directa. Esta característica tiene que ver con la gran influencia que ejerce el clima sobre las especies vegetales. Así, por ejemplo, la disponibilidad de recursos esenciales para su desarrollo, como es el agua, condicionan fuertemente la distribución espacial de las mismas. Los patrones espaciales de la GPP para ambos años son similares. No obstante, para resaltar patrones espaciales diferentes entre ellos, se han calculado imágenes diferencia de ambas variables entre los dos años (2008-2011) (figura 3.4). Los valores negativos indican una mayor GPP/precipitación en 2011.

Al comparar ambas imágenes diferencia se observa cierta coherencia espacial entre ambos, especialmente en la zona de Galicia, Castilla León, Cataluña, Comunidad Valenciana, Islas Baleares y las zonas costeras de Málaga y Granada. La precipitación está ligada muy directamente con el desarrollo y productividad de la vegetación, especialmente en zonas mediterráneas donde la disponibilidad de este recurso no está siempre garantizada. En el caso contrario, no parece observarse una correspondencia tan clara entre las diferencias de GPP y precipitaciones en las provincias de Asturias, Cantabria, País Vasco y Navarra. Se ha de tener en cuenta que para la óptima realización de la fotosíntesis es tan importante la cantidad de precipitación como el momento en el que esta ocurre, información que no queda puesta de manifiesto en la figura 3.4 (b).

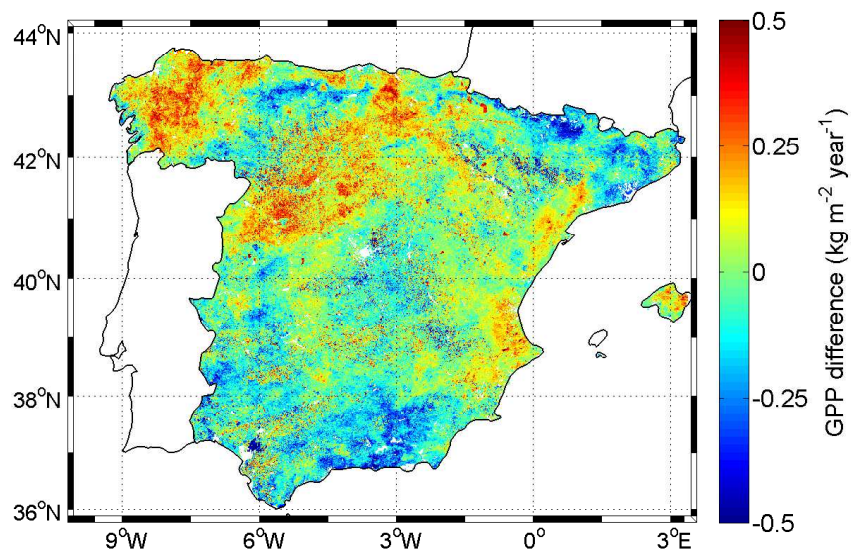


(a)

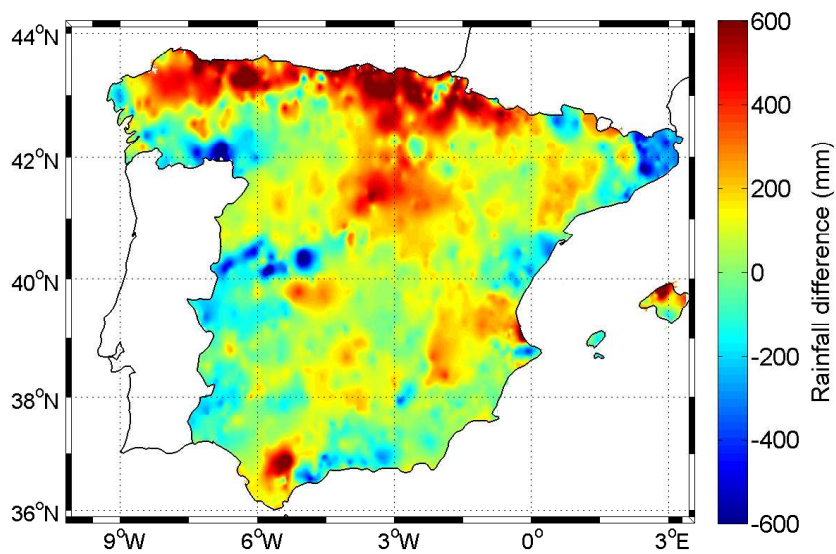


(b)

Figura 3.3. GPP anual calculada para los años 2008 (a) y 2011 (b).



(a)



(b)

Figura 3.4. Mapa de diferencias anuales de GPP (a) y de precipitación anual (b) entre los años 2008 y 2011.

3.2. Validación de resultados

La validación directa hace referencia a la determinación de la incertidumbre del modelo a partir de la comparación con datos de referencia [102]. Los datos de referencia corresponden normalmente a estimaciones *in situ* del parámetro de interés. Estos están distribuidos espacialmente sobre un área coincidente con la resolución del producto a validar o incluso mayor [103], de forma que caractericen la variabilidad espacial de la zona validada. Sin embargo, para productos de media y baja resolución como los desarrollados en esta tesis, este procedimiento es difícil de llevar a cabo debido a que los productos se extienden a escala regional y global. En este caso se requiere la adquisición de medidas *in situ* sobre diferentes zonas de estudio distribuidas globalmente y que capturen la variabilidad del modelo y de diferentes tipos de cubiertas. Tal como veremos a continuación, el número de datos procedentes de torres EC no es muy elevado. Por este motivo, se complementará la validación con una intercomparación del producto de GPP obtenido con otros productos existentes (MODIS y DMP) para la zona de estudio (validación indirecta).

La nomenclatura empleada para diferenciar las diversas estimaciones de GPP es: GPP_{OPT} para la GPP calculada con el modelo optimizado desarrollado en esta tesis, GPP_{EC} para la GPP estimada a partir de datos de torres *eddy covariance* medidos *in situ*, GPP_{MODIS} para la GPP del producto operacional de MODIS y GPP_{DMP} para la GPP del producto operacional de Copernicus.

3.2.1. Validación directa utilizando datos de torres EC

La validación a partir de medidas *in situ* consiste en la comparación directa entre los productos derivados mediante teledetección y dichas medidas de referencia. Hay que señalar que la validación directa de la GPP, en sentido estricto, no es posible por la imposibilidad de medirla directamente mediante un sensor. A pesar de este problema, su estimación a partir de medidas de NEP en torres EC suele ser la forma habitual de comparar la calidad de las

medidas estimadas con datos de satélite u otros métodos. Por simplicidad, en este trabajo nos referiremos a la estimaciones mediante datos de torres EC como medidas *in situ*. Se han utilizado siete torres disponibles para el área de estudio en los años 2008 y 2011, que pertenecen a la iniciativa europea *European Fluxes Database Cluster* (<http://www.europe-fluxdata.eu/>). El objetivo fundamental de la iniciativa *European Eddy Fluxes Database Cluster* (EEFDC) es mejorar la estandarización, integración y colaboración entre las bases de datos que forman parte de los diferentes proyectos de investigación europeos. Para ello, la EEFDC busca impulsar la creación de una infraestructura única que aglutine medidas de flujos entre los ecosistemas y la atmósfera pero empleando unas herramientas de intercambio y procesado de los datos comunes de alta calidad.

A continuación se da una descripción de las características principales de los sitios considerados con estimaciones de torres EC:

Aguamarga (AGU). Esta torre se encuentra ubicada en Almería dentro del Parque Natural de Cabo de Gata ($36^{\circ} 56' 26'' N$, $2^{\circ} 1' 48'' W$) a una altitud de 200 m sobre el nivel del mar. La vegetación cubre aproximadamente el 60 % de la superficie y está dominada por *Stipa tenacissima*, con una altura comprendida entre 60-100 cm. También coexisten especies herbáceas tales como *Chamaerops humilis*, *Rhamnus lycioides*, *Asparagus horridus*, *Olea europea var. sylvestris*, *Pistacia lentiscus* y *Rubia peregrina*. Es una de las zonas más áridas de España con una precipitación anual baja e irregular de unos 200 mm y temperaturas media anuales de 18 °C.

Cortes de Pallás (COR). La torre está ubicada en Valencia ($39^{\circ} 13' 26'' N$, $0^{\circ} 56' 11'' W$) a una altitud de 810 m, con temperaturas medias mensuales comprendidas entre 5 °C y 28 °C. Presenta un clima caracterizado por lluvias irregulares con una precipitación media anual de aproximadamente 600 mm. La vegetación predominante alrededor de la torre se compone de matorral bajo, con presencia de romero (*Rosmarinus officinalis*) y brezo (*Calluna vulgaris*).

Lanjarón (LAN). Está ubicada en la loma de Lanjarón, Sierra Nevada ($36^{\circ} 58' 20'' N$, $3^{\circ} 28' 26'' W$) a una altitud de 2300 m. Dicha zona fue afectada por un incendio en septiembre de 2005 dañando aproximadamente 1 500 ha de bosques de pinos (*Pinus sylvestris*) de reforestaciones realizadas hace 50 años. En el año 2008 el 89 % de los árboles se mantenían en pie aunque ninguno sobrevivió al incendio. La vegetación está compuesta mayoritariamente por herbáceas, siendo la *Genista versicolor*, *Festuca spp.* y *Sesamoides prostrata* las especies predominantes. El clima de la zona es mediterráneo subalpino con una precipitación anual media de 470 mm y una temperatura media anual de 8 °C.

Lagunaseca (LAG). Esta torre se ubica en la vertiente septentrional de Sierra Nevada ($37^{\circ} 5' 53'' N$, $2^{\circ} 57' 57'' W$) a unos 39 km de la costa mediterránea y a 2 300 m de altura sobre el nivel del mar. El 45 % de la superficie está cubierta por las especies vegetales *Festuca indigesta* y *Cytisus Purgans* que alcanzan una altura de 20 cm y cubren el 45 % de la superficie. El 35 % corresponde al lecho y el resto es suelo desnudo. El clima se caracteriza por una precipitación media anual de 800 mm y una de temperatura media anual de 6 °C.

Llano de los Juanes (JUA). La torre está ubicada en la sierra de Gádor ($36^{\circ} 55' 41'' N$, $2^{\circ} 45' 1'' W$), a una altitud de 1 600 m y con una temperatura media anual de 12 °C. La precipitación media anual es de 475 mm, con lluvias que se concentran principalmente durante el otoño e invierno. Es una meseta cubierta de matorral de alta montaña, formado por dos especies vegetales dominantes, la *Festuca scariosa* y la *Genista pumila*. Cubren en total una media del 46 % del suelo.

Las Majadas del Tiétar (MAJ). La torre de Majadas de Tiétar en Cáceres ($39^{\circ} 56' 29'' N$, $5^{\circ} 46' 2'' W$) se encuentra a una altitud de 258 m y presenta una temperatura máxima mensual de 35 °C y mínima de 1 °C. La precipitación anual es de 572 mm, con lluvias más bien escasas e irregulares, que se concentran en primavera y en otoño. La vegetación natural en la zona corresponde a la serie mesomediterránea *Pyro bourgeanae-Querceto rotundifoliae quercetum*, cuyo manejo ha resul-

tado en un bosque de encinas adhesionado con un sotobosque dominado por unas pocas especies (*Cistus ladanifer*, *Lavandula pedunculata* y *Ulex europaeus*).

Sueca (SUE) La torre de Sueca está ubicada en el parque natural de la Albufera. Es una zona extensa de cultivos de arroz (15 000 ha) ($39^{\circ} 56' 29'' N$, $5^{\circ} 46' 2'' W$). El clima es mediterráneo sub-árido con veranos cálidos y secos y con pocos excesos hídricos en invierno. Las temperaturas medias y la precipitación anual son $17^{\circ}C$ y 550 mm, respectivamente. El cultivo de arroz es muy homogéneo temporalmente (momento de cultivo e inundación) y en las técnica de cultivo (no ha cambiado significativamente en los últimos 200 años). Los campos se vacían de marzo a abril, se inundan en mayo para sembrarlos y se vacían de nuevo en agosto para recolectar el arroz a principios/mitad de septiembre.

Las principales características de cada una de las zonas de estudio se muestran resumidas en la tabla 3.1. Las torres se ubican mayoritariamente en zonas de vegetación escasa, a excepción del arrozal de Sueca y la dehesa de las Majadas del Tiétar.

Tabla 3.1. Datos más relevantes de las torres *eddy covariance*. OSH hace referencia a matorrales dispersos, CSH a matorrales densos, CRP a cultivos (arroz) y EBF a bosques perennes de hoja ancha.

Nombre	Latitud ($^{\circ}$)	Longitud ($^{\circ}$)	Cubierta
Aguamarga (AGU)	36.941	-2.033	OSH
Cortes de Pallás (COR)	39.224	-0.903	OSH
Lagunaseca (LAG)	37.098	-2.966	OSH
Llano de los Juanes (JUA)	36.927	-2.752	OSH
Lanjarón (LAN)	36.972	-3.474	OSH
Las Majadas del Tiétar (MAJ)	39.941	-5.773	CSH/EBF
Sueca (SUE)	39.276	-0.315	CRP

La tabla 3.2 muestra los estadísticos obtenidos (MBE, MAE, RMSE, rMBE, rMAE, rRMSE, R^2) para cada una de la torres EC de validación en el año 2008. La definición de dichos estadísticos se puede encontrar en el apéndice E. A nivel climático, de acuerdo con los calendarios meteorológicos que suministra AEMet [104], el año 2008 se caracterizó por ser cálido, algo seco (en la vertiente cantábrica) y más húmedo de lo normal en la costa mediterránea.

Tabla 3.2. Resultados de la validación directa para los píxeles donde se ubican cada una de las torres EC disponibles en el año 2008¹.

Torre EC	MBE	MAE	RMSE	rMBE	rMAE	rRMSE	R^2
Aguamarga	0.12	0.37	0.47	0.14	0.42	0.53	0.50
Lagunaseca	0.00010	0.38	0.52	0.00083	0.32	0.43	0.80
Lanjarón	-0.29	0.50	0.64	-0.23	0.40	0.51	0.60
Llano de los Juanes	0.30	0.39	0.48	0.47	0.62	0.76	0.68
Las Majadas del Tiétar	0.23	0.80	1.0	0.081	0.28	0.38	0.80
Sueca	0.34	1.0	1.5	0.086	0.26	0.38	0.96

Los resultados de la tabla 3.2 muestran, en general, un alto potencial explicativo de la GPP calculada frente a las estimaciones a partir de datos *in situ* en torres EC. La varianza explicada por el modelo, R^2 , oscila entre un 50 % para la torre de Aguamarga y un 96 % para el arrozal de Sueca. En general, los errores son elevados especialmente en términos absolutos y cuadráticos, llegando al 60 % de rMAE y al 77 % de rRMSE para la torre de Llano de los Juanes. Esto se relaciona con el hecho de que, a escalas temporales diarias, la GPP presenta variaciones muy abruptas relacionadas principalmente con la PAR y con problemas en la correcta determinación de la respiración² [29]. Se pueden destacar los excelentes resultados obtenidos en términos de sesgo (MBE) para Lagunaseca ($\simeq 0$) y en términos de MAE ($\simeq 0.8 \text{ g m}^{-2} \text{ day}^{-1}$) y RMSE ($\simeq 1.1 \text{ g m}^{-2} \text{ day}^{-1}$) para las Majadas del Tiétar. En la figura 3.5 se observa cómo la evolución anual estimada por el modelo optimizado reproduce adecuadamente la serie temporal estimada a partir de datos EC para todos los píxeles, de ahí los altos valores de R^2 . Cabe destacar las discrepancias en los perfiles temporales al final de verano

¹Se adopta el convenio de proporcionar los errores con dos cifras significativas.

²La GPP de las torres se estima a partir de la NEP y de la respiración.

y otoño en las torres de Lagunaseca y Lanjarón. La fuente de estas discrepancias es la f_{APAR} , que indica ausencia de vegetación cuando a nivel de suelo se ha observado que la existencia de pequeños episodios de precipitaciones hace crecer de nuevo la vegetación herbácea aunque sea ligeramente [105].

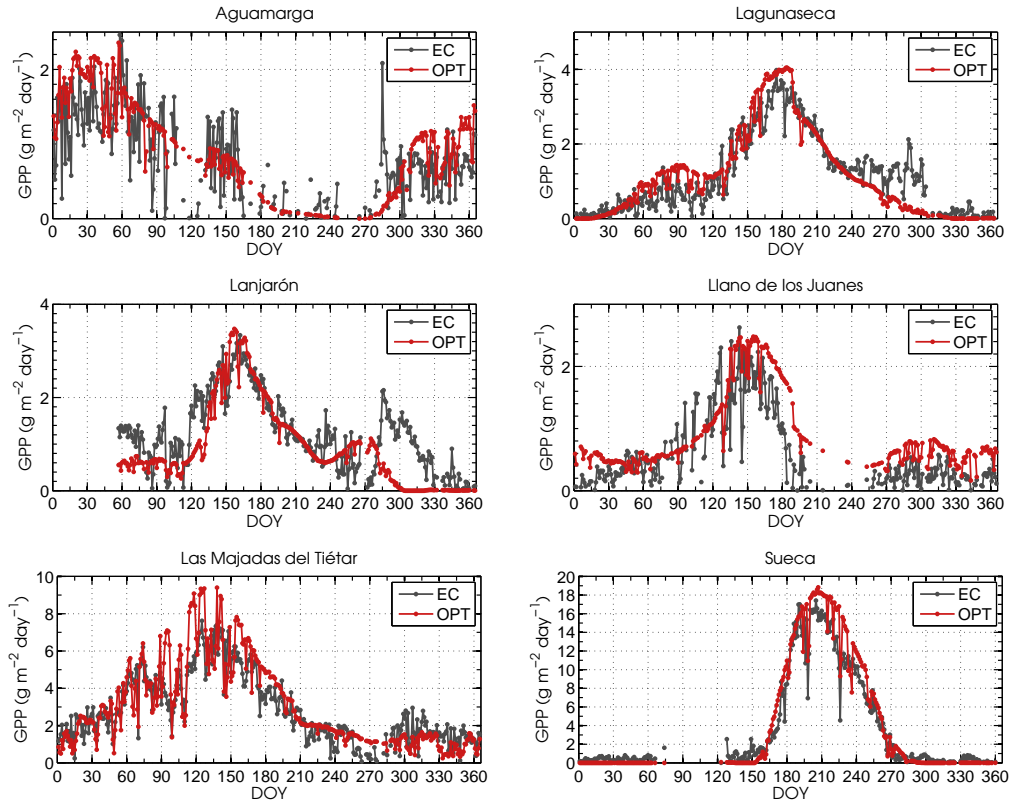


Figura 3.5. Evolución temporal de la GPP calculada con el modelo optimizado y la GPP procedente de las torres EC para el año 2008.

En los distintos diagramas de dispersión entre GPP_{OPT} y GPP_{EC} (figura 3.6) se observan pendientes cercanas a la unidad (0.8 en el peor caso para Aguamarga y 1 para el mejor caso de Lagunaseca) y ordenadas cercanas a cero (0.39 en el peor caso para Llano de los Juanes y -0.01 para Lagunaseca en el mejor caso); esto hace que los datos estimados se distribuyan a lo largo de la bisectriz como cabría esperar en un modelo que obtenga resultados óptimos. Los distintos diagramas también indican que en las torres de menor productividad las dispersiones son mayores. Esto es atribuible a que en estos sitios de validación la GPP es menor y coexiste con eventos de ventilación

de CO_2 del subsuelo. Este efecto no es considerado adecuadamente por la metodología habitual para la estimación de la GPP mediante datos EC, lo que puede constituir una fuente de error [106].

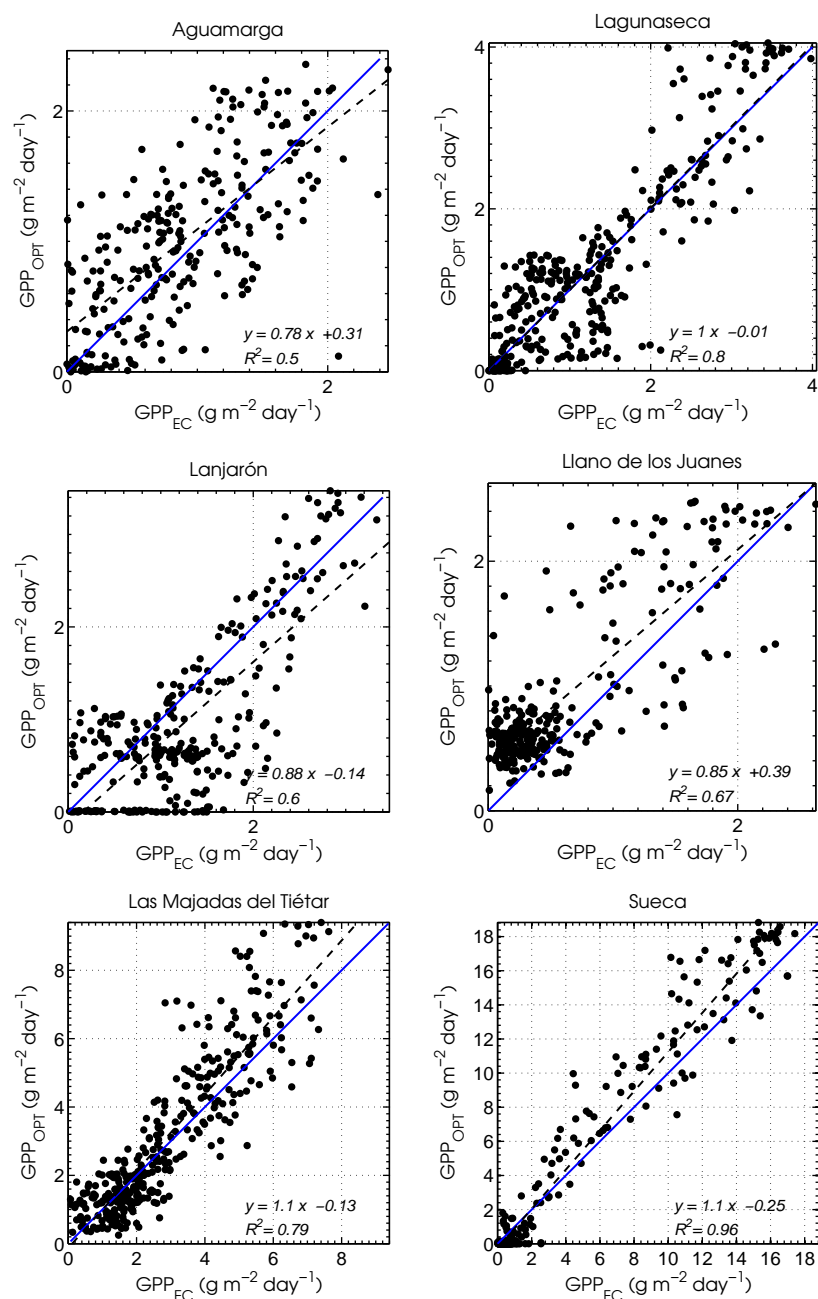


Figura 3.6. Diagramas de dispersión entre la GPP calculada con el modelo optimizado y la GPP procedente de todas las torres EC consideradas en el año 2008.

Por otra parte, la tabla 3.3 muestra los estadísticos obtenidos para cada uno de los sitios de validación en el año 2011. La GPP del modelo optimizado para ese año emplea datos de PAR estimados a partir de datos de satélite. A nivel climático, y de acuerdo con los calendarios meteorológicos que suministra AEMet, el año 2011 [107] se caracterizó por ser un año húmedo y muy cálido.

Tabla 3.3. Resultados de la validación directa para cada una de las torres EC disponibles para el año 2011.

Zona de estudio	MBE	MAE	RMSE	rMBE	rMAE	rRMSE	R^2
Aguamarga	-0.14	0.40	0.54	-0.10	0.29	0.40	0.70
Cortes de Pallás	-0.17	0.54	0.75	-0.094	0.31	0.43	0.80
Llano de los Juanes	0.27	0.37	0.54	0.43	0.55	0.81	0.75
Las Majadas del Tiétar	-0.015	0.80	0.99	-0.0045	0.25	0.30	0.86

Los estadísticos obtenidos para el año 2011 (tabla 3.3) son similares a los mostrados para el año 2008, lo que pone de manifiesto su consistencia temporal. El uso de estimaciones de PAR a partir de datos satélite o a partir de datos puntuales no tiene un impacto muy significativo en la calidad de la GPP calculada. Este resultado, por otro lado, era esperable ya que la validación llevada a cabo en las publicaciones (apéndices A y B) así lo indicaba. Aún así, hay que remarcar que en las torres de las que se disponen de datos para los dos años considerados simultáneamente, esto es Aguamarga, Llano de los Juanes y las Majadas del Tiétar, se observa un pequeño incremento de la correlación del 2011 respecto a 2008. Este incremento podría atribuirse a una mejor estimación de la PAR debido al uso de datos de satélite. Se ha comprobado (apéndice B) que el producto de satélite ofrece correlaciones temporales ligeramente más altas cuando se valida con datos *in situ*, mientras que el resto de variables de entrada del modelo optimizado son idénticas.

El resto de torres presentan correlaciones más altas (e.g., $R = 0.93$ en Majadas del Tiétar), indicando que el modelo captura la evolución temporal y la GPP en términos absolutos. Cuando se comparan estas estimaciones con las de 2008, se observa que existe un incremento de la GPP máxima para el año 2011 en la torre de las Majadas del Tiétar, probablemente relacionado

con una mayor disponibilidad de agua ya que se registró un incremento total de las precipitaciones en 2011 de 110 mm anuales (15%) respecto a 2008. Las series temporales y los diagramas de dispersión para el año 2011 se muestran en las figuras 3.7 y 3.8. En este caso también existe un buen acuerdo con las estimaciones *in situ*, mostrando una distribución próxima a la bisectriz (especialmente destacan los buenos resultados de Aguamarga y las Majadas del Tiétar).

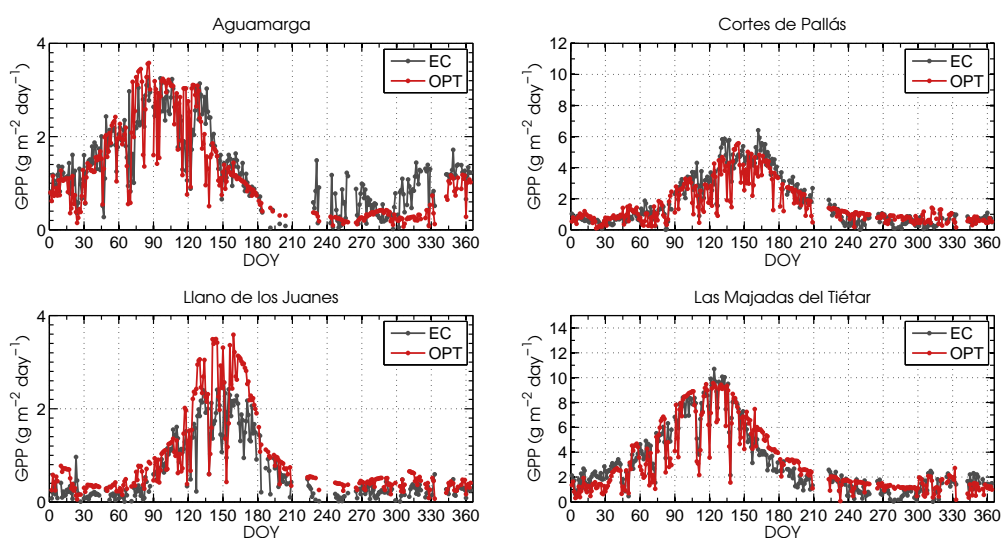


Figura 3.7. Evolución temporal de las estimaciones de la GPP calculada con el modelo optimizado y la GPP procedente de las torres EC para el año 2011.

3.2.2. Validación indirecta comparando con otros productos operacionales

La validación indirecta permite evaluar el producto derivado mediante teledetección a partir de la comparación con otros productos actuales obtenidos a partir de otros sensores u otros algoritmos. Este método complementa a la validación directa en cuanto a que proporciona información de la precisión y consistencia del producto estimado respecto a los actualmente en uso. La validación indirecta permitirá obtener información sobre el rango de variabilidad de cada producto, su precisión y comportamientos anómalos que se pongan de relieve al intercompararse entre ellos. Esta validación sólo

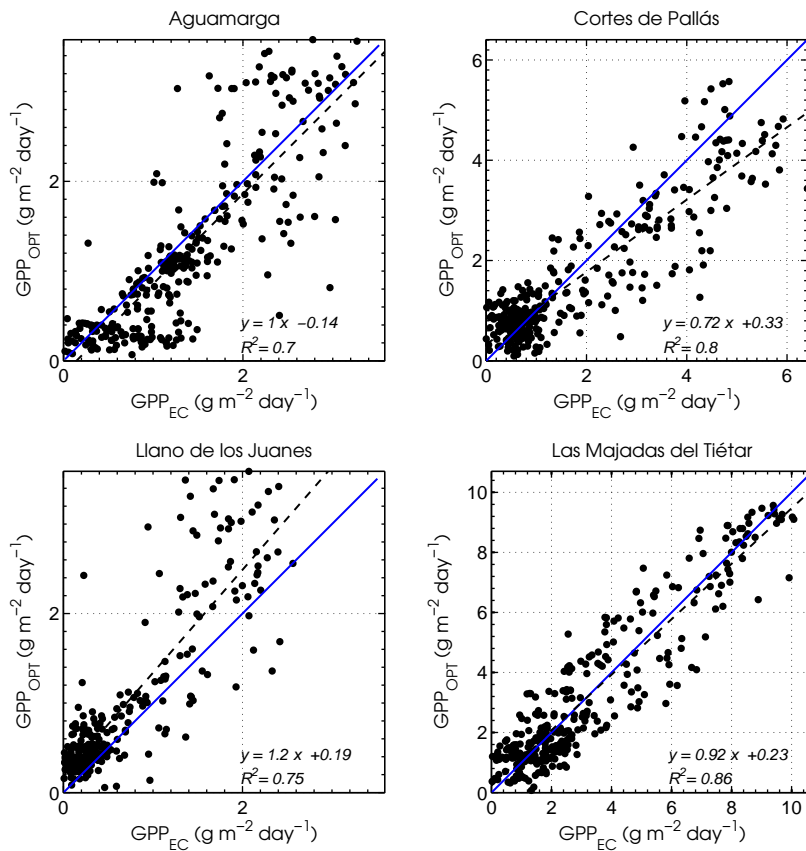


Figura 3.8. Diagramas de dispersión entre la GPP calculada con el modelo optimizado y la GPP procedente de todas las torres EC consideradas en el año 2011.

se llevará a cabo para la GPP del año 2011, pues el producto DMP no está disponible para el otro periodo (2008).

El primero de los productos considerados es el desarrollado por MODIS para el cálculo de la GPP (MOD17A2), disponible en <http://modis.gsfc.nasa.gov/>. El algoritmo de MODIS emplea también un PEM con un factor corrector de la eficiencia por bajas temperaturas (T_{min}) y otro por estrés hídrico que se estima a partir del déficit de presión de vapor (VPD) [6]. La resolución temporal es de 8 días y la espacial de 1 km. La validación y verificación de este producto es una línea activa de investigación y existen dos grandes proyectos que consideran múltiples tipos de cubierta vegetal, múltiples climas [108] y datos de torres EC de la red americana AmeriFlux [109]. Hay que remarcar que, aunque existen validaciones

globales, los estudios citados se centran principalmente en ecosistemas sin déficit hídrico. Así que los resultados de MODIS en estas zonas (matorrales o herbáceas), tan abundantes en el clima mediterráneo, es una tarea que aún no se ha realizado exhaustivamente.

El otro producto considerado para la validación indirecta es el desarrollado dentro de la iniciativa europea Copernicus, denominado “productividad de materia seca” (*Dry Matter Productivity*, DMP) [110] y está disponible en la dirección <http://land.copernicus.eu>. Este producto también emplea un PEM, pero considera únicamente un factor corrector de temperatura, a diferencia del modelo optimizado y el de MODIS, que cuantifican tanto estrés térmico como hídrico. El producto DMP modeliza la eficiencia únicamente como una función no lineal en forma de campana asimétrica dependiente de la temperatura media diaria, alcanzando la máxima eficiencia a 22 °C y dejando de fijarse carbono para temperaturas menores de 0 °C o superiores a 40 °C. Se calcula a una resolución espacial de 1 km y temporal de 10 días y, en el momento de redacción de esta tesis, no ha sido sometido a ningún proceso de validación. Dicho producto no proporciona directamente estimaciones de GPP, pero se ha relacionado con esta aplicando los factores de escala necesarios que aparecen en la bibliografía [23], [111] para de esta forma poder realizar la intercomparación. Por lo tanto, los resultados obtenidos de la comparación con este producto han de ser considerados con una mayor cautela.

En la tabla 3.4 se muestran algunas características generales de los productos intercomparados.

Tabla 3.4. Características de los productos de GPP utilizados en la intercomparación.

Producto	Res. esp.	Res. temp.	Factores
MODIS (MOD17A2)	1 km	8 días	T_{\min} , VPD
Copernicus (DMP)	1 km	10 días	T
Optimizado (OPT)	1 km	1 día	T_{\min} , C_{ws}

Previamente a la validación indirecta, se ha realizado una comprobación de la calidad de los dos productos operacionales. El objetivo es estimar sus errores también sobre los datos *in situ* disponibles en 2011. En la tabla 3.5 se muestran los estadísticos obtenidos de los tres productos de GPP (GPP_{MODIS} , GPP_{DMP} y GPP_{OPT}).

Tabla 3.5. Resultados de la comprobación de la calidad para los píxeles donde se ubican cada una de las torres EC disponibles en el año 2011. Se muestran los estadísticos obtenidos de los tres productos considerados GPP_{MODIS} , GPP_{DMP} y GPP_{OPT} .

Torre EC	Producto	MBE	MAE	RMSE	R^2
Aguamarga	GPP_{OPT}	-0.14	0.40	0.54	0.70
	GPP_{MODIS}	-0.22	0.47	0.53	0.64
	GPP_{DMP}	1.4	1.4	1.701	0.58
Cortes de Pallás	GPP_{OPT}	-0.17	0.54	0.75	0.80
	GPP_{MODIS}	0.14	0.80	0.98	0.76
	GPP_{DMP}	3.1	3.1	3.3	0.70
Llano de los Juanes	GPP_{OPT}	0.29	0.37	0.54	0.75
	GPP_{MODIS}	0.95	0.96	1.1	0.57
	GPP_{DMP}	2.3	2.3	2.6	0.50
Las Majadas del Tiétar	GPP_{OPT}	0.015	0.80	0.99	0.86
	GPP_{MODIS}	-0.73	1.3	1.7	0.76
	GPP_{DMP}	2.0	2.2	2.7	0.80

La tabla 3.5 pone de manifiesto que, para las cuatro zonas, el modelo optimizado captura mejor (coeficientes de determinación más altos) la evolución temporal de las estimaciones con datos *in situ*. MODIS obtiene marginalmente mejores resultados que el modelo optimizado en términos de RMSE para Aguamarga y MBE para Cortes de Pallás. Por otro lado, el producto de GPP_{DMP} presenta en casi todos los casos las peores estimaciones, obteniendo en muchos casos errores del mismo orden que la magnitud a estimar. Solo mejora en términos del coeficiente de determinación a la GPP_{MODIS} en las Majadas del Tiétar.

En la figura 3.9 se muestran los perfiles temporales de los tres productos de GPP, en $g\ m^{-2}\ día^{-1}$. Los productos GPP_{MODIS} y GPP_{DMP} no coinciden temporalmente (debido a su diferente resolución temporal), por lo que

se han calculado las diferencias con las estimaciones *in situ* y el modelo optimizado (a resolución diaria) de forma independiente. A partir de una primera inspección visual, la figura 3.9 pone de manifiesto algunos aspectos que los estadísticos anteriores ya habían indicado. Así, por ejemplo, se observa que la GPP_{DMP} sobreestima de forma sistemática la GPP en todas las torres, mientras que la GPP_{MODIS} subestima los máximos de GPP para todas las torres excepto para Llano de los Juanes y, en general, muestra una caída demasiado gradual de la GPP respecto del máximo (en torno a principios mayo). Este descenso de la GPP es incompatible con el descenso más drástico y que se observa, por ejemplo, en Cortes de Pallás, Llano de los Juanes y las Majadas del Tiétar. Los perfiles del modelo optimizado sí capturan este descenso más rápido de la GPP, porque incluyen el factor C_{ws} que es capaz de incorporar convenientemente la información relevante para describir situaciones de estrés hídrico. Esta puede ser una de las razones más determinantes de la mejora de R^2 de la GPP_{OPT} frente a los productos GPP_{MODIS} y GPP_{DMP} . El producto de GPP_{DMP} no posee ningún factor para capturar este efecto, que es importante para el área de estudio de este trabajo, mientras que la GPP_{MODIS} posee un factor dependiente del déficit de presión de vapor para modelizarlo. Estos resultados están en concordancia con los de otros autores, que también habían indicado que el factor de GPP_{MODIS} puede ser la opción más óptima para incluir los efectos del estrés hídrico en PEMs [10],[4]. Así pues, un balance de agua, como el que subyace en el factor C_{ws} , podría resultar más conveniente.

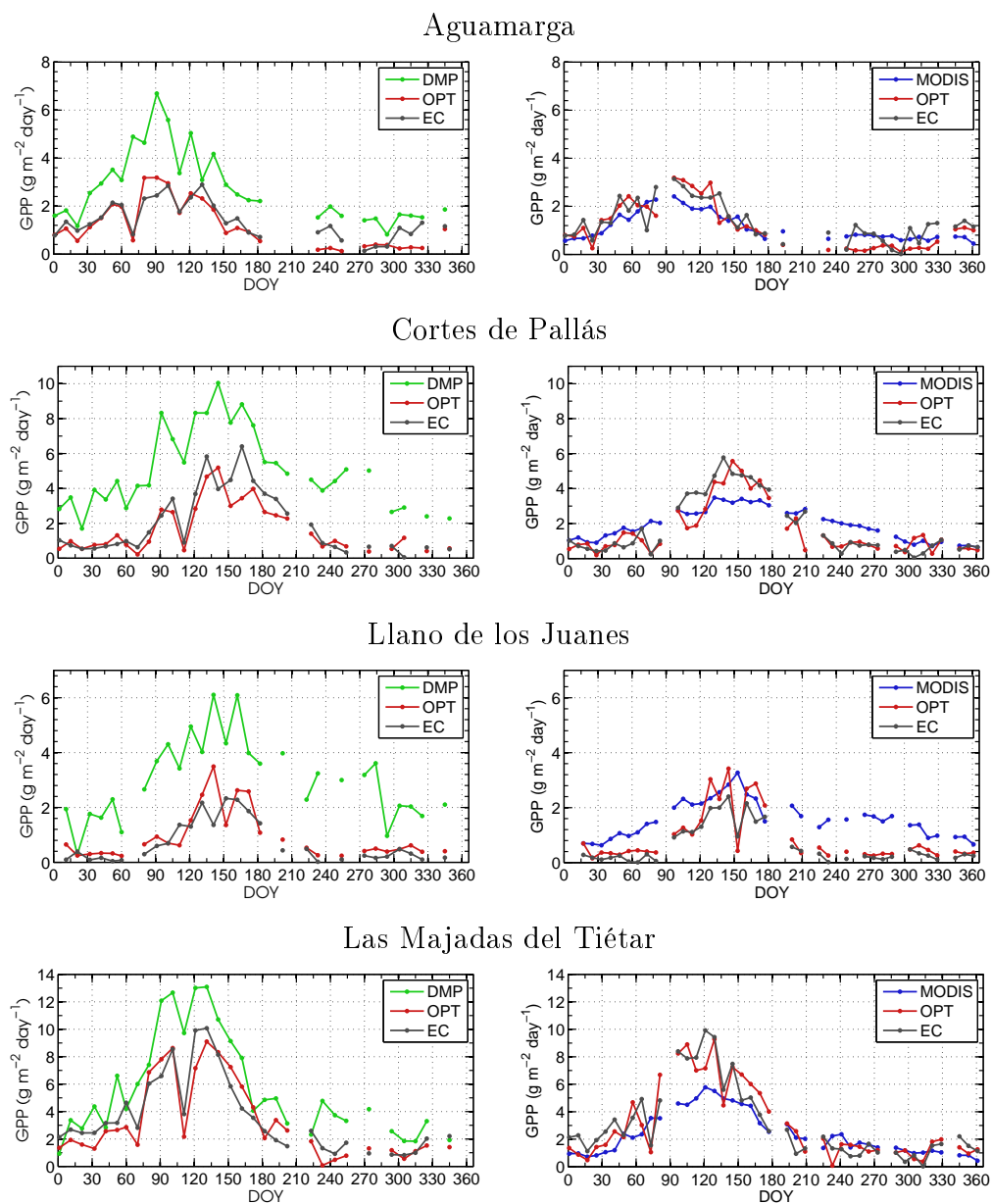
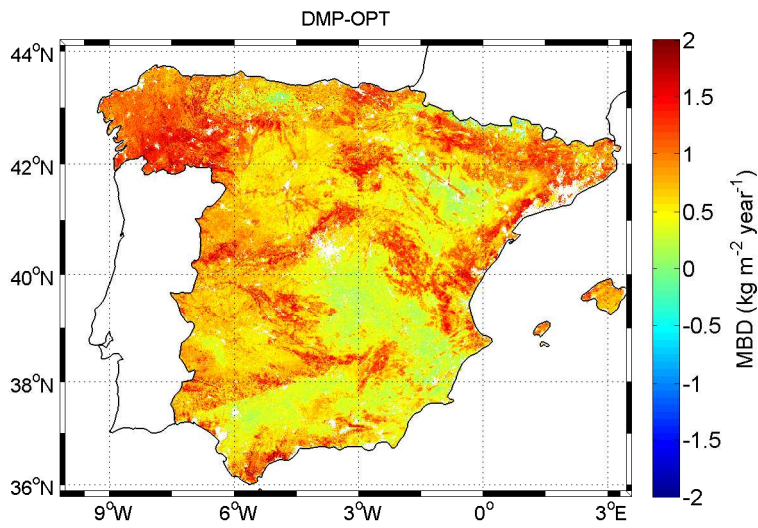


Figura 3.9. Evolución temporal de la GPP de los tres productos operacionales y las estimaciones con datos de torres EC para el año 2011.

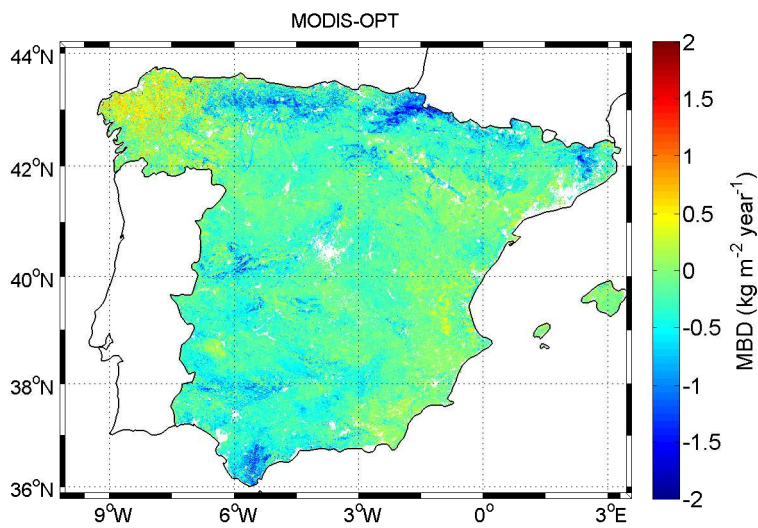
Una vez realizada la comprobación de la calidad de los tres productos de GPP sobre las torres EC, se muestran a continuación los resultados de su intercomparación. Previamente se ha realizado el cálculo de los valores anuales de GPP del resto de productos. Dado que existen huecos en las

series de GPP_{DMP} y GPP_{MODIS} , se han rellenado empleando una técnica de interpolado lineal simple y, posteriormente, se han integrado mediante la técnica trapezoidal. La GPP_{OPT} no presenta huecos porque su f_{APAR} , imposible de estimar en presencia persistente de nubes o nieve, ha sido suavizada e interpolada a escala diaria (sección 2.1.2).

La figura 3.10 muestra un mapa de diferencias entre la GPP anual. La GPP_{DMP} muestra una clara sobreestimación (colores cálidos) con respecto a nuestros cálculos (casi del 100 % si comparamos con la figura 3.3), sobre todo en las zonas de mayor productividad anual. Por otro lado, cuando se compara la GPP_{OPT} con la GPP_{MODIS} se observa que, en general, no hay diferencias tan marcadas. Las diferencias más significativas (en colores fríos) aparecen en zonas de vegetación densa como en la selva de Irati, Cornisa Cantábrica, Pirineo Catalán en el norte y el Parque Natural los Alcornocales en el sur de la península (en zona de Cádiz/Málaga). En las zonas de vegetación especialmente densa, donde la clasificación la indica como bosques de hoja ancha, la GPP_{OPT} obtiene valores más elevados. Así pues, estas diferencias parecen estar relacionadas con discrepancias entre clasificaciones que afectan directamente a los correspondientes valores de ε_{max} .



(a)

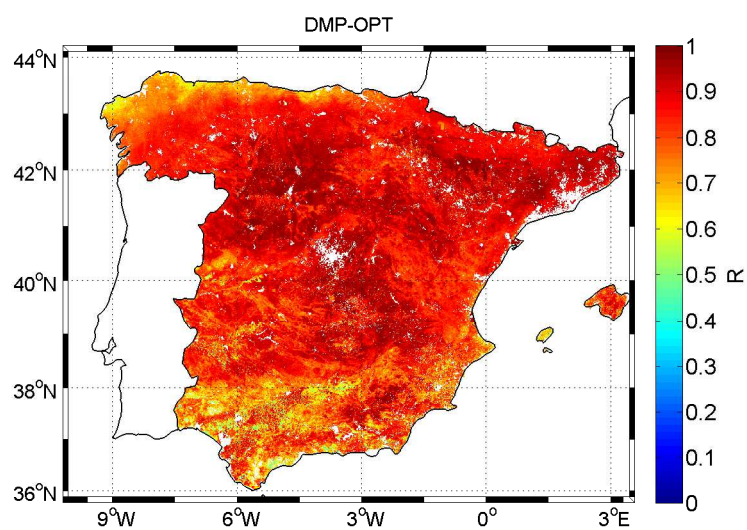


(b)

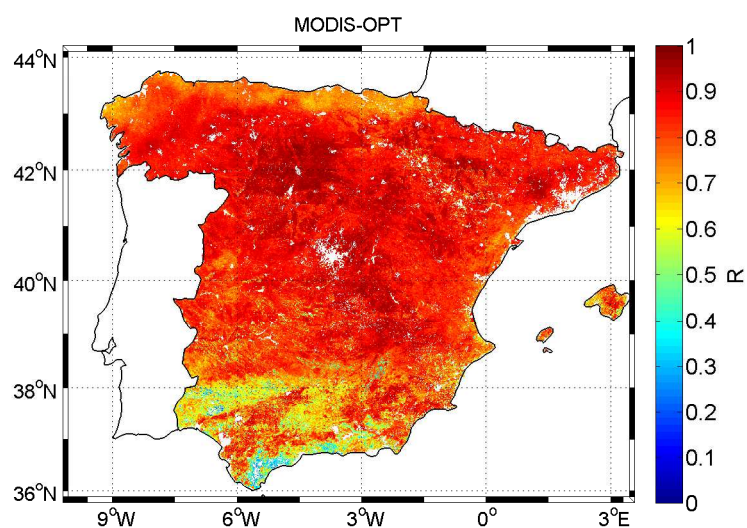
Figura 3.10. Cálculo de los sesgos (MBD) entre el modelo optimizado y DMP (a) y el modelo optimizado y MODIS (b) para la GPP anual en 2011³.

³En esta sección hablamos de diferencias (MBD) en lugar de errores (MBE), porque la comparación se realiza entre productos similares y no con datos *in situ* de torres EC.

Adicionalmente al análisis de los sesgos en términos de GPP anual de los diferentes productos, también se ha evaluado la consistencia temporal de las series. Para ello, se han calculado los coeficientes de correlación a nivel de imagen entre las diferentes series temporales (GPP_{OPT} , GPP_{DMP} y GPP_{MODIS}). Los resultados se muestran en la figura 3.11 e indican que las correlaciones GPP_{OPT}/GPP_{MODIS} y GPP_{OPT}/GPP_{DMP} son altas en gran parte del territorio, sobre todo en la parte central. Cuando se comparan dichas imágenes con mapas típicos de precipitación anual [101], se observa que las correlaciones decrecen en general en el norte y en algún sitio puntual del sur coincidiendo con unos mayores registros pluviométricos (precipitaciones medias anuales superiores a 700 mm). Las diferencias podrían deberse, fundamentalmente, a la repercusión de la nubosidad en el cálculo de la PAR. La GPP_{OPT} se ha calculado con una PAR obtenida de MSG, que ofrece un óptimo seguimiento de las nubes pero no hay garantía de que sea así en los otros dos productos. Se ha descartado que el descenso de la correlación sea debido a la f_{APAR} porque, como se muestra en [49], la evolución temporal es consistente entre diferentes productos operacionales. Las correlaciones más bajas entre el producto MODIS y el nuestro aparecen en el sur. Estas diferencias pueden deberse fundamentalmente a problemas con la correcta incorporación de los efectos del estrés hídrico. Como ya se ha mencionado, la VPD no es la mejor forma de monitorizar el estrés hídrico en bosques mediterráneos y, en la zona sur, es donde más se manifiestan sus efectos.



(a)



(b)

Figura 3.11. Consistencia temporal entre las estimaciones obtenidas por el modelo optimizado y los modelos operacionales DMP (a) y MODIS (b).

Para concluir se remarcan los siguientes aspectos de la validación directa e indirecta del modelo optimizado:

- La calidad de las estimaciones del modelo optimizado es buena tanto cuando se utiliza PAR a partir de datos de satélite como cuando se utilizan imágenes de PAR derivadas con datos puntuales de estación.
- Existe un buen acuerdo (altas correlaciones) entre las series temporales de GPP a partir de datos EC y a partir del modelo optimizado (llegando a 0.98 para Sueca en 2008 y 0.92 para la torre de las Majadas en 2011).
- Los diagramas de dispersión muestran relaciones lineales entre ambas estimaciones, con pendientes cercanas a la unidad (e.g. Lagunaseca en 2008 y Aguamarga en 2011) y sesgos próximos a cero (e.g. Lagunaseca y las Majadas del Tiétar en 2008 y Aguamarga en 2011), lo que indica la idoneidad de nuestra metodología para las zonas de validación estudiadas.
- Las correlaciones más bajas entre la GPP_{OPT} y GPP_{EC} se presentan en las zonas con vegetación escasa (GPP baja). Por otra parte, GPP_{EC} es más difícil de estimar en las mismas por la existencia de eventos de ventilación de CO_2 en el subsuelo. Dicho efecto no se modeliza adecuadamente en los protocolos utilizados para procesar los flujos y presentan un peso relativo considerable en este tipo de cubiertas. Por otro lado, GPP_{OPT} presenta diferencias al suponer $f_{APAR} = 0$ en ciertas ocasiones en las que está documentado que no debería de ser así.
- La comprobación de la calidad de las estimaciones de los productos GPP_{MODIS} y GPP_{DMP} sobre datos de torres EC indica que el producto DMP obtiene unos sesgos muy elevados (llegando al 100 % en el caso LLano de los Juanes). MODIS, sin embargo, los presenta similares a los del modelo optimizado. En términos de varianza capturada por el modelo, la GPP_{OPT} supera significativamente al resto de productos operacionales.

- La validación indirecta indica que la GPP_{OPT} obtiene unos valores intermedios respecto a los dos productos comparados pero más cercanos a los del producto de GPP_{MODIS} , aunque se muestran diferencias significativas con los bosques de hoja ancha.
- El análisis de la consistencia temporal de las series de GPP de los tres productos revela un acuerdo general, pero se observa una disminución de la R para las zonas de mayor precipitación en el norte y de mayor estrés hídrico en sur (sobre todo con MODIS). En el norte estas discrepancias pueden deberse a diferencias en la PAR y en el sur a las diferentes alternativas para caracterizar el estrés hídrico de cada uno de los productos de GPP.

3.3. Discusión

El objetivo de este apartado es analizar el comportamiento del modelo optimizado en el área de estudio. Para ello, primero se va a estudiar, en términos del coeficiente de correlación en las torres disponibles, el modo en que las diferentes variables aportan “potencial explicativo”⁴. Son conocidas las limitaciones de dicho estadístico (solo es sensible a relaciones lineales, las variables pueden estar correlacionadas entre ellas, etc), pero su interpretación sencilla y utilidad, a pesar de sus limitaciones, ha propiciado su uso en este trabajo. Adicionalmente se procederá a analizar, a escala de imagen y para los dos años calculados, la contribución relativa de las variables de entrada del modelo. Esto permitirá observar la distribución espacial de la relevancia de los parámetros involucrados en su cálculo.

⁴Se ha entremetido el término potencial explicativo porque estrictamente el porcentaje de varianza capturada por el modelo sería R^2 y no R pero, el paso de un estadístico a otro es directo.

3.3.1. Análisis del potencial explicativo del modelo utilizando datos *in situ*

El procedimiento llevado a cabo a consiste en correlacionar las series temporales de las estimaciones mediante datos de torres EC y el producto de las variables de entrada que componen el modelo empleado. Este análisis, independientemente de los sesgos que puedan haber, nos muestra el grado de información temporal que incorpora progresivamente al modelo la inclusión de nuevas variables. En primer lugar, se muestran estas correlaciones para las estimaciones del año 2008 (tabla 3.6) y, posteriormente, para el año 2011 (tabla 3.7).

Tabla 3.6. Valores de la correlación (R) en función de las variables incluidas en el modelo para las seis torres en 2008. AGU hace referencia a la torre de Aguamarga, LAG a Lagunaseca, JUA a Llano de los Juanes, LAN a Lanjarón, MAJ a las Majadas del Tiétar y SUE a la torre de Sueca.

Variables incluidas	AGU	LAG	LAN	JUA	MAJ	SUE
(PAR)	-0.24	0.78	0.52	0.65	0.42	0.73
(PAR, f_{APAR})	0.69	0.90	0.72	0.83	0.86	0.98
(PAR, f_{APAR} , T_{min})	0.69	0.90	0.72	0.84	0.82	0.98
(PAR, f_{APAR} , T_{min} , C_{ws})	0.71	0.90	0.77	0.82	0.89	0.98

Para el año 2008 se observa que, en el caso de Aguamarga, la PAR no es un parámetro que por sí mismo describa demasiada variabilidad temporal ($R \simeq -0.2$) de la GPP anual. No es el caso de la f_{APAR} que, por el contrario, es el principal descriptor de evolución temporal de la misma (pasamos de un valor de correlación de -0.2 a 0.7 cuando es incluida en el modelo). Por otro lado, los dos factores inhibidores de la eficiencia (T_{min} y C_{ws}) producen mejoras sutiles en las estimaciones. En el caso de Lagunaseca se aprecia el efecto inverso, la PAR es el mayor aporte de información al modelo (0.8), la f_{APAR} mejora ligeramente las estimaciones (0.9) y, por último, T_{min} y C_{ws} prácticamente no tienen ningún peso en el modelo. En las zonas de Lanjarón y la dehesa de Majadas del Tiétar se observa una mayor contribución del factor de estrés hídrico. Para el caso de la torre de arrozal de Sueca, su evolución temporal viene bien descrita por la PAR y la f_{APAR} , ya que al ser

irrigada no se han incluido los efectos del factor C_{ws} y el factor corrector por bajas temperaturas (T_{min}) actúa cuando el valor de la GPP es ya, por ausencia de vegetación, nulo.

Tabla 3.7. Valores de la correlación (R) en función de las variables incluidas en el modelo para las cuatro torres en 2011. AGU hace referencia a la torre de Aguamarga, COR a Cortes de Pallás, JUA a Llano de los Juanes y MAJ a las Majadas del Tiétar.

Variables incluidas	AGU	COR	JUA	MAJ
(PAR)	0.26	0.59	0.53	0.42
(PAR, f_{APAR})	0.83	0.82	0.83	0.88
(PAR, f_{APAR} , T_{min})	0.83	0.83	0.83	0.86
(PAR, f_{APAR} , T_{min} , C_{ws})	0.84	0.89	0.87	0.93

Para el año 2011 se observan unos comportamientos similares del modelo frente a las estimaciones de GPP con datos EC. La PAR tiene una importancia marginal en las torres de Aguamarga (ya ocurría en el 2008). La f_{APAR} es la variable más significativa para las cinco zonas, destacando el incremento de correlación observado en Aguamarga y las Májadas del Tiétar (de 0.3 a 0.8 y de 0.4 a 0.9 respectivamente). Esto es debido a que en ambos sitios la vegetación presenta una gran estacionalidad. En Aguamarga debido al rápido desarrollo en primavera con las lluvias y su deterioro en verano por la sequía y las altas temperaturas y, en las Majadas del Tiétar, por la coexistencia de un sotobosque que presenta una alta estacionalidad (por la misma razón que en Aguamarga) con vegetación perenne que presenta escasa estacionalidad. En cuanto al factor T_{min} , no se observa una gran contribución del mismo. Conviene remarcar que las zonas de estudio no están ubicadas en zonas especialmente frías y, por lo tanto, la corrección es casi inexistente. Por último, en el caso del factor C_{ws} , se observan contribuciones positivas al modelo para todas las zonas de estudio en este año, destacando la dehesa de las Majadas del Tiétar. Para mostrar más claramente el impacto del factor de estrés hídrico en las estimaciones del modelo, en la figura 3.12 se muestran, a modo de ejemplo, los perfiles de los datos de estimaciones EC, el modelo sin el factor C_{ws} y el modelo completo.

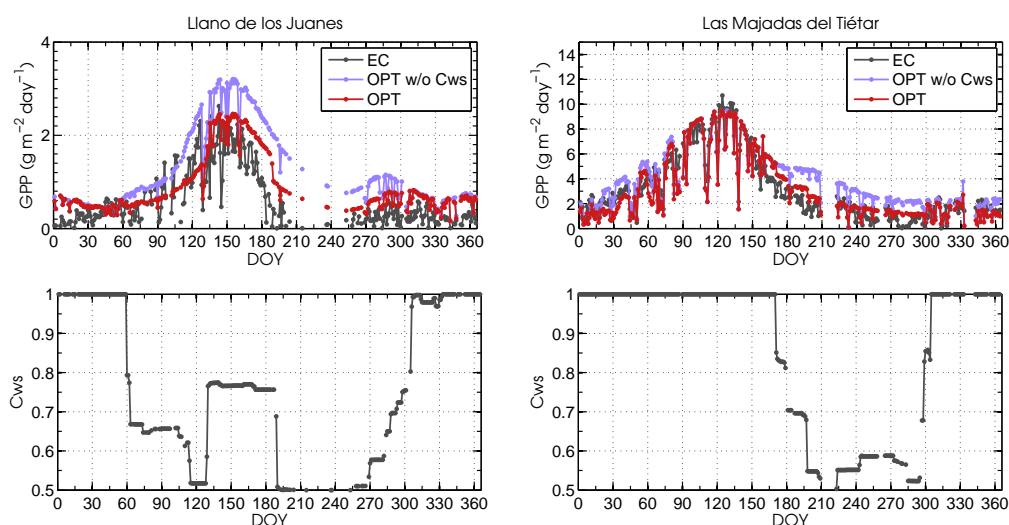


Figura 3.12. Perfiles temporales de GPP para el año 2008 (Llano de los Juanes) y 2011 (Majadas del Tiétar) estimados con datos de torres EC (EC), con el modelo optimizado (OPT) y con el el modelo optimizado eliminado el efecto del factor de estrés hídrico (OPT w/o C_{ws}). En las figuras inferiores se muestran los correspondientes factores de estrés hídrico para cada una de las zonas de estudio.

Los resultados indican, cualitativamente en las figuras 3.12 y cuantitativamente en las tablas 3.6 y 3.7, que el efecto del factor C_{ws} es significativo, ya que elimina sesgos y, adicionalmente, en función de la severidad de la sequía, incorpora en el cálculo de la GPP sus efectos a corto plazo. En el caso de Llano de los Juanes, una zona muy árida, la vegetación está prácticamente sometida a estrés desde la primavera hasta finales del otoño. Aunque se aprecia un incremento de su C_{ws} a partir de abril (debido a un episodio de lluvias en marzo), la productividad es prácticamente nula a finales de junio, probablemente porque la cobertura vegetal se ha degradado por las condiciones de extrema aridez a la que está sometida esta zona en este periodo. Por otro lado, en las Majadas del Tiétar el efecto del estrés hídrico es únicamente evidente en el periodo más cálido y seco del año, es decir, durante el verano. El factor C_{ws} , en este caso, es una corrección que afecta únicamente a la época estival. En términos relativos, su contribución siempre será menor a la de otras variables que presentan una variación anual como la PAR y la f_{APAR} . Destaca, sobre todo, la influencia de la f_{APAR} en especies con acusadas fenologías. Así pues, se puede concluir que, a nivel de

las cuatro zonas de trabajo consideradas, la PAR junto con la f_{APAR} son las variables que mayor potencial explicativo tienen, seguidas del factor C_{ws} y por último, con una contribución casi despreciable, el factor por temperaturas mínimas (T_{min}), siendo la importancia de estas dos últimas de un orden de magnitud inferior.

3.3.2. Análisis a nivel de imagen de la contribución de las variables de entrada al modelo optimizado

En este último apartado se expone una descripción de la respuesta del modelo en toda el área de estudio a nivel de imagen. Para ello se analiza a nivel espacial el grado de información que aportan cada uno de los *inputs* empleados en el cálculo final de la GPP. Se han analizado, al igual que en el apartado anterior, las correlaciones de las estimaciones del modelo optimizado frente a versiones en las que se han eliminado algunos de los *inputs* de la ecuación de Monteith (ecuación 1.2). El procedimiento consiste en la obtención de los incrementos de correlación que supone la adición de cada una de las variable de entrada. De esta forma se evalúa de forma incremental el potencial explicativo introducido por los *inputs* del modelo optimizado en el cálculo de la GPP. El análisis se ha realizado tanto sobre las series calculadas para el año 2008 como para el año 2011. La expresión matemática del cálculo de las correlaciones se muestra en las siguientes expresiones:

$$\begin{aligned}
 R_{\text{PAR}} &= R(\text{GPP}_{\text{OPT}}, \text{PAR}) \\
 R_{f_{\text{APAR}}} &= R(\text{GPP}_{\text{OPT}}, \text{PAR } f_{\text{APAR}}) - R_{\text{PAR}} \\
 R_{C_{\text{ws}}} &= R(\text{GPP}_{\text{OPT}}, \text{PAR } f_{\text{APAR}} C_{\text{ws}}) - R_{f_{\text{APAR}}} \\
 R_{T_{\text{min}}} &= 1 - R(\text{GPP}_{\text{OPT}}, \text{PAR } f_{\text{APAR}} C_{\text{ws}})
 \end{aligned} \tag{3.1}$$

donde R hace referencia al coeficiente de correlación.

En la figura 3.13 se muestran los resultados de dicho análisis. El comportamiento del modelo depende de las condiciones ambientales y fisiológicas de

la vegetación, que han determinado el escenario sobre el que se realizan los cálculos. Por esta razón, se ha incluido información adicional que permita la discusión de los resultados. En este sentido, en la figura 3.14, se muestran resultados del trabajo llevado a cabo por *Martínez & Gilabert* [101] sobre la fenología en España mediante la técnica multiresolución *wavelets* empleando una serie temporal de catorce años (1989-2002) de imágenes AVHRR NDVI. Se han incluido el mes cuando se alcanza el máximo valor del NDVI (t_{max}) y la amplitud anual del ciclo fenológico ($\Delta NDVI$), por considerar que esta información es de utilidad para facilitar la explicación de algunos de los comportamientos en la GPP_{OPT} . Hay que tener en cuenta que el cálculo de dichos parámetros ha sido realizado con otro sensor (AVHRR) y para otro periodo, por lo que la extrapolación de las conclusiones de dicho trabajo a nuestros resultados se debe hacer con cierta cautela.

En general, en la figura 3.13 se observa que para ambos años la variable más influyente en el cálculo de la GPP es la PAR, seguida de la f_{APAR} . La f_{APAR} complementa espacialmente a la PAR de forma que, donde no domina el potencial explicativo de la PAR, domina la f_{APAR} . Estos dos *inputs* del modelo por sí mismos son capaces de capturar gran parte de la dinámica temporal y espacial de la GPP. En otra escala, como se muestran en los ejes de las correlaciones, aparecen los factores que caracterizan el estrés hídrico y el de bajas temperaturas, siendo más influyente el primero. El incremento en la correlación proporcionado por estos factores es pequeño y la razón de esto se debe, en parte, a que no están actuando durante toda la serie temporal, al contrario que la PAR o la f_{APAR} en cubiertas con fenologías marcadas. Estos factores son un ajuste fino en el cálculo de la GPP y, aunque en términos de correlación su efecto no sea muy elevado, su impacto en los errores, como el sesgo, puede ser muy importante (figura 3.12). Estos resultados coinciden, en parte, con los resultados en torres EC, ya que se mostraba que tanto la PAR como la f_{APAR} eran las variables más influyentes sin que la PAR destacase claramente frente a la f_{APAR} . Este comportamiento se debe fundamentalmente a que las pocas torres EC disponibles no son representativas de todo el área de estudio. Por esta razón, consideramos que es de interés realizar este análisis cubriendo todo el área de estudio.

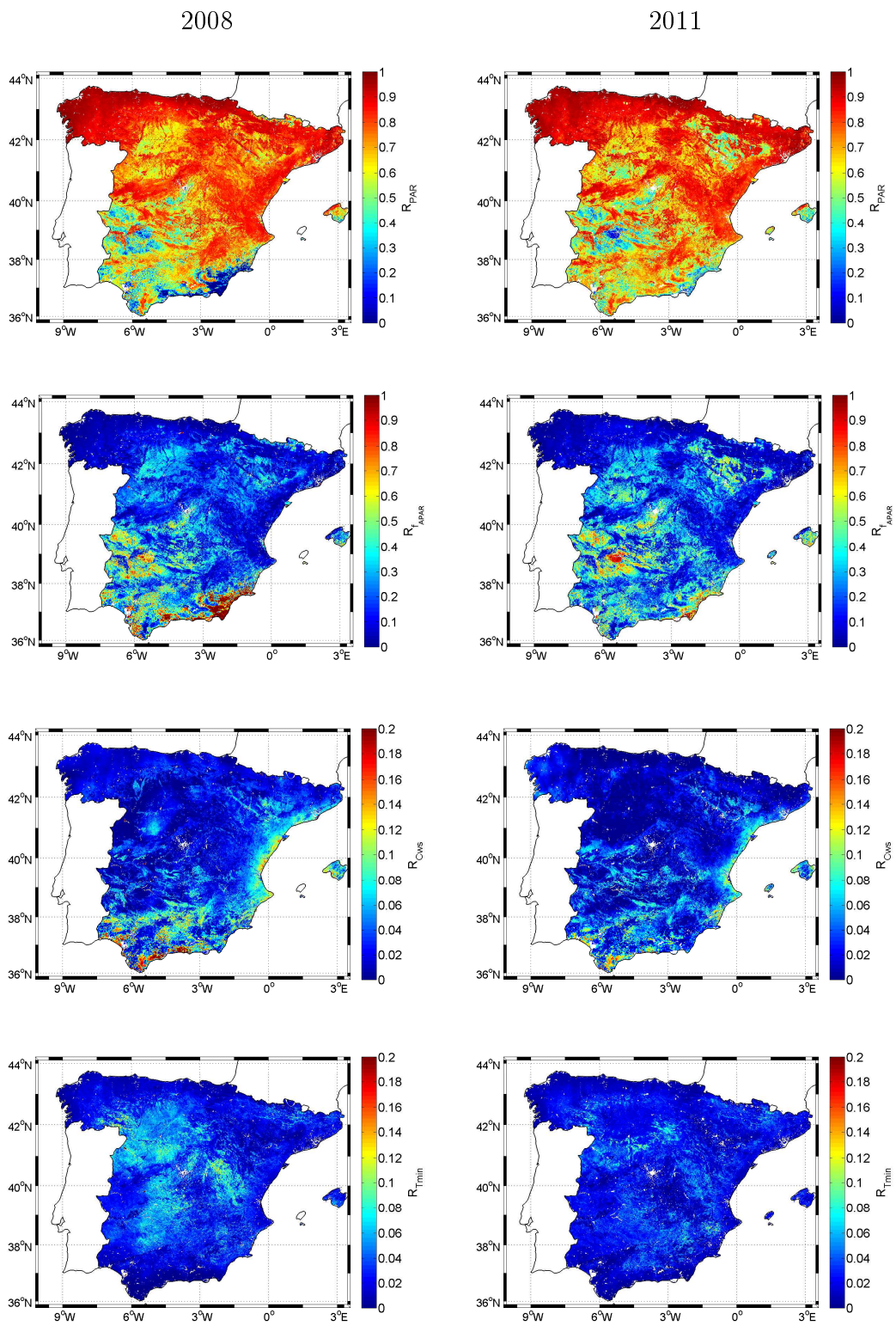


Figura 3.13. Correlaciones temporales respecto al modelo optimizado para cada una de las variables involucradas.

La comparación cualitativa de las correlaciones obtenidas con los parámetros fenológicos incluidos en la figura 3.14 indica que la PAR gobierna la GPP cuando la variabilidad fenológica es baja ($\Delta\text{NDVI} \ll$), a excepción de algunas zonas de la Cornisa Cantábrica. La razón de este comportamiento diferente radica en que, aunque exista una fenología importante en ellas, hay un alto registro de precipitaciones (gran variabilidad en la PAR debido a la presencia de nubes) y los máximos de NDVI (junio/julio) se producen coincidiendo con el máximo solar. Al no existir un desfase temporal importante entre los máximos de la PAR y la f_{APAR} (presenta una evolución temporal similar al NDVI), las dos variables (correlacionadas) pueden ser empleadas para explicar la evolución temporal del ciclo anual de la GPP en esta zona. Cuando se analiza la $R_{f_{\text{APAR}}}$ se observa que la combinación de la información del momento de máximo desarrollo junto con la amplitud del ciclo fenológico es la clave de su potencial explicativo. En este caso, se observa lo contrario, amplitudes elevadas en la fenología de la vegetación se relacionan con una mayor importancia de la variable f_{APAR} . Este carácter se agudiza más aún a medida que el máximo fenológico se separa del máximo solar y donde las precipitaciones son más escasas (contrariamente a lo que ocurría con la R_{PAR}).

En cuanto al patrón espacial de $R_{C_{\text{ws}}}$, se observa un gradiente noroeste-sureste coincidiendo con el gradiente conocido de precipitaciones y temperaturas existente en la península. $R_{C_{\text{ws}}}$ presenta sus valores más bajos en el NO, donde las temperaturas son más suaves y el régimen de precipitaciones presenta registros elevados. Mientras que en la zona más árida del área de estudio (SE), en condiciones casi desérticas, las altas temperaturas diarias y las escasas precipitaciones anuales hacen que la varianza del factor C_{ws} sea mayor y $R_{C_{\text{ws}}}$ aumente por tanto su valor. Además se observa que, en general, en zonas con fenología baja y máximos más tardíos (zona de Cádiz), se registra la máxima influencia del factor de estrés hídrico. Esto se debe a que dicho *input* se activa en general a mitad de primavera y deja de influir a final de verano, alcanzando su valor mínimo en julio/agosto coincidiendo con el periodo más seco y cálido del año. Si la vegetación no ha sido dañada, es en este periodo cuando las condiciones meteorológicas son más adversas para la realización de fotosíntesis debido al déficit hídrico,

la alta irradiación y las altas temperaturas. Por último, el factor de estrés por bajas temperaturas T_{\min} presenta una menor importancia, pero hay que destacar que su influencia es mayor en la zona central de país donde suelen registrarse las temperaturas mínimas más bajas.

Cuando se buscan diferencias entre los dos años analizados (2008 y 2011), se observa que en general los patrones espaciales de las correlaciones se mantienen. Se pueden destacar como diferencias significativas el incremento ($\simeq 0.3$) de $R_{f_{\text{APAR}}}$ de 2008 a 2011 en la zona de cultivos irrigados al noreste de la imagen y una acusada disminución ($\simeq 0.4$) en el caso de la sureste. En la figura 3.15 se muestra la media de las precipitaciones acumuladas mensuales para ambos años, observándose que para el año 2008 presenta una mayor variabilidad (se aprecia una evolución temporal mensual más errática que en el 2011). La otra variable mostrada en la figura son las temperaturas mínimas medias. Se observa que en 2008 se registra en la mayor parte de los meses temperaturas mínimas inferiores que en 2011, destacando un aumento de las temperaturas más suave a principios de año y un descenso más rápido a partir de octubre que en 2011. Estas características diferenciadas de las principales variables que influyen en el cálculo de los factores de estrés son las principales causantes de las diferencias observadas entre ambos años de estudio.

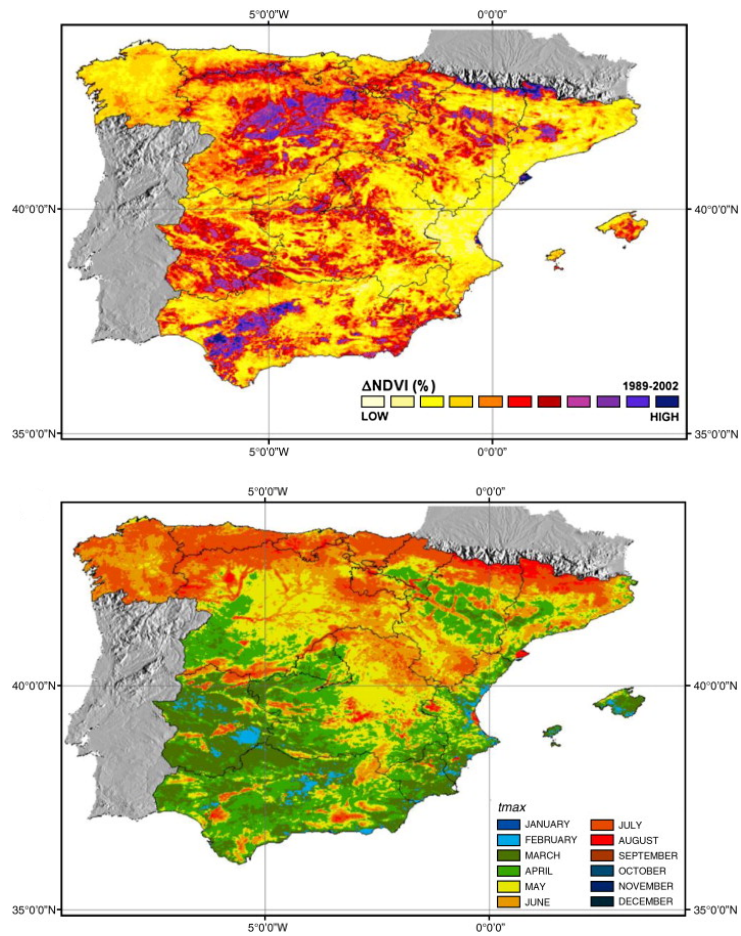


Figura 3.14. Parámetros fenológicos de la vegetación ΔNDVI y t_{max} (amplitud del ciclo fenológico y mes del máximo de la vegetación, respectivamente) obtenidos mediante la técnica *wavelets* en series temporales del índice de vegetación NDVI. Fuente: *Martínez & Gilabert* [101]

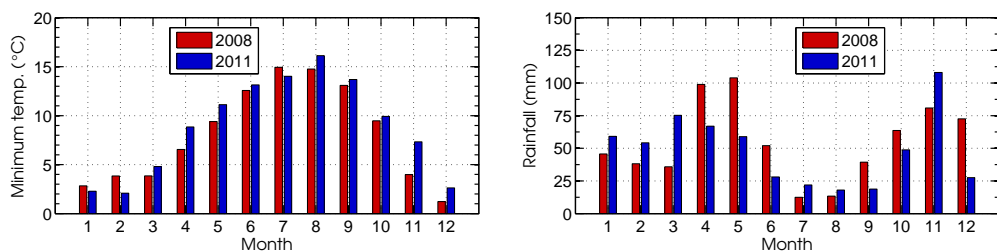


Figura 3.15. Precipitación acumulada mensual y temperatura mínima media mensual para los dos años en los que se ha calculado la GPP.

Así, aunque se aprecia el gradiente norte-sur de la $R_{C_{ws}}$ coincidiendo con el conocido gradiente de temperaturas y de precipitaciones de la península, también es significativo el aumento de la correlación de dicho factor y el de T_{\min} en 2008 frente a 2011. En el caso de C_{ws} esta diferencia es atribuible a una mayor variabilidad de la serie temporal de la precipitación como se ha indicado anteriormente. Sin embargo, en el caso de T_{\min} , la influencia de las temperaturas mínimas mensuales en ambos años es más confusa. Se ha de tener en cuenta que el factor T_{\min} ejerce su influencia para todos los tipos de cubierta para temperaturas mínimas diarias inferiores a 8 °C. De acuerdo con los valores promedio mostrados en la figura 3.15, las temperaturas mínimas mensuales fueron en general inferiores, destacando especialmente un descenso más rápido de las temperaturas a finales de 2008.

Las ideas expuestas acerca del comportamiento del modelo empleado constituyen una serie de rasgos generales, pero se ha realizado un análisis adicional de la figura 3.2 en zonas seleccionadas y representativas dentro de la alta heterogeneidad del área de trabajo de esta tesis (España).

La figura 3.16 muestra dicha selección que, debido a sus diferentes características climáticas y de vegetación, son de interés para ser tratadas con detalle. A continuación se muestran sus rasgos más relevantes:

- La zona “A” se corresponde con el Parque Natural Los Alcornocales, la mayor masa conservada y productiva de alcornocal de la Península Ibérica. Está situada en la provincia de Cádiz y parte de Málaga. Destaca por la humedad proveniente de la costa, que se acumula formando bosques de niebla en valles estrechos y profundos. Está compuesta principalmente por un denso alcornocal, laurisilva, quejigos y roble andaluz.
- La zona “B” se ubica en Almería. En esta zona se tiene una marcada tendencia a la aridez, que incrementa a medida que nos aproximamos a la costa y disminuye la altitud. La vegetación deriva a espartales que dan al paisaje un claro carácter desértico acrecentado por la fuerte acción antrópica.

- La zona “C” se corresponde con un bosque mediterráneo localizado al norte de la provincia de Castellón. Al igual que en la zona anterior, tiene que enfrentarse a condiciones de aridez, pero en este caso menores. Es por esto que el paisaje está compuesto por una baja densidad de pinos por debajo de los cuáles se extienden plantas aromáticas como romeros, salvias, lavanda, madroños, lentisco, jaras, etc.
- La zona “D” corresponde a la Selva de Irati, localizada en el Pirineo oriental navarro y constituye el segundo hayedo-abetal más extenso y mejor conservado de Europa. El alto índice de lluvias que registra esta zona pirenaica provoca que el paisaje esté poblado de arroyos y torrentes, que surcan su abrupto relieve y propician el desarrollo exuberante de la vegetación.
- La zona “E” corresponde a una zona de cultivos de secano en la zona de Castilla y León (Valladolid). Los cultivos de secano ocupan casi todas las tierras de cultivo en esta comunidad, siendo los principales: los cereales (cebada y trigo), los girasoles, la vid y las legumbres que también se cultivan en regadío.
- La zona “F” es representativa de los bosques en la provincia de Asturias. Esta zona pertenece a la Iberia húmeda y tiene una gran masa forestal. Predominan castaños y pinos en las cotas medias y los eucaliptos se instalan en zonas bajas.
- Por último la zona “G” es representativa de una amplia zona en Extremadura cerca del embalse del Zújar, compuesto principalmente por amplias extensiones de pastizal y labor de secano, predominando el primero.

En la zona “A” se observa que es una zona fuertemente dominada por la PAR, así como una de las zonas de más influencia del factor C_{ws} . Aquí se registran importantes precipitaciones y existe una abundante vegetación, pero el periodo de estrés hídrico se prolongó un mes más en 2008 que en 2011, de ahí el ligero incremento en el potencial explicativo de C_{ws} que muestran las figuras. El factor T_{min} no presenta una gran importancia ya que en nuestro análisis de temperaturas mensuales mínimas, solo en febrero

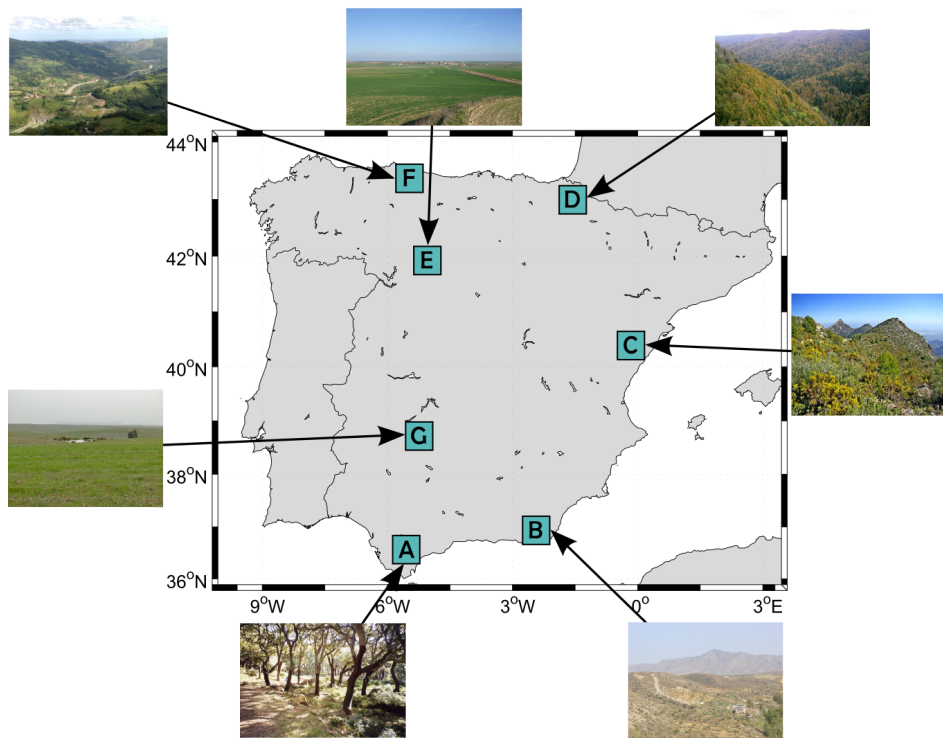


Figura 3.16. Zonas seleccionadas para la discusión de resultados.

de 2011 se alcanzaron temperaturas por debajo de los 10 °C. Si se observa la línea de costa hacia al este, se aprecian también diferencias significativas. Estas se deben principalmente a que, en términos de la precipitación total anual (acumulada), en 2008 fue hasta un 30 % inferior a la de 2011.

Cuando se estudia la zona “B” destaca el incremento de la $R_{f_{APAR}}$ de 2008 frente a 2011. La razón parece ser un comportamiento diferente en la f_{APAR} para ambos años (figura 3.17). En esta zona, aunque la vegetación es muy escasa, el inicio prematuro del déficit hídrico parece ser el causante de un desarrollo significativamente menor al comienzo del año 2008. Por otro lado, el aumento de las precipitaciones al final del año, parece explicar el mayor desarrollo al final del mismo. La mayor importancia de la f_{APAR} se debe a que el principal control sobre la GPP lo ejerce la disponibilidad de una vegetación que llega prácticamente a desaparecer al completo en el periodo de estival, cuando la radiación y la temperatura alcanzan su máximo. Otro aspecto a remarcar es que el C_{ws} , aunque presenta una importante variabilidad, no parece presentar un potencial explicativo importante, ya

que la respuesta de la vegetación al estrés hídrico es casi inmediata. Este tipo de cubiertas vegetales no poseen raíces profundas y son especialmente vulnerables a la sequía, lo que hace que la aparición de daños estructurales en la vegetación sea tan rápida que hacen coincidir las series temporales de la f_{APAR} y C_{ws} .

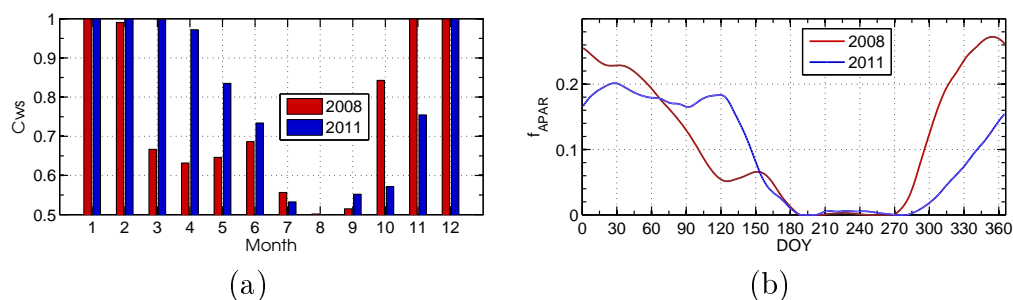


Figura 3.17. Perfil mensual del factor de estrés hídrico (a) y de f_{APAR} (b) para la zona de estudio “B” (Almería) y los años 2008 y 2011.

En la zona “C”, ubicada en la zona de levante, se observa un predominio claro de la influencia de la PAR sobre el resto de entradas que se aplican al modelo. Cerca de la costa mediterránea predominan cultivos perennes (naranjos) y en la zona escogida el bosque mediterráneo, aunque este no es tan árido y degradado como el del zona “B”, presentando una variabilidad temporal baja. En este contexto vemos que la variabilidad de la GPP viene controlada principalmente por la disponibilidad de PAR y por la influencia del estrés hídrico. La precipitación en la zona fue inferior para el año 2008 (530 mm anuales) que en 2011 (650 mm anuales), pero la principal razón del incremento del $R_{C_{\text{ws}}}$ está asociada con un perfil más irregular (mayor varianza temporal de la serie) de la precipitación en 2008, como muestra la figura 3.15 a escala de imagen.

En las zonas “D” y “F” se observa un comportamiento similar. Se caracterizan por una acusada fenología relacionada con la mayor presencia de vegetación de hoja caduca. El valor de la f_{APAR} llega prácticamente a duplicarse desde el mínimo (justo antes del *bloom* primaveral cuyo valor es $\simeq 0.5$) hasta alcanzar el valor máximo a mediados de junio ($\simeq 0.95$) para el caso de la selva de Irati y ligeramente inferior ($\simeq 0.85$) para el bosque de Asturias. La evolución temporal de la vegetación de la escena es simultánea, aunque más abrupta, a la evolución temporal de la PAR. Es por esto que,

al incluir en el modelo la f_{APAR} , el incremento de la correlación no es muy significativo. El resto de *inputs* presentan una importancia baja en el cálculo de la GPP: el estrés hídrico es casi inexistente, al menos en los dos años considerados, haciendo que C_{ws} rara vez tome valores diferentes de la unidad, y el factor T_{min} , que solo se activa cuando gran parte o la totalidad de la cubierta vegetal ha perdido sus hojas, no tiene efecto prácticamente porque son precisamente las bajas temperaturas las que controlan el inicio del proceso de caída de la hoja en la vegetación caduca.

La figura 3.18 muestra información acerca de la zona de estudio “E”. Se aprecia cómo existe una marcada fenología con un máximo estrecho ubicado en abril/mayo como indica también la figura 3.14. Esta amplitud del ciclo fenológico, no conectada de forma directa con el máximo de irradiación anual, hace que la PAR y la f_{APAR} sean las responsables de explicar la evolución temporal de la GPP calculada. Adicionalmente es destacable que en esta zona de estudio adquiere una mayor relevancia el factor T_{min} . La razón de esto es que durante todo el año 2008 se dan temperaturas mínimas inferiores a los 12 °C (umbral de activación del factor para cultivos) y son, en general, inferiores a las de 2011. A partir del umbral indicado, el factor comienza a decrecer la eficiencia (llegando a inhibir totalmente la fotosíntesis a los -8 °C) que, por la forma en la que está implementado (mayor influencia a temperaturas más bajas), hace que sea más patente el incremento de $R_{T_{\text{min}}}$ para el 2008 frente a 2011.

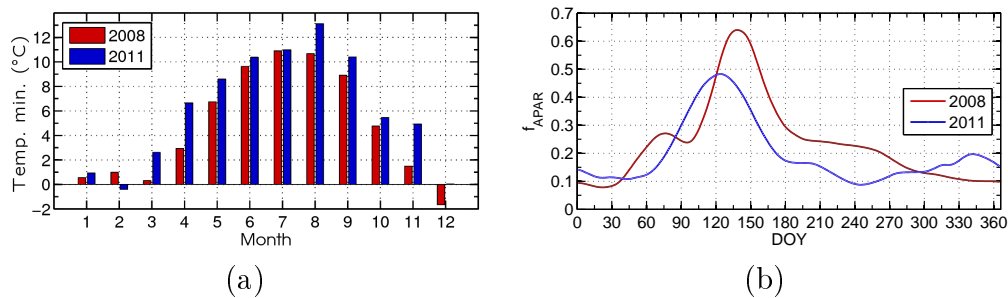


Figura 3.18. Perfil mensual de las temperaturas mínimas mensuales (a) y de f_{APAR} (b) para la zona de estudio “E” (Valladolid) y los años 2008 y 2011.

Por último en la zona “G” se aprecia un incremento de la $R_{f_{\text{APAR}}}$, una casi nula influencia del factor de estrés hídrico y una influencia significativa del factor T_{min} , con un comportamiento interanual similar al mostrado en

la zona “E”. La razón del gran potencial explicativo de la f_{APAR} es el fuerte desacoplo de esta con la PAR, ya que probablemente debido a la presencia de pastizales, esta zona muestra el máximo fenológico alrededor de febrero como se observa en la figura 3.14 (también se ha comprobado sobre el perfil temporal de la f_{APAR} y ambos coinciden). Por otra parte, la influencia de T_{min} se debe a que el máximo del ciclo fenológico para este tipo de vegetación se registra durante el invierno.

En definitiva, de esta sección se puede concluir lo siguiente:

- Los parámetros más influyentes para el cálculo de la GPP en el modelo optimizado son la PAR y la f_{APAR} . El potencial explicativo de la PAR destaca especialmente en especies vegetales con baja fenología y poco sometidas a estrés. Este parámetro caracteriza la gran variabilidad espacial de la GPP a escala diaria y el ciclo anual marcado de la actividad fotosintética, debido a su evolución anual asociada con variaciones en la irradiación solar disponible en el techo de la atmósfera.
- La PAR posee una gran influencia sobre el desarrollo fenológico y, por tanto, sobre la f_{APAR} . De manera que ambas variables suelen estar correlacionadas en cubiertas vegetales con fenologías marcadas si no existe algún factor que desacople su evolución temporal (estrés hídrico, térmico, falta de nutrientes, etc). Así pues, cuando la evolución temporal de ambas variables está desacoplada, la f_{APAR} toma una mayor relevancia en el modelo controlando la productividad casi exclusivamente.
- En otro orden de magnitud, se encuentran los factores correctores de la ε_{max} por bajas temperaturas (T_{min}) y por estrés hídrico (C_{ws}). Estos dos factores presentan una importancia cualitativa menor porque actúan en periodos concretos y reducidos del año.
- El factor T_{min} presenta sus valores más extremos en invierno y en zonas geográficas donde se registran las temperaturas mínimas diarias más bajas. Aunque la PAR es mínima en este periodo, si la vegetación presenta su máximo fenológico en invierno y las temperaturas son

suficiente bajas esto hace que, aún ejerciendo un efecto de segundo orden, la influencia de T_{\min} sea apreciable en los valores de la GPP.

- El factor C_{ws} tiene, en general, una mayor influencia que T_{\min} . Destaca su patrón espacial de influencia a nivel de imagen en zonas cálidas, secas y con una vegetación con su máximo fenológico en los meses de verano. Este factor es clave para poder incorporar al modelo las diferencias interanuales de GPP asociadas con la disponibilidad de recursos hídricos y que son típicas en las zonas mediterráneas.
- Cuando la vegetación es poco resistente al estrés hídrico, los daños estructurales se producen de forma más inmediata y afectan a la f_{APAR} . Como consecuencia, el C_{ws} presenta una menor influencia en estos casos.

Conclusions

In this dissertation an optimized production efficiency model (PEM) to estimate the gross primary production (GPP) over Mediterranean ecosystems at regional scale has been developed. Daily GPP images have been obtained for 2008 and 2011 with a 1 km spatial resolution over peninsular Spain. The inputs for the model (f_{APAR} , PAR and ε) have been obtained using different approaches from *in situ* data, which must be spatially extended, and/or from remote sensing data. Finally, the optimized GPP product has been validated by means of a comparison with *in situ* GPP estimates from the EC data (direct validation) and an inter-comparison with MODIS and Copernicus GPP products (indirect validation).

The main conclusions in relation to the calculation of the inputs of PEM and the validation of the GPP product are summarized below.

1. *Inputs of the production efficiency model*

- a) f_{APAR} . In LSA-SAF the statistical method proposed by Roujean & Bréon (1995) was adopted and validated to operationally deliver f_{APAR} from SEVIRI. This method has been shown to yield intermediate f_{APAR} values (in terms of bias) when compared to different operational f_{APAR} products, and hence, in this dissertation, it has been the method of choice to retrieve daily f_{APAR}

from MODIS BRDF observations. Raw f_{APAR} time series often have gaps and low quality/noisy data due to cloud or snow contamination. The de-noising and gap filling of the series has been carried out using a local regression method (*loess*) that takes into account the quality flags from MODIS data. This adapted *loess* has outperformed other methods to capture series envelope.

- b) PAR. Different methodologies to calculate the PAR from meteorological stations and remote sensing data have been developed. On the one hand, three models have been developed and validated to estimate the daily irradiation (the PAR corresponds to 46 % of the total flux density of solar radiation) from meteorological stations data. The BC model has presented a tendency to under-estimate high global solar irradiation values. The KRR model has shown significantly improved accuracy of its irradiation estimates. Yet, the ANN method, with a MAE = $2.3 \text{ MJ m}^{-2} \text{ day}^{-1}$, has performed even better and, therefore, it has been selected to obtain the solar irradiation images from the temperature and precipitation images derived by ordinary kriging of the data from the 400 stations. On the other hand, the PAR has also been estimated from remote sensing data by the improvement and validation of the DSSF product derived from SEVIRI/MSG images. The improved DSSF product has exhibited a better performance in the estimation of the daily irradiation (bias of $-0.12 \text{ MJ m}^{-2} \text{ day}^{-1}$ and a mean absolute difference of $1.0 \text{ MJ m}^{-2} \text{ day}^{-1}$) than the results from the ANN model. Therefore, the methodology based on remote sensing data has been proposed whenever DSSF images were available (2011) and the ANN model has been used otherwise (2008). It has been observed that the use of DSSF images results in slight improvements in the GPP estimations, although they invariably have a high quality regardless of the methodology used.
- c) ε . The light-use efficiency has been derived from the product of ε_{max} , the maximum light use efficiency of the ecosystem, and two terms that account for the reduction in photosynthesis due to

thermal (T_{\min}) and water stress (C_{ws}). The ε_{\max} value has been assigned according to the land cover type. For this purpose, a hybrid land-cover map has been considered. Special attention has been paid to the two stress terms. In order to monitor the inter-annual differences in productivity due to water stress two alternative approaches have been analysed: (i) the spectral indices (PRI, NDWI, SIWSI, NDII7) calculated from remote sensing images, and (ii) the C_{ws} , which is based on a water balance and computed from meteorological. The results have demonstrated a better performance of the water balance and, therefore, the proposed optimized model incorporates C_{ws} to model the reduction in light use efficiency due to water stress.

2. Validation of the optimized GPP

- a) *Direct validation.* In general, the comparison between the GPP_{OPT} calculated using the optimized model developed in this dissertation and the GPP estimates from EC data has evidenced a good agreement with correlations up to 0.98 in 2008 (e.g. Sueca) and 0.92 in 2011 (e.g. Majadas). However, a disagreement has been observed on sparse vegetation areas with low productivity. The problem has been attributed to the inaccurate modelling of the CO₂ fluxes from EC data due to the soil respiration events that play a significant role in this type of land covers. Very rarely, GPP_{OPT} has failed in some areas where the f_{APAR} estimation has returned incorrect zero values.
- b) *Indirect validation.* The optimized model developed in this dissertation has generated GPP_{OPT} values that lay between the MODIS and Copernicus GPP products. The agreement with the MODIS product has been particularly good. The analysis of the temporal consistency of these three GPP products has revealed a good agreement among them. However, a decrease of the correlation has been observed on those areas with high precipitation (Northern Spain) and high water stress (Southern Spain). Whereas the discrepancies on the northern region could be explained

due to differences in the PAR, the discrepancies on the southern region could be related to the different way the GPP products account for the water stress.

c) *Analysis of the importance of the inputs in the optimized model.*

This analysis has demonstrated that the most significant inputs in the optimized model are PAR and f_{APAR} . The importance of PAR is emphasized on vegetation canopies with low phenology and not affected by stress. However, the f_{APAR} plays a major role when the temporal evolutions of these variables do not coincide. The relevance of the thermal (T_{min}) and the water (C_{ws}) stress terms has been quantified as a lower order of magnitude, with C_{ws} usually being more important than T_{min} . It has been stressed the major relevance of the C_{ws} over hot and dry areas with maximum development during summer months. The C_{ws} has revealed as a key variable to account for the inter-annual variability of the productivity that can be associated with water resources availability typical of the Mediterranean landscapes.

Referencias

- [1] J.L. Monteith, “Solar radiation and productivity in tropical ecosystems,” *Journal of Applied Ecology*, vol. 9, no. 3, pp. 747–766, 1972.
- [2] N.C. Coops, T. Hilker, F.G. Hall, C.J. Nichol, and G.G. Drolet, “Estimation of light-use efficiency of terrestrial ecosystems from space: a status report,” *BioScience*, vol. 60, no. 10, pp. 788–797, 2010.
- [3] C.B. Field, “Ecological Scaling of Carbon Gain to Stress and Resource,” in *Response of Plants to Multiple Stresses*, H.A. Mooney, W.E. Winner, and E.J. Pell, Eds., chapter 2, pp. 35–65. Academic Press, New York, 1991.
- [4] M.F. Garbulsky, J. Peñuelas, D. Papale, J. Ardö, M.L. Goulden, G. Kiely, A.D. Richardson, E. Rotenberg, E.M. Veenendaal, and I. Filella, “Patterns and controls of the variability of radiation use efficiency and primary productivity across terrestrial ecosystems,” *Global Ecology and Biogeography*, vol. 19, no. 2, pp. 253–267, 2010.
- [5] S.J. Goetz and S.D. Prince, “Modelling terrestrial carbon exchange and storage: evidence and implications of functional convergence in light-use efficiency,” *Advances in ecological research*, vol. 28, pp. 57–92, 1999.
- [6] F.A. Heinsch, M. Reeves, P. Votava, S. Kang, C. Milesi, M. Zhao, J. Glassy, W.M. Jolly, R. Loehman, C.F. Bowker, et al., “GPP and NPP (MOD17A2/A3) Products NASA MODIS Land Algorithm,” *MOD17 User’s Guide*, pp. 1–57, 2003.

-
- [7] F. Veroustraete, H. Sabbe, and H. Eerens, “Estimation of carbon mass fluxes over Europe using the C-Fix model and Euroflux data,” *Remote Sensing of Environment*, vol. 83, no. 3, pp. 376–399, 2002.
- [8] F. Veroustraete, H. Sabbe, D.P. Rasse, and L. Bertels, “Carbon mass fluxes of forests in Belgium determined with low resolution optical sensors,” *International Journal of Remote Sensing*, vol. 25, no. 4, pp. 769–792, 2004.
- [9] F. Maselli, A. Barbati, M. Chiesi, G. Chirici, and P. Corona, “Use of remotely sensed and ancillary data for estimating forest gross primary productivity in Italy,” *Remote sensing of environment*, vol. 100, no. 4, pp. 563–575, 2006.
- [10] F. Maselli, D. Papale, N. Puletti, G. Chirici, and P. Corona, “Combining remote sensing and ancillary data to monitor the gross productivity of water-limited forest ecosystems,” *Remote Sensing of Environment*, vol. 113, no. 3, pp. 657–667, 2009.
- [11] M. Iqbal, *An introduction to solar radiation*, Academic Press, Orlando, FL, 1983.
- [12] J-L Roujean and F-M Breon, “Estimating PAR absorbed by vegetation from bidirectional reflectance measurements,” *Remote Sensing of Environment*, vol. 51, no. 3, pp. 375–384, 1995.
- [13] A. Moreno, F.J. García-Haro, B. Martínez, and M.A. Gilabert, “Noise reduction and gap filling of FAPAR time series using an adapted local regression filter,” *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing (under revision)*.
- [14] A. Pérez-Hoyos, F.J. García-Haro, and J. San-Miguel-Ayanz, “A methodology to generate a synergetic land-cover map by fusion of different land-cover products,” *International Journal of Applied Earth Observation and Geoinformation*, vol. 19, pp. 72–87, 2012.
- [15] IPCC, “IPCC Fourth Assessment Report: Climate Change 2007,” Tech. Rep., Intergovernmental Panel on Climate Change, 2007.
- [16] D.A. Clark, S. Brown, D.W. Kicklighter, J.Q. Chambers, J.R. Thomlinson, and J. Ni, “Measuring net primary production in forests: concepts and field methods,” *Ecological Applications*, vol. 11, no. 2, pp. 356–370, 2001.

-
- [17] I.C. Prentice, G.D. Farquhar, M.J.R. Fasham, M.L. Goulden, M. Heimann, H.S. Kheshi, L. Quere, R.J. Scholes, D.W.R. Wallace, D. Archer, et al., “The carbon cycle and atmospheric carbon dioxide,” 2001.
- [18] P. Smithson, K. Addison, and K. Atkinson, *Fundamentals of the physical environment*, Routledge, 2002.
- [19] F.S. Chapin III, P.A. Matson, and H.A. Mooney, *Principles of terrestrial ecosystem ecology*, Springer, 2011.
- [20] R.H. Waring and S.W. Running, *Forest ecosystems: analysis at multiple scales*, Elsevier, 2010.
- [21] G.M. Lovett, J.J. Cole, and M.L. Pace, “Is net ecosystem production equal to ecosystem carbon accumulation?,” *Ecosystems*, vol. 9, no. 1, pp. 152–155, 2006.
- [22] F.S. Chapin III, G.M. Woodwell, J.T. Randerson, E.B. Rastetter, G.M. Lovett, D.D. Baldocchi, D.A. Clark, M.E. Harmon, D.S. Schimel, R. Valentini, et al., “Reconciling carbon-cycle concepts, terminology, and methods,” *Ecosystems*, vol. 9, no. 7, pp. 1041–1050, 2006.
- [23] R.H. Waring, J.J. Landsberg, and M. Williams, “Net primary production of forests: a constant fraction of gross primary production?,” *Tree Physiology*, vol. 18, no. 2, pp. 129–134, 1998.
- [24] S. Luyssaert, M. Reichstein, E-D Schulze, I.A. Janssens, B.E. Law, D. Papale, D. Dragoni, M.L. Goulden, A. Granier, W.L. Kutsch, et al., “Toward a consistency cross-check of eddy covariance flux-based and biometric estimates of ecosystem carbon balance,” *Global Biogeochemical Cycles*, vol. 23, no. 3, pp. n/a–n/a, 2009.
- [25] J. Chen, K.D. Brosofske, A. Noormets, T.R. Crow, M.K. Bresee, J.M. Le Moine, E.S. Euskirchen, S.V. Mather, and D. Zheng, “A working framework for quantifying carbon sequestration in disturbed land mosaics,” *Environmental Management*, vol. 33, no. 1, pp. 210–221, 2004.

- [26] P.S. Curtis, P.J. Hanson, P. Bolstad, C. Barford, J.C. Randolph, H.P. Schmid, and K.B. Wilson, “Biometric and eddy-covariance based estimates of annual carbon storage in five eastern North American deciduous forests,” *Agricultural and Forest Meteorology*, vol. 113, no. 1, pp. 3–19, 2002.
- [27] G.G. Burba and D. Anderson, *A brief practical guide to eddy covariance flux measurements: principles and workflow examples for scientific and industrial applications*, Li-Cor Biosciences, 2010.
- [28] D.D. Baldocchi, B.B. Hincks, and T.P. Meyers, “Measuring biosphere-atmosphere exchanges of biologically related gases with micrometeorological methods,” *Ecology*, vol. 69, no. 5, pp. 1331–1340, 1988.
- [29] M. Reichstein, E. Falge, D. Baldocchi, D. Papale, M. Aubinet, P. Berbigier, C. Bernhofer, N. Buchmann, T. Gilmanov, A. Granier, et al., “On the separation of net ecosystem exchange into assimilation and ecosystem respiration: review and improved algorithm,” *Global Change Biology*, vol. 11, no. 9, pp. 1424–1439, 2005.
- [30] D. Papale, M. Reichstein, M. Aubinet, E. Canfora, C. Bernhofer, W. Kutsch, B. Longdoz, S. Rambal, R. Valentini, T. Vesala, et al., “Towards a standardized processing of Net Ecosystem Exchange measured with eddy covariance technique: algorithms and uncertainty estimation,” *Biogeosciences*, vol. 3, no. 4, pp. 571–583, 2006.
- [31] I. McCallum, W. Wagner, C. Schmullius, A. Shvidenko, M. Obersteiner, S. Fritz, S. Nilsson, et al., “Satellite-based terrestrial production efficiency modeling,” *Carbon Balance and Management*, vol. 4, no. 1, pp. 1–14, 2009.
- [32] K.P. Tan, K.D. Kanniah, and A.P. Cracknell, “A review of remote sensing based productivity models and their suitability for studying oil palm productivity in tropical regions,” *Progress in Physical Geography*, vol. 36, no. 5, pp. 655–679, 2012.
- [33] C.S. Potter, J.T. Randerson, C.B. Field, P.A. Matson, P.M. Vitousek, H.A. Mooney, and S.A. Klooster, “Terrestrial ecosystem production: a process model based on global satellite and surface data,” *Global Biogeochemical Cycles*, vol. 7, no. 4, pp. 811–841, 1993.

- [34] S.D. Prince and S.N. Goward, "Global primary production: a remote sensing approach," *Journal of biogeography*, vol. 22, pp. 815–835, 1995.
- [35] X. Xiao, D. Hollinger, J. Aber, M. Goltz, E.A. Davidson, Q. Zhang, and B. Moore III, "Satellite-based modeling of gross primary production in an evergreen needleleaf forest," *Remote Sensing of Environment*, vol. 89, no. 4, pp. 519–534, 2004.
- [36] A. Ruimy, L. Kergoat, A. Bondeau, et al., "Comparing global models of terrestrial net primary productivity (NPP): Analysis of differences in light absorption and light-use efficiency," *Global Change Biology*, vol. 5, no. S1, pp. 56–64, 1999.
- [37] W. Yuan, S. Liu, G. Zhou, G. Zhou, L.L. Tieszen, D. Baldocchi, C. Bernhofer, H. Gholz, A.H. Goldstein, M.L. Goulden, et al., "Deriving a light use efficiency model from eddy covariance flux data for predicting daily gross primary production across biomes," *Agricultural and Forest Meteorology*, vol. 143, no. 3, pp. 189–207, 2007.
- [38] C. Wu, X. Han, J. Ni, Z. Niu, and W. Huang, "Estimation of gross primary production in wheat from in situ measurements," *International Journal of Applied Earth Observation and Geoinformation*, vol. 12, no. 3, pp. 183–189, 2010.
- [39] D.A. Sims, A.F. Rahman, V.D. Cordova, B.Z. El-Masri, D.D. Baldocchi, P.V. Bolstad, L.B. Flanagan, A.H. Goldstein, D.Y. Hollinger, L. Misson, et al., "A new model of gross primary productivity for North American ecosystems based solely on the enhanced vegetation index and land surface temperature from MODIS," *Remote Sensing of Environment*, vol. 112, no. 4, pp. 1633–1646, 2008.
- [40] N.C. Coops, R.H. Waring, and J.J. Landsberg, "Assessing forest productivity in Australia and New Zealand using a physiologically-based model driven with averaged monthly weather data and satellite-derived estimates of canopy photosynthetic capacity," *Forest Ecology and Management*, vol. 104, no. 1, pp. 113–127, 1998.
- [41] Y. Peng and A.A. Gitelson, "Remote estimation of gross primary productivity in soybean and maize based on total crop chlorophyll content," *Remote Sensing of Environment*, vol. 117, no. 15, pp. 440–448, 2012.

- [42] M. Meroni, M. Rossini, L. Guanter, L. Alonso, U. Rascher, R. Colombo, and J. Moreno, "Remote sensing of solar-induced chlorophyll fluorescence: Review of methods and applications," *Remote Sensing of Environment*, vol. 113, no. 10, pp. 2037–2051, 2009.
- [43] J.M. Chen, "Canopy architecture and remote sensing of the fraction of photosynthetically active radiation absorbed by boreal conifer forests," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 34, no. 6, pp. 1353–1368, 1996.
- [44] Y. Knyazikhin, J.V. Martonchik, R.B. Myneni, D.J. Diner, and S.W. Running, "Synergistic algorithm for estimating vegetation canopy leaf area index and fraction of absorbed photosynthetically active radiation from MODIS and MISR data," *Journal of Geophysical Research: Atmospheres (1984–2012)*, vol. 103, no. D24, pp. 32257–32275, 1998.
- [45] N. Gobron, F. Mélin, B. Pinty, M.M. Verstraete, J-L Widlowski, and G. Bucini, "A global vegetation index for SeaWiFS: Design and applications," in *Remote Sensing and Climate Modeling: Synergies and Limitations*, Martin Beniston and Michel M. Verstraete, Eds., vol. 7 of *Advances in Global Change Research*, pp. 5–21. Springer Netherlands, 2001.
- [46] N. Gobron, B. Pinty, O. Aussedat, J.M. Chen, W.B. Cohen, R. Fensholt, V. Gond, K.F. Huemmrich, T. Lavergne, F. Mélin, et al., "Evaluation of fraction of absorbed photosynthetically active radiation products for different canopy radiation transfer regimes: Methodology and results using Joint Research Center products derived from SeaWiFS against ground-based estimations," *Journal of Geophysical Research: Atmospheres (1984–2012)*, vol. 111, no. D13, 2006.
- [47] F.J. García-Haro, F. Camacho, and J. Meliá, "Vegetation Parameters Validation Report (VEGA VR)," Tech. Rep., SAF/LAND/UV/VR VEGA/2.1, January 2008, 91 pp. Available online at <http://landsaf.meteo.pt>, 2008.
- [48] United States. National Bureau of Standards and Fred Edwin Nicodemus, *Geometrical considerations and nomenclature for reflectance*, vol. 160, US Department of Commerce, National Bureau of Standards Washington, D. C, 1977.

- [49] B. Martínez, F. Camacho, A. Verger, F.J. García-Haro, and M.A. Gilabert, “Intercomparison and quality assessment of MERIS, MODIS and SEVIRI FAPAR products over the Iberian Peninsula,” *International Journal of Applied Earth Observation and Geoinformation*, vol. 21, pp. 463–476, 2013.
- [50] S.N. Goward, B. Markham, D.G. Dye, W. Dulaney, and J. Yang, “Normalized difference vegetation index measurements from the Advanced Very High Resolution Radiometer,” *Remote sensing of environment*, vol. 35, no. 2, pp. 257–277, 1991.
- [51] J.O. Pettorelli, N. and Vik, A. Mysterud, J-M Gaillard, C.J. Tucker, and N.C. Stenseth, “Using the satellite-derived NDVI to assess ecological responses to environmental change,” *Trends in Ecology & Evolution*, vol. 20, no. 9, pp. 503–510, 2005.
- [52] H. Kobayashi and D.G. Dye, “Atmospheric conditions for monitoring the long-term vegetation dynamics in the Amazon using normalized difference vegetation index,” *Remote Sensing of Environment*, vol. 97, no. 4, pp. 519–525, 2005.
- [53] B.N. Holben, “Characteristics of maximum-value composite images from temporal AVHRR data,” *International Journal of Remote Sensing*, vol. 7, no. 11, pp. 1417–1434, 1986.
- [54] P. Jonsson and L. Eklundh, “Seasonality extraction by function fitting to time-series of satellite sensor data,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 40, no. 8, pp. 1824–1832, 2002.
- [55] W.S. Cleveland and S.J. Devlin, “Locally weighted regression: an approach to regression analysis by local fitting,” *Journal of the American Statistical Association*, vol. 83, no. 403, pp. 596–610, 1988.
- [56] J. Chen, P. Jönsson, M. Tamura, Z. Gu, B. Matsushita, and L. Eklundh, “A simple method for reconstructing a high-quality NDVI time-series data set based on the Savitzky–Golay filter,” *Remote sensing of Environment*, vol. 91, no. 3, pp. 332–344, 2004.
- [57] P. Jönsson and L. Eklundh, “TIMESAT-A program for analyzing time-series of satellite sensor data,” *Computers & Geosciences*, vol. 30, no. 8, pp. 833–845, 2004.

- [58] A. Savitzky and M.J.E. Golay, "Smoothing and differentiation of data by simplified least squares procedures.," *Analytical chemistry*, vol. 36, no. 8, pp. 1627–1639, 1964.
- [59] D.A. Sampson and H.L. Allen, "Light attenuation in a 14-year-old loblolly pine stand as influenced by fertilization and irrigation," *Trees-Structure and Function*, vol. 13, no. 2, pp. 80–87, 1998.
- [60] J. Ross and M. Sulev, "Sources of errors in measurements of PAR," *Agricultural and forest meteorology*, vol. 100, no. 2-3, pp. 103–125, 2000.
- [61] V. Badescu, *Modeling Solar Radiation at the Earth's Surface: Recent Advances*, Springer, 2008.
- [62] X. Liu, X. Mei, Y. Li, Y. Zhang, Q. Wang, J.R. Jensen, and J.R. Porter, "Calibration of the Ångström-Prescott coefficients (a, b) under different time scales and their impacts in estimating global solar radiation in the Yellow River basin," *Agricultural and Forest Meteorology*, vol. 149, no. 3-4, pp. 697–710, 2009.
- [63] J. Almorox and C. Hontoria, "Global solar radiation estimation using sunshine duration in Spain," *Energy Conversion and Management*, vol. 45, no. 9-10, pp. 1529–1535, 2004.
- [64] K.L. Bristow and G.S. Campbell, "On the relationship between incoming solar radiation and daily maximum and minimum temperature," *Agricultural and Forest Meteorology*, vol. 31, no. 2, pp. 159–166, 1984.
- [65] S.S. Haykin, *Neural networks and learning machines*, vol. 10, Prentice Hall Upper Saddle River, NJ, 2009.
- [66] E. Alpaydin, *Introduction to Machine Learning (Adaptive Computation and Machine Learning)*, The MIT Press, 2004.
- [67] P.J. Durrant, *winGamma TM: a non-linear data analysis and modelling tool with applications to flood prediction*, Ph.D. thesis, Department of Computer Science, Cardiff University, Wales, UK, 2001.
- [68] H. Alsamamra, J.A. Ruiz-Arias, D. Pozo-Vázquez, and J. Tovar-Pescador, "A comparative study of ordinary and residual kriging techniques for mapping global solar radiation over southern Spain," *Agricultural and Forest Meteorology*, vol. 149, no. 8, pp. 1343–1357, 2009.

- [69] P.E. Van Laake, , and G.A. Sanchez-Azofeifa, “Simplified atmospheric radiative transfer modelling for estimating incident PAR using MODIS atmosphere products,” *Remote Sensing of Environment*, vol. 91, no. 1, pp. 98–113, 2004.
- [70] D. Cano, J-M Monget, M. Albuisson, H. Guillard, N. Regas, and L. Wald, “A method for the determination of the global solar radiation from meteorological satellite data,” *Solar Energy*, vol. 37, no. 1, pp. 31–39, 1986.
- [71] C. Rigollier, M. Lefèvre, and L. Wald, “The method Heliosat-2 for deriving shortwave solar radiation from satellite images,” *Solar Energy*, vol. 77, no. 2, pp. 159–169, 2004.
- [72] R. Meyer, C. Hoyer, C. Schillings, F. Trieb, E. Diedrich, and M. Schroedter, “SOLEMI: A New satellite-based service for high-resolution and precision solar radiation data for Europe, Africa and Asia,” in *ISES Solar World Congress 2003*, 2003.
- [73] R.E. Bird and R.L. Hulstrom, “Simplified clear sky model for direct and diffuse insolation on horizontal surfaces,” Tech. Rep., Solar Energy Research Inst., Golden, CO (USA), 1981.
- [74] B. Geiger, C. Meurey, D. Lajas, L. Franchistéguy, D. Carrer, and J-L Roujean, “Near real-time provision of downwelling shortwave radiation estimates derived from satellite observations,” *Meteorological Applications*, vol. 15, no. 3, pp. 411–420, 2008.
- [75] M.G.R. Cannell and J.H.M. Thornley, “Modelling the components of plant respiration: some guiding principles,” *Annals of Botany*, vol. 85, no. 1, pp. 45–54, 2000.
- [76] B.D. Cook, P.V. Bolstad, J.G. Martin, F.A. Heinsch, K.J. Davis, W. Wang, A.R. Desai, and R.M. Teclaw, “Using light-use and production efficiency models to predict photosynthesis and net carbon exchange during forest canopy disturbance,” *Ecosystems*, vol. 11, no. 1, pp. 26–44, 2008.
- [77] P. Ciais, M. Reichstein, N. Viovy, A. Granier, J. Ogée, V. Allard, M. Aubinet, N. Buchmann, C. Bernhofer, A. Carrara, et al., “Europe-wide reduction in primary productivity caused by the heat and drought in 2003,” *Nature*, vol. 437, no. 7058, pp. 529–533, 2005.

- [78] H. Bolle, M. Eckardt, D. Koslowsky, F. Maselli, J.M. Miralles, M. Menenti, F.S. Olesen, L. Petkov, and I. Rasool, *Mediterranean land-surface processes assessed from space*, Springer, 2006.
- [79] M.J. Aitkenhead and I.H. Aalders, “Automating land cover mapping of Scotland using expert system and knowledge integration methods,” *Remote Sensing of Environment*, vol. 115, no. 5, pp. 1285–1295, 2011.
- [80] T.R. Loveland, B.C. Reed, J.F. Brown, D.O. Ohlen, Z. Zhu, L. Yang, and J.W. Merchant, “Development of a global land cover characteristics database and IGBP DISCover from 1 km AVHRR data,” *International Journal of Remote Sensing*, vol. 21, no. 6-7, pp. 1303–1330, 2000.
- [81] M.C. Hansen, R.S. DeFries, J.R.G. Townshend, and R. Sohlberg, “Global land cover classification at 1 km spatial resolution using a classification tree approach,” *International Journal of Remote Sensing*, vol. 21, no. 6-7, pp. 1331–1364, 2000.
- [82] E. Bartholomé and A.S. Belward, “GLC2000: a new approach to global land cover mapping from Earth observation data,” *International Journal of Remote Sensing*, vol. 26, no. 9, pp. 1959–1977, 2005.
- [83] M.A. Friedl, D. Sulla-Menashe, B. Tan, A. Schneider, N. Ramankutty, A. Sibley, and X. Huang, “MODIS Collection 5 global land cover: Algorithm refinements and characterization of new datasets,” *Remote Sensing of Environment*, vol. 114, no. 1, pp. 168–182, 2010.
- [84] J-L Kaptué Tchuenté, A.T. and Roujean and S.M. De Jong, “Comparison and relative quality assessment of the GLC2000, GLOBCOVER, MODIS and ECOCLIMAP land cover data sets at the African continental scale,” *International Journal of Applied Earth Observation and Geoinformation*, vol. 13, no. 2, pp. 207–219, 2011.
- [85] M. Bossard, J. Feranec, and J. Otahel, “CORINE land cover technical guide: Addendum 2000,” 2000.
- [86] P. Bicheron, P. Defourny, C. Brockmann, L. Schouten, C. Vancutsem, M. Huc, S. Bontemps, M. Leroy, F. Achard, M. Herold, et al., “GLOBCOVER: products description and validation report,” *ME-DIAS France, Toulouse*, 2008.

- [87] A. Pérez-Hoyos, *Desarrollo de metodologías para la comparación, validación e integración de información procedente de productos de cobertura del suelo*, Ph.D. thesis, PhD Dissertation, University of Valencia, Spain, 2011.
- [88] M.E. Jensen and H.R. Haise, “Estimating evapotranspiration from solar radiation,” *Proceedings of the American Society of Civil Engineers, Journal of the Irrigation and Drainage Division*, vol. 89, pp. 15–41, 1963.
- [89] W. Bilger and O. Björkman, “Role of the xanthophyll cycle in photoprotection elucidated by measurements of light-induced absorbance changes, fluorescence and photosynthesis in leaves of *hedera canariensis*,” *Photosynthesis Research*, vol. 25, no. 3, pp. 173–185, 1990.
- [90] J.A. Gamon, J. Peñuelas, and C.B. Field, “A narrow-waveband spectral index that tracks diurnal changes in photosynthetic efficiency,” *Remote sensing of environment*, vol. 41, no. 1, pp. 35–44, 1992.
- [91] J. Peñuelas, J. Llusia, J. Pinol, and I. Filella, “Photochemical reflectance index and leaf photosynthetic radiation-use-efficiency assessment in Mediterranean trees,” *International Journal of Remote Sensing*, vol. 18, no. 13, pp. 2863–2868, 1997.
- [92] G.P. Asner, R.E. Martin, K.M. Carlson, U. Rascher, and P.M. Vitousek, “Vegetation–climate interactions among native and invasive species in hawaiian rainforest,” *Ecosystems*, vol. 9, no. 7, pp. 1106–1117, 2006.
- [93] M. F Garbulsky, J. Peñuelas, J.A. Gamon, Y. Inoue, and I. Filella, “The photochemical reflectance index (PRI) and the remote sensing of leaf, canopy and ecosystem radiation use efficiencies: a review and meta-analysis,” *Remote Sensing of Environment*, vol. 115, no. 2, pp. 281–297, 2011.
- [94] G.G. Drolet, E.M. Middleton, K.F. Huemmrich, F.G. Hall, B.D. Amiro, A.G. Barr, T.A. Black, J.H. McCaughey, and H.A. Margolis, “Regional mapping of gross light-use efficiency using MODIS spectral indices,” *Remote Sensing of Environment*, vol. 112, no. 6, pp. 3064–3078, 2008.

- [95] J-M Gre, P. Ceccato, B. Pinty, S. Tarantola, and N. Gobron, “Designing a spectral index to estimate vegetation water content from remote sensing data: Part 1 Theoretical approach,” *Remote Sensing of Environment*, vol. 82, pp. 188–197, 2002.
- [96] P.J Zarco-Tejada, C. Rueda, and S.L Ustin, “Water content estimation in vegetation with MODIS reflectance data and model inversion methods,” *Remote Sensing of Environment*, vol. 85, no. 1, pp. 109–124, 2003.
- [97] F.M. Danson and P. Bowyer, “Estimating live fuel moisture content from remotely sensed reflectance,” *Remote Sensing of Environment*, vol. 92, no. 3, pp. 309–321, 2004.
- [98] B-C Gao, “NDWI a normalized difference water index for remote sensing of vegetation liquid water from space,” *Remote sensing of environment*, vol. 58, no. 3, pp. 257–266, 1996.
- [99] R. Fensholt and I. Sandholt, “Derivation of a shortwave infrared water stress index from MODIS near-and shortwave infrared data in a semiarid environment,” *Remote Sensing of Environment*, vol. 87, no. 1, pp. 111–121, 2003.
- [100] M.A. Rubio, D. Riaño, Y.B. Cheng, and S.L. Ustin, “Estimation of canopy water content from MODIS using artificial neural networks trained with radiative transfer models,” *6th EMS/6th ECAC*, 2006.
- [101] B. Martínez and M.A. Gilabert, “Vegetation dynamics from NDVI time series analysis using the wavelet transform,” *Remote Sensing of Environment*, vol. 113, no. 9, pp. 1823–1842, 2009.
- [102] C. Justice, A. Belward, J. Morisette, P. Lewis, J. Privette, and F. Baret, “Developments in the ‘validation’ of satellite sensor products for the study of the land surface,” *International Journal of Remote Sensing*, vol. 21, no. 17, pp. 3383–3390, 2000.
- [103] M. Göckede, T. Foken, M. Aubinet, M. Aurela, J. Banza, C. Bernhofer, J-M Bonnefond, Y. Brunet, A. Carrara, R. Clement, et al., “Quality control of CarboEurope flux data—Part 1: Coupling footprint analyses with flux data quality assessment to evaluate sites in forest ecosystems,” *Biogeosciences*, vol. 5, no. 2, pp. 433–450, 2008.
- [104] AEMet, “Calendario Meteorológico,” Tech. Rep., Ministerio de Medio Ambiente y Medio Rural y Marino, 2009.

- [105] B.R. Reverter, E.P. Sánchez-Cañete, V. Resco, P. Serrano-Ortiz, C. Oyonarte, and A.S. Kowalski, “Analyzing the major drivers of NEE in a Mediterranean alpine shrubland,” *Biogeosciences*, vol. 7, no. 9, pp. 2601–2611, 2010.
- [106] A. Rey, E. Pegoraro, C. Oyonarte, A. Were, P. Escribano, and J. Raimundo, “Impact of land degradation on soil respiration in a steppe (*Stipa tenacissima*) semi-arid ecosystem in the SE of Spain,” *Soil Biology and Biochemistry*, vol. 43, no. 2, pp. 393–403, 2011.
- [107] AEMet, “Calendario Meteorológico,” Tech. Rep., Ministerio de Medio Ambiente y Medio Rural y Marino, 2012.
- [108] D.P. Turner, W.D. Ritts, W.B. Cohen, S.T. Gower, S.W. Running, M. Zhao, M.H. Costa, A.A. Kirschbaum, J.M. Ham, S.R. Saleska, et al., “Evaluation of MODIS NPP and GPP products across multiple biomes,” *Remote Sensing of Environment*, vol. 102, no. 3, pp. 282–292, 2006.
- [109] J. Xiao, Q. Zhuang, B.E. Law, J. Chen, D.D. Baldocchi, D.R. Cook, R. Oren, A.D. Richardson, S. Wharton, S. Ma, et al., “A continuous measure of gross primary production for the conterminous United States derived from MODIS and AmeriFlux data,” *Remote Sensing of Environment*, vol. 114, no. 3, pp. 576–591, 2010.
- [110] B. B. Smets, H. Eerens, T. Jacobs, and A. Royer, “Product User Manual, Dry Matter Productivity (DMP),” Tech. Rep., GIO GLOBAL LAND COMPONENT-LOT I “OPERATION OF THE GLOBAL LAND COMPONENT”, 29 pp. Available on-line at <http://land.copernicus.eu/>, 2013.
- [111] G.L. Ajtay, P. Ketner, and P. Duvigneaud, “Terrestrial primary production and phytomass,” *The global carbon cycle*, vol. 13, pp. 129–182, 1979.

Apéndice **A**

Publication I: Mapping daily global solar irradiation over Spain: A comparative study of selected approaches



Mapping daily global solar irradiation over Spain: A comparative study of selected approaches

A. Moreno, M.A. Gilabert^{*}, B. Martínez

Departament de Física de la Terra i Termodinàmica, Universitat de València, Dr. Moliner, 50. 46100-Burjassot, Valencia, Spain

Received 17 December 2010; received in revised form 9 May 2011; accepted 25 May 2011
Available online 22 June 2011

Communicated by: Associate Editor Christian Gueymard

Abstract

Three methods to estimate the daily global solar irradiation are compared: the Bristow–Campbell (BC), Artificial Neural Network (ANN) and Kernel Ridge Regression (KRR). BC is an empirical approach based on air maximum and minimum temperature. ANN and KRR are non-linear approaches that use temperature and precipitation data (which have been selected as the best combination of input data from a gamma test). The experimental dataset includes 4 years (2005–2008) of daily irradiation collected at 40 stations and temperature and precipitation data collected at 400 stations over Spain. Results show that the ANN method produces the best global solar irradiation estimates, with a mean absolute error $2.33 \text{ MJ m}^{-2} \text{ day}^{-1}$. Daily maps of solar irradiation over Spain at 1-km spatial resolution are produced by applying the ANN method to temperature and precipitation maps generated from ordinary kriging. © 2011 Elsevier Ltd. All rights reserved.

Keywords: Daily global solar irradiation maps; Bristow–Campbell; Artificial Neural Network; Kernel Ridge Regression

1. Introduction

The global solar irradiation is an essential variable with respect to many fields, including agriculture and renewable energy applications. In particular, the energy of solar radiation fixed in chemical form by green plants through photosynthesis is, with very rare exceptions, the primary energy source for living organisms (Smithson et al., 2002). The rate of total organic material fixed by photosynthesis is called the gross primary production. The fraction that remains in the plant after respiration losses is termed net primary production. The increase in biomass with time is a measure of net ecosystem production. That is, changes in biomass from year to year indicate the amount of energy or carbon fixed by photosynthesis and incorporated into an ecosystem. The procedures proposed to estimate vegetation production from remotely sensed data (Maselli et al., 2006,

2009; Veroustraete et al., 2002) require spatially continuous solar radiation data (maps).

Methodologies based on spatial interpolation methods, such as kriging (Alsamamra et al., 2009), allow us to obtain spatially continuous databases from isolated stations measurements. However, despite the large spectrum of applications demanding solar radiation data, direct measurements are not widely available due to the cost, maintenance and calibration requirements of the measuring equipment. In fact, the number of meteorological stations recording solar radiation is much lower than those recording temperature, precipitation, sunshine duration and relative humidity data. This limited availability of radiation data urges the development of computational procedures to estimate solar radiation from other available meteorological data (Mellit et al., 2005; Robaa, 2009; Tymvios et al., 2005), mainly when the final objective of the work is to generate solar radiation maps. In this case, the interpolation techniques allow us to obtain those spatially continuous databases based on spatially interpolation methods (such as kriging)

^{*} Corresponding author. Tel.: +34 963 543 118; fax: +34 963 543 385.
E-mail address: m.amparo.gilabert@uv.es (M.A. Gilabert).

and the accuracy of the estimations is highly influenced by the number of isolated-stations measurements (Huges and Lettenmaier, 1981). When relief is relevant, these procedures can be refined by means of the inclusion of additional topographic information (Pons and Ninyerola, 2008).

The global solar irradiation H arriving to the Earth surface is a function of the extraterrestrial value H_0 and the attenuation caused by the atmospheric constituents (represented by their total transmittance $\tau \propto H/H_0$). The daily extraterrestrial solar irradiation H_0 can be evaluated assuming a horizontal surface (Iqbal, 1983), as

$$H_0 = \frac{24}{\pi} I_{SC} f \left[\frac{\pi}{180} \varpi_s (\sin \delta \sin \varphi) + (\cos \delta \cos \varphi \cos \omega_s) \right] \quad (1)$$

where I_{SC} is the solar constant, f is defined as the square of the ratio between the mean Sun–Earth distance and the Sun–Earth distance for a given day, ϖ_s is the sunrise hour angle (in degrees), δ is the sun declination (angular position of the sun at solar noon respect to the equator plane, north positive, in degrees), and φ is the latitude of the site.

The atmospheric transmittance τ can be evaluated from both physical and empirical models. The first ones explicitly account for the interaction radiation-matter in the atmosphere, including absorption and scattering processes. These models require numerous data inputs (some of them not easily available) and are highly influenced by the input accuracy (Badescu, 2008). In contrast, empirical models estimate the atmospheric transmittance from available meteorological and geographical data such as sunshine hours, air temperature, latitude, precipitation, relative humidity, and cloudiness. Since these latter models do not explicitly explore the physics underlying the radiative transfer process, their main drawback is that they are site-specific and use empirical coefficients that need to be frequently updated and optimized.

Daily global solar irradiation can be determined by sunshine duration. Most solar irradiance models (Almorox et al., 2005; Liu et al., 2009a) use the Ångström–Prescott equation (Ångström, 1924; Prescott, 1940) to describe the extinction of radiation due to cloud cover. This is a linear relationship between the ratio of average daily global irradiation to the corresponding value on a completely clear day and the ratio of average daily sunshine duration to the maximum possible sunshine duration. Customary derivation leads to relative sunshine duration as a natural parameter in this type of correlation. Sunshine duration can be easily measured, and reliable data are widely available. Ångström empirical coefficients are time and space dependent and require site calibration (Almorox and Hontoria, 2004; Bakirci, 2009; Liu et al., 2009a; Menges et al., 2006).

The Ångström–Prescott equation has been modified to incorporate other surface meteorological parameters such as the fractional cloud amount and the daily extreme temperatures (Paulescu, 2008). Bristow and Campbell (1984) established an empirical equation for daily global solar irradiation at the Earth surface using daily air temperature amplitude. Other empirical procedures to estimate

atmospheric transmittance and solar radiation via advanced non-linear relations may include additional meteorological variables. This is the case of the Artificial Neural Network and the Kernel Ridge Regression procedures described in Section 2, which present a number of advantages over the classical models (Alpaydin, 2004; Bishop, 1996; Haykin, 1999). An interesting feature of these procedures is the fact that the non-linear relations among the variables are not postulated *a priori* but derived from the data.

The purpose of the present research is twofold. First, three methods to estimate the daily global solar irradiation on a horizontal surface are developed and compared: the Bristow–Campbell (BC), Artificial Neural Network (ANN) and Kernel Ridge Regression (KRR). A gamma test is applied to select the best combination of input data for the ANN and KRR methods. The conclusion from this test is that incorporation of the precipitation data leads to a significant improvement of the solar irradiation estimation. Second, maps of solar irradiation over Spain at 1-km spatial resolution are produced applying the best methodology (namely, ANN) to temperature and precipitation maps generated from ordinary kriging. The experimental dataset includes 4 years (2005–2008) of daily irradiation collected at about 40 stations and daily temperature and precipitation data collected at about 400 stations. An accuracy test of the global solar irradiation estimation over Spain is rather demanding because of its wide range of topographic and climatic characteristics, and therefore our conclusions are expected to be relevant to other geographical regions.

2. Material and methods

2.1. Study area

Spain is located in southwestern Europe on the Iberian Peninsula (Fig. 1). Spain has remarkable landscape diversity due to its relief, climate, geological features and biogeographical location (transition between the Eurosiberian and Mediterranean regions) (Alcaraz et al., 2006; Vicente-Serrano et al., 2008). Elevation ranges from sea level to 3479 m (Mulhacén in the Sierra Nevada, at the south). The study area has a climate with hot and dry summers in the interior and wet and cool summers along the coastlines, especially along the Atlantic coast. The precipitation regime is dominated in winter by the advection of Atlantic air masses, while in summer it is determined by convective processes, which also depend on land surface conditions. In the Mediterranean coast the climate has dry and hot conditions in summer resulting from subtropical anticyclones. In this coast, maximum rainfall occurs in autumn and spring and annual rainfall ranges between 350 and 500 mm (Immerzeel et al., 2009).

2.2. Data

The experimental dataset used includes 4 years (2005–2008) of daily global solar irradiation (H), maximum and

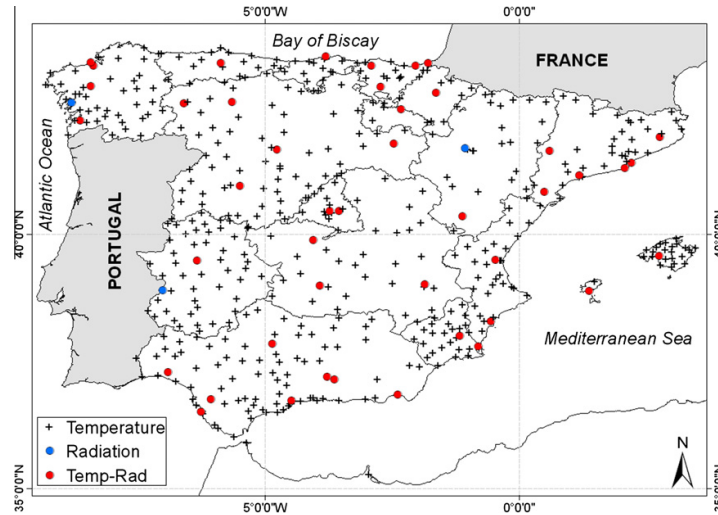


Fig. 1. Study area with the location of the meteorological stations.

minimum temperatures (T_{\max} and T_{\min}), and precipitation (P) measured at the surface (supplied by the Spanish Meteorological Agency AEMet). Around 40 stations record radiation data, whereas ca. 400 stations record temperature and precipitation (see Fig. 1). As mentioned above, this is the main reason to select procedures based on temperature and/or precipitation data to map global solar irradiation. The 4 years considered in the study present higher temperature than the mean that characterize the period 1970–2000, being 2005 and 2006 extremely hot. In relation to precipitation, 2005 was the driest year and 2008 the wettest one.

2.3. Estimation of the atmospheric transmittance

2.3.1. Method 1: Bristow-Campbell (BC)

The daily air temperature amplitude $\Delta T = T_{\max} - T_{\min}$ can be used as an indicator of overcast conditions that impact on the transmittance of solar radiation through the atmosphere. The models that estimate solar irradiation via air temperature data (Paulescu, 2008) are rather popular because these measurements are routinely recorded at weather stations (while radiation measurements are scarce). Most of them are based on the original Bristow and Campbell (1984) empirical relation between the daily global solar irradiation at the Earth surface H and ΔT

$$\frac{H}{H_0} = a[1 - \exp(-b(\Delta T)^c)] \quad (2)$$

or its alternative (Weiss et al., 2001)

$$\frac{H}{H_0} = a[1 - \exp(-bD^c)] \quad (3)$$

where $D(j) = T_{\max}(j) - 0.5[T_{\min}(j) + T_{\min}(j+1)]$ is a temperature amplitude evaluated considering the average of

two successive days ($j, j+1$) as the minimum temperature. The values of empirical coefficients a , b and c are sensitive to the location because ΔT is influenced by local weather conditions (Liu et al., 2009b; Weiss et al., 2001) and their determination requires a global solar radiation database (Ball et al., 2004).

2.3.2. Method 2: Artificial Neural Network (ANN)

A traditional possibility to extend the linear regression framework is the use of Artificial Neural Networks (ANNs). These are connected structures of neurons organized in layers, where a neuron is just a linear regression followed in general by a non-linear function (Haykin, 1999). ANN procedures allow us to estimate atmospheric transmittance and global solar radiation at the surface using advanced non-linear modeling that can incorporate additional meteorological data such as precipitation. These procedures are very versatile and have already been applied for the prediction of tropospheric ozone concentration (Gómez et al., 2006), rainfall-runoff processes, etc. Several studies have shown that ANN procedures are most advantageous to estimate incoming solar radiation data (López et al., 2005; Mellit et al., 2005; Rehman and Mohandes, 2008; Tymvios et al., 2005).

In this work, the classical Multilayer Perceptron (MLP) has been used. MLP consists of a set of interconnected elementary processing units (neurons). Each connection to a neuron has an adjustable weight factor associated with it. The transformation carried out by the output layer neuron is linear whereas every neuron of the hidden layers carries out a non-linear transformation of its inputs x_i ($i = 1, \dots, m$) defined by the following expression:

$$y = \tanh\left(\sum_{i=1}^m w_i x_i + u\right) \quad (4)$$

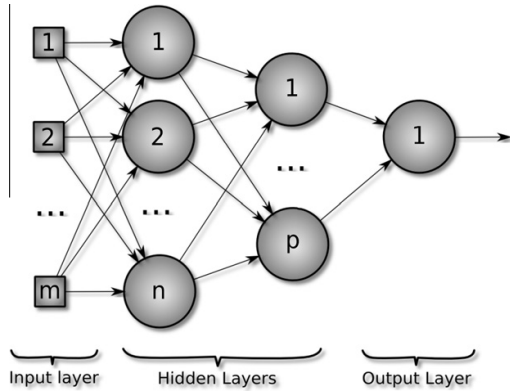


Fig. 2. Multilayer perceptron (MLP) structure.

where w_i are the synaptic weights of the connections and u is the bias coefficient. The tanh is a common choice because it is non-linear and differentiable.

The inputs are first transformed (i.e., standardized) to have zero mean and variance equal to one. The neurons are organized into layers with no feedback or lateral connections. It consists of a lowermost input layer, some intermediate or hidden layers, and an output layer at the top. The input signal propagates through the network in a forward direction, on a layer-by-layer basis (see Fig. 2). It has been demonstrated that a MLP with two hidden-layers can model any relation between two sets of variables (Haykin, 1999). This ability of generating the correct input/output relation is achieved after a training process of the network in which the synaptic weights w_i and the bias coefficient are determined. We have used the iterative Levenberg–Marquardt algorithm (Lourakis, 2005; Madsen et al., 2004), which shows a better performance than other classical algorithms (e.g., backpropagation, Kermani et al., 2005). This algorithm requires testing different initializations because it is based on local search of the minimum of a function and the problem of falling into a local minimum has to be avoided. When the current solution is far from the minimum, the algorithm behaves like a steepest descent method (that is, minimization along the gradient direction), but guaranteed to converge. When the current solution is close to the minimum, it becomes a Gauss–Newton method (it uses a quadratic model to speed up the process of finding the minimum).

We have trained 150 ANNs with different architectures (in particular, with different numbers of neurons in the two hidden layers) using randomly chosen 60% of the available daily temperature amplitude and daily precipitation data as neuron inputs (training data set). Another 25% of the data (validation data set) are used to stop the training in the case of generalization loss of the ANNs. The performance of these architectures has been tested by comparing their predictions to the remaining 15% of the meteorological data (test data set). The “network committee” consisting

on the six better ANNs (lowest error over the validation data set) finally produces the output by averaging their estimates. This averaged result gives better and more stable estimates than using a unique network (Haykin, 1999).

2.3.3. Method 3: Kernel Ridge Regression (KRR)

Our experimental dataset (also known as training data) consists of N sets of values $(x_{1,i}, \dots, x_{m,i}, y_i)$ where $y_i = H_i$ and $\vec{x}_i \equiv [x_{1,i}, \dots, x_{m,i}]^T$ are the measured solar irradiation and meteorological data, respectively. The inputs (\vec{x}_i, y_i) are first transformed (i.e., standardized) to have zero mean and variance equal one. Although the relation between meteorological data and solar irradiation is nonlinear, a linear relation can be used in a higher dimensional space between solar irradiation and a new set of variables $\vec{\Phi}_i \equiv [\phi_1(\vec{x}_i), \dots, \phi_D(\vec{x}_i)]^T$ where the $\phi_k(\vec{x}_i)$ ($1 \leq k \leq D$) are real functions of the meteorological data (Guyon, 2005). The number D of functions used (i.e. the dimensions of the Hilbert space) is higher than the number m of direct meteorological data and can be even higher than the number of samples N . The linear regression equation is then

$$y = \sum_{k=1}^D w_k \phi_k(\vec{x}) + \rho \quad (5)$$

where w_k ($1 \leq k \leq D$) are the regression coefficients (or weights) and ρ is a Gaussian noise (i.e. a random variable with zero mean and finite variance).

Since the variables ϕ_k can be correlated, ridge regression must be used instead of ordinary least squares minimization. That is, the cost function to be considered is

$$C = \sum_{i=1}^N \left(y_i - \sum_{k=1}^D w_k \phi_k(\vec{x}_i) \right)^2 + \lambda \sum_{k=1}^D w_k^2 \\ = \sum_{i=1}^N \left(y_i - \vec{w}^T \cdot \vec{\Phi}_i \right)^2 + \lambda \|\vec{w}\|^2 \quad (6)$$

where the second term in the r.h.s of Eq. (6) is a regularization term and it remains to be determined how to choose the regularization parameter $\lambda > 0$. Minimization of the cost function allows us to determine the vector $\vec{w} \equiv [w_1, \dots, w_D]^T$ of regression coefficients as

$$\vec{w} = (\lambda \vec{I}_D + \vec{\Phi} \cdot \vec{\Phi}^T)^{-1} \cdot \vec{\Phi} \cdot \vec{Y} \quad (7)$$

where $\vec{Y} \equiv [y_1, \dots, y_N]^T$, \vec{I}_D is the $(D \times D)$ identity matrix, the component ji of the $(D \times N)$ $\vec{\Phi}$ is $\phi_j(\vec{x}_i)$; i.e., the i th column of matrix $\vec{\Phi}$, and the i th row of the $(N \times D)$ matrix $\vec{\Phi}^T$, is $\vec{\Phi}_i \equiv [\phi_1(\vec{x}_i), \dots, \phi_D(\vec{x}_i)]^T$.

Making use of the identity $(\lambda \vec{I}_D + \vec{\Phi} \cdot \vec{\Phi}^T)^{-1} \cdot \vec{\Phi} = \vec{\Phi} \cdot (\lambda \vec{I}_N + \vec{\Phi}^T \cdot \vec{\Phi})^{-1}$, a particular case of the Representer’s theorem, Eq. (7) can be transformed to $\vec{w} = \vec{\Phi} \cdot \vec{z}$ or

$$w_k = \sum_{i=1}^N \alpha_i \phi_k(\vec{x}_i) \quad (8)$$

where $\vec{\alpha} \equiv \vec{\beta} \cdot \vec{y}$ and $\vec{\beta}^{-1} \equiv \lambda \vec{I}_N + \vec{\Phi}^T \cdot \vec{\Phi}$. The components of the vector $\vec{\alpha}$ are known as dual variables and $\vec{K} \equiv \vec{\Phi}^T \cdot \vec{\Phi}$ is known as the $(N \times N)$ kernel matrix, whose component ij is $K_{ij} = (\vec{\Phi}^T \cdot \vec{\Phi})_{ij} = \sum_{k=1}^D \phi_k(\vec{x}_i) \phi_k(\vec{x}_j)$. The complexity of Eq. (8) lies in the evaluation of the inverse of the $(N \times N)$ matrix $\vec{\beta}^{-1}$; which is easier than the inverse of the $(D \times D)$ matrix in Eq. (7) when D is larger than N . The estimated (or predicted) value of the solar irradiation at a new position where the meteorological variables are $\vec{x} = [x_1, \dots, x_m]^T$ is

$$y(\vec{x}) = \sum_{k=1}^D w_k \phi_k(\vec{x}) = \sum_{i=1}^N \sum_{k=1}^D \alpha_i \phi_k(\vec{x}_i) \phi_k(\vec{x}) \quad (9)$$

or $y = \vec{w}^T \cdot \vec{\Phi}(\vec{x}) = \vec{\alpha}^T \cdot \vec{\Phi}^T \cdot \vec{\Phi}(\vec{x})$ in matrix form, where $\vec{\Phi}(\vec{x}) \equiv [\phi_1(\vec{x}), \dots, \phi_D(\vec{x})]^T$. Eq. (9) is the central equation of Kernel Ridge Regression (KRR), the dual representation of ridge regression (Shawe-Taylor and Cristianini, 2004).

A crucial point in the implementation of KRR is the choice of the kernel matrix. Although the theoretical background of the Kernel Ridge Regression explained above is based on choosing a set of functions $\phi_k(\vec{x})$ ($1 \leq k \leq D$) (as a basis of Hilbert space), it is customary (Keerthi and Lin, 2003) to use a Gaussian function (or Radial Basis Function) to define the components of the kernel matrix as follows

$$K_{ij} \equiv K(\vec{x}_i, \vec{x}_j) \equiv \exp\left(-\frac{\|\vec{x}_i - \vec{x}_j\|^2}{2\sigma^2}\right) \quad (10)$$

where $\sigma > 0$ is a kernel parameter to be optimized. Thus, the final expression for the estimation of solar irradiation is

$$y(\vec{x}) = \sum_{i=1}^N \alpha_i K(\vec{x}_i, \vec{x}) \quad (11)$$

where $\alpha_i = \sum_{j=1}^N \beta_{ij} y(\vec{x}_j)$, $(\vec{\beta}^{-1})_{ij} \equiv \lambda \delta_{ij} + K_{ij}$ and δ_{ij} is Kronecker's delta.

2.4. Ordinary kriging (OK)

Interpolation procedures allow us to obtain spatially continuous databases from stations measurements. The kriging techniques are a particular class of stochastic interpolation procedures (Alsamamra et al., 2009), which have become widely used (Webster and Oliver, 2001) because they have shown considerable advantages with respect to deterministic interpolation techniques, e.g., in rainfall (Buytaert et al., 2006; Tabios and Salas, 1985) and temperature estimation (Chuanyan et al., 2005; Jarvis and Stuart, 2001).

In the kriging techniques, a continuous random variable (RV) $Z(\vec{u})$ (i.e., spatial values of a random function) is estimated in the study area A at any unsampled location \vec{u} using the measurements at different locations, $Z(\vec{u}_\alpha)$,

$\alpha = 1, \dots, n$ (Goovaerts, 1997; Isaaks and Srivastava, 1989). The different kriging estimators are variants of the basic linear regression estimator $Z^*(\vec{u})$ defined as:

$$Z^*(\vec{u}) - m(\vec{u}) = \sum_{\alpha=1}^{n(\vec{u})} \lambda_\alpha(\vec{u}) [Z(\vec{u}_\alpha) - m(\vec{u}_\alpha)] \quad (12)$$

where $\lambda_\alpha(\vec{u})$ is the weight assigned to datum $Z(\vec{u}_\alpha)$ interpreted as a realization of the RV $Z(\vec{u}_\alpha)$ and the quantities $m(\vec{u})$ and $m(\vec{u}_\alpha)$, which are the expected values of the RVs $Z(\vec{u})$ and $Z(\vec{u}_\alpha)$. The n weights are obtained by solving a system of linear equations that must satisfy two conditions. First, weights must be such that the estimate given in Eq. (12) is unbiased, $E\{Z^*(\vec{u}) - Z(\vec{u})\} = 0$. Second, the weights must minimize the variance $\sigma^2(\vec{u}) = \text{var}\{Z^*(\vec{u}) - Z(\vec{u})\}$. The values of the weights depend mainly on the spatial configuration of the data. Observations at locations \vec{u}_α that are close to the unsampled location \vec{u} have larger weights $\lambda_\alpha(\vec{u})$ than those that are further away.

Different kriging estimators vary depending on the model adopted for $m(\vec{u})$ and $m(\vec{u}_\alpha)$ (usually referred as trend) and the possibility of including another RV on the estimates. The simple kriging (SK) considers the trend as known and constant throughout the study area A . The ordinary kriging (OK) accounts for local fluctuations of the mean by limiting the domain of stationarity to a small spatial window $W(\vec{u})$ (i.e., local neighborhood) within the study area A . The OK estimator is thus written as a linear combination of the $n(\vec{u})$ random variables $Z(\vec{u}_\alpha)$:

$$Z^*(\vec{u}) = \sum_{\alpha=1}^{n(\vec{u})} \lambda_\alpha^{\text{OK}}(\vec{u}) Z(\vec{u}_\alpha) \quad (13)$$

where the unknown local mean $m(\vec{u})$ is filtered from the linear estimator by forcing the kriging weights to sum to 1. The accuracy of the OK depends on the validity of the assumptions about stationarity of the mean and the variance as well as on the $z(\vec{u}_\alpha)$ measurements.

In order to determine the weights at unsampled locations, the variogram model is required (Woodcock et al., 1988). The variogram is the main tool used in Geostatistics in order to account for the degree of spatial correlation (i.e., spatial continuity) of a random variable as a function of the separation distance and direction. In practice, the effectiveness of kriging depends upon the appropriate selection of the model variogram parameters and how representative the observation points are of the phenomenon (Martínez et al., 2010).

In this paper, we have used the OK to calculate temperature and precipitation maps. As an example, Fig. 3 shows maximum temperature, minimum temperature and precipitation images calculated from January 1st, 2008 measurements.

The performance of OK is evaluated by the leave-one-out cross validation technique (Cawley and Talbot, 2003). The OK is applied at each location, $\alpha = 1, \dots, n$, using the actual measurement of the specific location as the validation data, and the remaining observations as

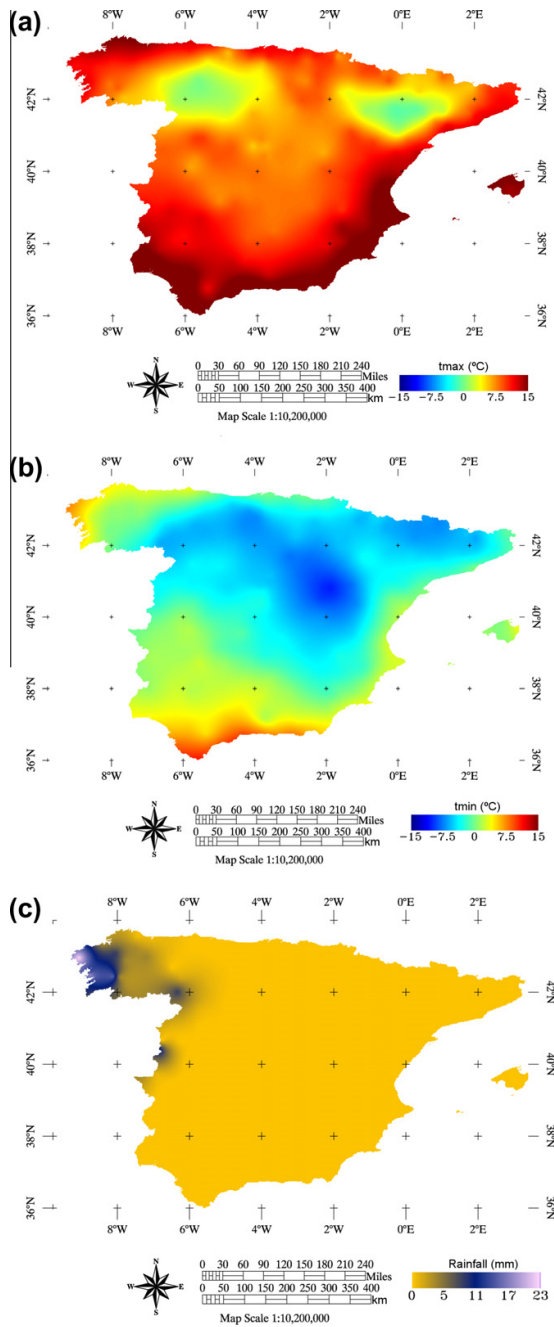


Fig. 3. Example of images obtained by kriging from the meteorological stations' data obtained on January 1st, 2008: (a) maximum temperature, (b) minimum temperature, (c) precipitation.

the training data. This is repeated such that each observation in the sample is used once as the validation data. The difference between the actual measurement and its

estimation gives an indication of how well this measurement follows the trend determined by the measured values at neighboring locations. The mean of the squared standardized cross-validation error is computed. This error provides an indication of the adequacy of the model and of its parameters.

In the actual case, the mean squared cross-validation error that has been obtained in temperature (T_{min} , T_{max}) is about 0.93 °C, and for rainfall is about 1.94 mm. This difference highlights the inherent difficulty of obtaining spatially continuous precipitation databases due to its spatial erratic behavior.

3. Results and discussion

The three methods are trained and tested using the irradiation data measured at 40 stations (where temperature and precipitation data are also available). In the BC method, 80% of the data is used for training (i.e. for finding the relationships between temperature, precipitation and global irradiation) and the remaining 20% is used for testing. In the ANN and KRR methods, 60% of the data are used for training. Another 25% of the data are used to optimize the parameters λ and σ in the case of KRR or to stop the training in the case of generalization loss of the ANNs. The remaining 15% is used for testing. After training and testing, these methods are compared in order to determine the best one (in terms of lower error, higher correlation coefficient, normality of the residuals, etc.). The best method is finally applied to the spatially continuous datasets obtained by ordinary kriging of the meteorological data measured at the 400 stations.

Due to the non-linear character of the ANN and KRR methods, a gamma test (see Appendix A) is applied before the training in order to determine the best combination of meteorological variables. The results of this test show that using the precipitation data, in addition to daily air temperature amplitude and extraterrestrial solar irradiation, improves significantly the accuracy of the irradiation estimate from these methods. Then, $\vec{x}_i \equiv [x_{1,i}, \dots, x_{m,i}]^T = [T_{max,i}, \bar{T}_{min,i}, P_i, H_{0i}]$ in ANN and KRR methods, being \bar{T}_{min} the average minimum temperature of two successive days.

3.1. Implementation of the BC method

Instead of using literature values (Goodin et al., 1999), the empirical coefficients a , b and c of the BC model have been calculated for Spain by means of a non-linear least squares fit to Eq. (3) for each radiation station and for every year. Table 1 shows those coefficients for each year and their mean values.

The low variance values in Table 1 reveal the spatial consistency and homogeneity among the estimates from different stations.

The coefficient a represents the clear sky transmittance and ranges between 0.75 and 0.82, reaching higher values

Table 1
Empirical coefficients for the BC equation and their variances. Values for each year and their mean are shown.

	a	var a	b	var b	c	var c
2005	0.753	0.005	0.030012000	0.000000008	1.65	0.02
2006	0.756	0.012	0.030026500	0.000000009	1.66	0.06
2007	0.812	0.008	0.03007650	0.000000003	1.56	0.02
2008	0.817	0.018	0.0298451	0.00000018	1.60	0.06
Mean	0.785	0.011	0.0299900	0.00000005	1.62	0.04

in the years 2007 and 2008, colder and wetter than the other two. The coefficient b is the most stable (lowest variance values in %) and practically independent of the year. The coefficient c presents the highest variance values (from 1.2% to 3.8%) and a behavior dissimilar to the coefficient a , that is, it increases in the hottest and driest years. Therefore, it seems that there are two groups of coefficients presenting slight differences between them. The first one compressing the years 2005 and 2006, extremely hotter and the driest of the considered period, and the second one compressing 2007 and 2008, classified as hot years but registering a rather higher precipitation. It seems that the empirical coefficients a and b increase with temperature, and c decreases with precipitation. Nevertheless, as the discrepancies between them are so small, a mean value (see the variances for this case in Table 1) could also be taken without loss of information. This is an outstanding result because sometimes no irradiation data are available to test the models.

3.2. Implementation of the ANN method

ANNs are widely accepted as an alternative way to tackle complex and ill-defined problems. They can be trained to predict results from example, and are fault tolerant in the sense that they are able to handle noisy and incomplete data. They are able to deal with non-linear problems, and once trained can perform prediction at very high speed. A critical issue in developing ANN is generalization. ANN can suffer from either over-fitting or under-fitting the training set. To achieve generalization, the early stopping method was used. The validation set was used to test the performance of the network during training process. When the error of validation set reached a minimum, then training process stopped.

3.2.1. The ANN method consists on a three-step approach

1. An exhaustive initialization of parameters and a test of the different architectures: 150 ANNs were trained, trying different architectures and initializations. This step allowed determining the optimal complexity of the MLP appropriate to modeling daily global solar irradiation (determine the number of hidden layers and the number of neurons in each hidden layer). Because the algorithm to obtain the parameters is a local minimum

search one, 10 synaptic weight initializations were employed (Jain and Fanelli, 2000).

2. Selection of the models that present the best performance: As a result of a heuristic adjustment, only the six ANNs presenting the lowest errors over the validation data set were chosen (network committee).
3. Estimation of global solar radiation at the surface: Final estimation of solar irradiation by the network committee.

3.3. Implementation of the KRR method

A cross-validation methodology has been followed to optimize the values of the regularization parameter λ (varied logarithmically between 10^{-4} and 10^2) and the kernel parameter σ (between 10^{-1} and 10^3). The optimization involved a 20-fold-cross-validation using the validation data set.

The number N of samples used for the training of KRR is 60% of the data corresponding to the 40 stations recording solar irradiation on a daily basis during the 4 years analyzed. Since the $(N \times N)$ matrix $\beta^{-1} \equiv \lambda I_N + K$ has to be inverted, the use of such a large amount of data can be reconsidered. The final equation of KRR, Eq. (11), involves a sum of N terms with coefficients (known as dual variables) α_i . Most of these coefficients are so small that the corresponding contribution to the sum is negligible. Therefore, the computational effort can be significantly reduced without affecting the accuracy of the solar irradiation estimate by using a smaller amount of training data. The histogram of values of the dual variables α_i in Fig. 4 shows a distribution peaked around zero and tails corresponding to a small fraction of variables α_i with large positive and large negative values.

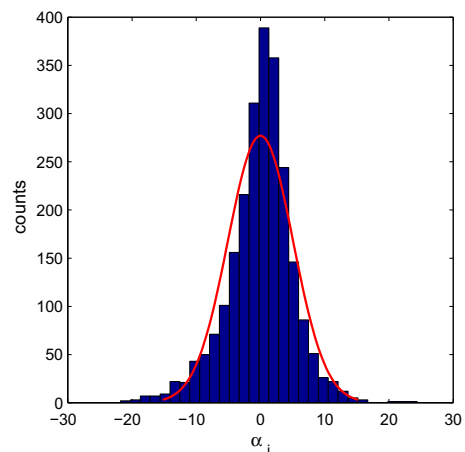


Fig. 4. Histogram of the dual variables α_i calculated in the training of KRR.

Table 2
Mean signed error (MSE), root mean squared error (RMSE), mean absolute error (MAE) and coefficient of determination (R^2) between estimated and measured daily solar irradiation during 4 years (2005–2008).

	MSE (MJ m ⁻² day ⁻¹)	RMSE (MJ m ⁻² day ⁻¹)	MAE (MJ m ⁻² day ⁻¹)	R^2
BC	0.43	4.80	3.29	0.71
ANN	-0.00	3.16	2.33	0.86
KRR	-0.01	3.33	2.45	0.85

3.4. Comparison of results from the three methods

In order to test the accuracy of the three methods described to estimate global solar irradiation, different error indicators have been evaluated as shown in Table 2. These errors are similar to the obtained by Fodor and Mika (2011), using a temperature-based S-shape method which considers separately days with and without precipitation, and somewhat lower than those found by Fortin

et al. (2008), using an ANN method not including precipitation data. Although the errors in Table 2 correspond to the complete data set (years 2005–2008), we also carried out the analysis considering only the year 2008 and no relevant differences were found in the parameter values of the three methods. Therefore, a 1-year dataset might be sufficient to obtain accurate irradiation estimates.

Table 2 shows that the two non-linear methods yield more accurate estimations than the classical BC method. ANN performs slightly better than KRR: ANN has lower values of RMSE and MAE, lower absolute value of MSE, and higher value of R^2 . The use of precipitation data in the ANN and KRR methods and their capability to find complex relationships between irradiation and the input variables, due to their non-linear character, leads to a substantial improvement in the irradiation estimates, in comparison to the BC method.

Fig. 5a shows a scatter plot of estimated versus observed values of solar irradiation including the three estimation methods. The clouds of points from ANN and KRR are

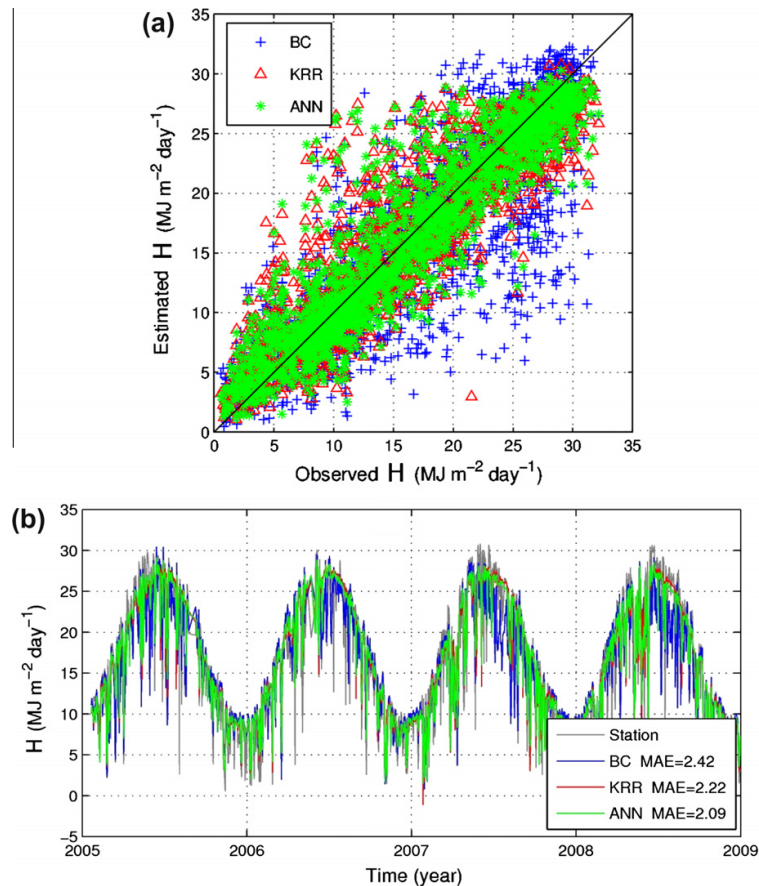


Fig. 5. (a) Performance of the models over the complete test data set from 2005 to 2008. (b) Performance of the models over train and test data in the Manises ground station.

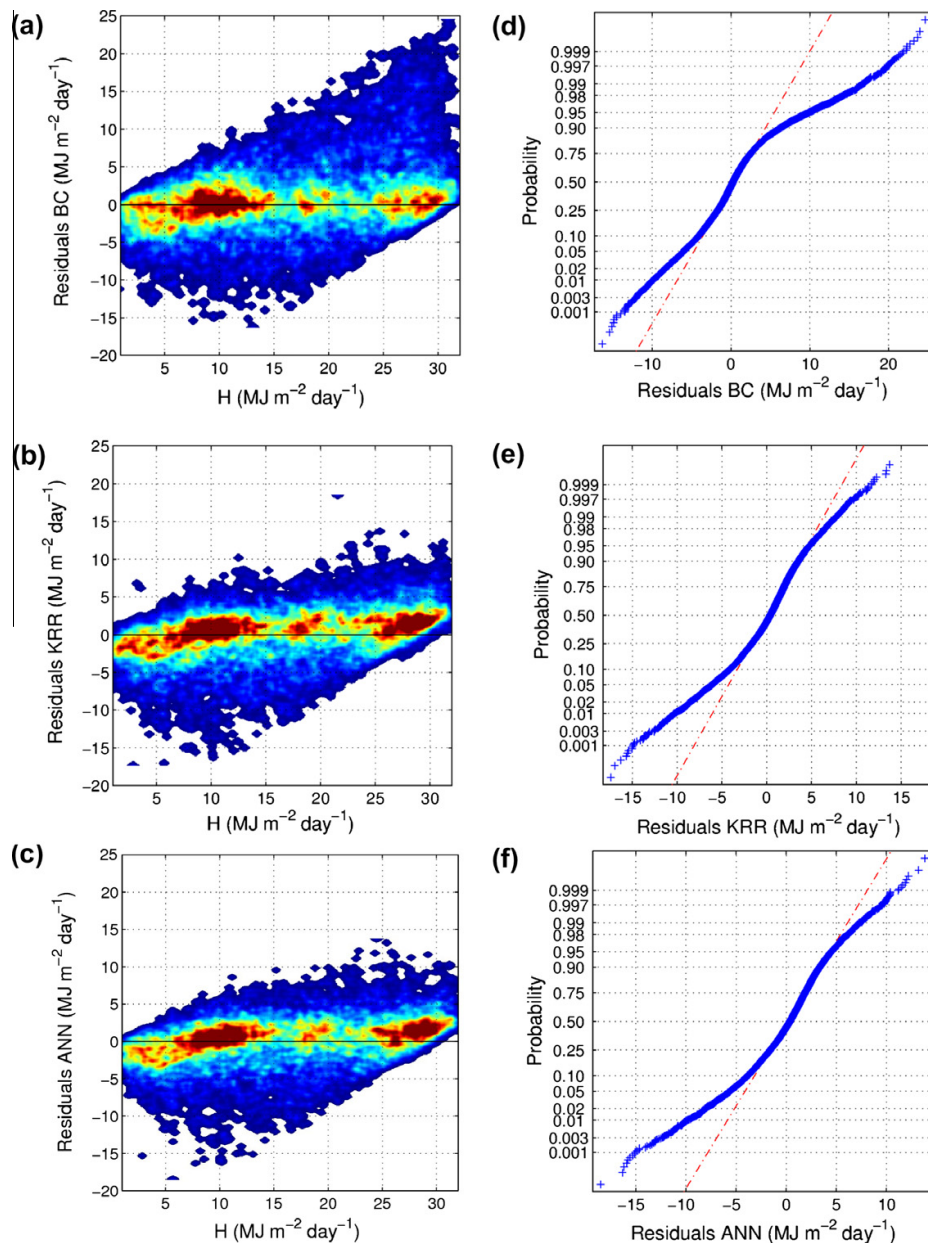


Fig. 6. Observed vs. residuals (a–c) and normal probability plots (d–f) for the three approaches. Red color indicates a higher density of points. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

closer to the 1:1 line than the BC cloud, as expected from the errors in Table 2. Fig. 5b shows the temporal profile of estimated and observed daily global irradiation corresponding to the Manises ground station (latitude 39°29'24" North and longitude 0°27'0" West). This figure shows a remarkable temporal consistency of the estimated profiles and confirms that the ANN and KRR methods have a higher accuracy.

Fig. 6a–c shows the analysis of residuals as scatter plots of the residuals vs. observed values. Since most residuals are (in absolute value) lower than 5 (predominant value), we conclude that there is no tendency between the value to be predicted and the obtained error. High residuals (>19 MJ m⁻² day⁻¹) are found for high observed values (>23 MJ m⁻² day⁻¹) in the case of the BC method. Therefore, the BC method presents a tendency to under-estimate

high global solar irradiation values (as concluded from the large positive residuals).

The normal probability plots in Fig. 6d–f serve to assess graphically whether the data follows a normal distribution. If the data were normal, the plot should be linear (as indicated by the dashed lines). The largest deviation from the linear behavior, especially for large positive values, is observed in the case of the BC method. Fig. 6e and f shows similar normal probability plots of the ANN and KRR methods, although the normality of the residuals from ANN is slightly better than those from KRR. Nevertheless, in both methods a fraction (ca. 15%) of the residuals lower than -4 and a fraction (ca. 5%) of the residuals larger than $+5$ deviate significantly from a normal distribution.

The results in Figs. 5 and 6 indicate that the ANN method yields the most accurate and lowest biased estimations of daily global irradiation, in comparison with the other two methods.

Fig. 7a and b shows two comparisons of the outcomes from BC vs. ANN and KRR vs. ANN. Fig. 7a shows that the BC method tends to overestimate the irradiation, especially for intermediate values, which might result from a bias of the BC method. Fig. 7b clearly evidences a good agreement between the estimates from the two non-linear methods, with a point cloud close to the 1:1 line.

3.5. Global solar irradiation maps from the ANN method

Daily images of global solar irradiation have been produced using the ANN method for the 4 years 2005–2008. Four images corresponding to different seasons for the year 2008 are shown in Fig. 8. Irradiation is clearly lower in autumn and winter, and higher in spring and summer. The presence of cloudy areas is more noticeable in spring and autumn (spots showing lower values). However, in sunny days the irradiation images look rather homogeneous.

4. Conclusions

Unlike temperature and precipitation, solar irradiation is recorded only at a limited number of weather stations in Spain. Solar irradiation maps cannot be obtained directly from the radiation measurements due to the low number of stations. However, irradiation can be estimated using extraterrestrial irradiation and the available meteorological data, such as air temperature and precipitation. In this work, three methods to estimate daily global solar irradiation on a horizontal surface have been developed and compared: the Bristow–Campbell (BC), Artificial Neural Network (ANN) and Kernel Ridge Regression (KRR). BC is a simple method based on air maximum and minimum temperature and using extraterrestrial irradiation. ANN and KRR are non-linear methods with a high capability to describe complex relationships between irradiation and the input variables. Their non-linear character, together with the additional use of precipitation data

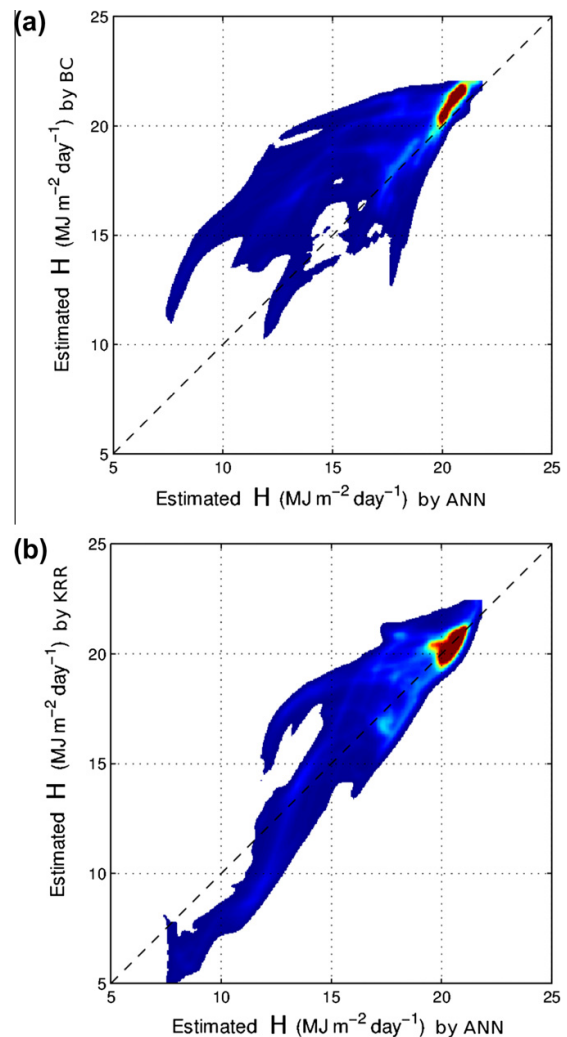


Fig. 7. Daily global irradiation estimated from (a) BC vs. ANN and (b) KRR vs. ANN methods. Red color indicates a higher density of points. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

(which significantly improves the outcomes as determined from a gamma test), makes them superior to the BC method.

The three methods are trained and tested using the irradiation data measured at 40 stations (where temperature and precipitation data are also available). In the BC method, 80% of the data (randomly chosen) is used for training (i.e. to calculate the empirical coefficients a , b and c for Spain) and the remaining 20% is used for testing. For the development of the ANN method, 150 ANNs with different architectures have been trained using 60% of the meteorological data (randomly chosen) as neuron inputs. Another 25% of the data are used to stop the training in

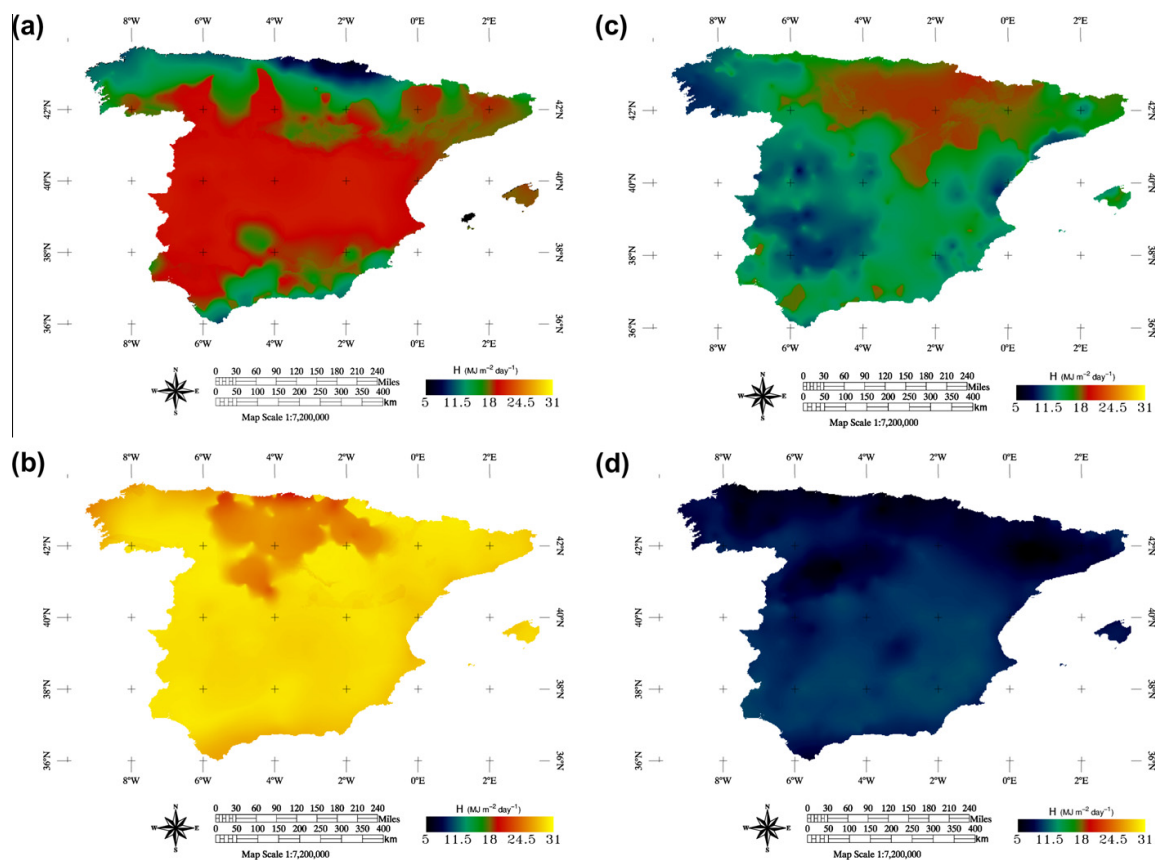


Fig. 8. Daily images of global solar irradiation over Spain produced using the ANN method for March 21th, 2008 (a), June 21th, 2008 (b), September 21th, 2008 (c), and December 21th, 2008 (d).

the case of generalization loss. The performance of these architectures has been tested by comparing their predictions to the remaining 15% of the data.

The six ANNs which have yielded a minimum mean square error have been chosen as the “network committee” and the final output has been evaluated as the mean of their results. The training of the KRR method has been performed using 60% of the data. Another 25% of the data have been used to optimize the values of the regularization parameter λ and the kernel parameter σ . The method has been tested using the remaining 15% of the data.

The mean signed error (MSE), root mean squared error (RMSE), mean absolute error (MAE) and the coefficient of determination (R^2) between estimated and measured daily solar irradiation during the 4 years have been calculated for the three methods. All of them are rather low, and similar or lower than the errors obtained in recently published papers on this topic (Rehman and Mohandes, 2008; Remesan et al., 2008). The ANN method, with a MAE = $2.33 \text{ MJ m}^{-2} \text{ day}^{-1}$, performs slightly better than the other two methods and, therefore, it has been selected to obtain the solar irradiation maps from temperature and

precipitation images derived by ordinary kriging from the 400 stations’ data. Daily global solar irradiation maps over Spain are derived from this methodology at 1-km spatial resolution.

Acknowledgments

This research was supported by the Spanish Ministry of Science and Innovation (Project ARTEMIS CGL2008-00381). The data were kindly provided by AEMet (Agencia Estatal de Meteorología, Spain). The authors thank Dr. J. Tamayo from AEMet for his assistance, and Dr. G. Valls for his useful comments on the non-linear methods used. Special thanks are due to the anonymous reviewers for their suggestions.

Appendix A. Gamma test

A Gamma test using winGamma™ software (Durrant, 2001) has been carried out to determine the best combination of input data to train the ANN and KRR methods. This test is also designed to solve the overtraining problem

Table A1
Gamma test results on daily global solar radiation measurements as function of different sets of meteorological data.

	Input variable combinations					
	$T_{\max}, \bar{T}_{\min}H_0$	$T_{\max}, T_{\min(t)}T_{\min(t+1)}, H_0$	$T_{\max}, \bar{T}_{\min}P, H_0$	T_{\max}, \bar{T}_{\min}	\bar{T}_{\min}, H_0	T_{\max}, H_0
<i>A</i>	0.9304	0.3907	0.1704	-0.7709	3.8643	1.9636
<i>V_{ratio}</i>	0.1804	0.1734	0.1405	0.3770	0.3144	0.2342
<i>SE</i>	0.0005	0.0005	0.0007	0.0011	0.0008	0.0005

associated with most non-linear modeling techniques (Remesan et al., 2008). The test provides information on the possibility of finding a generic smooth function *f* of the input variables $\vec{x} = [x_1, \dots, x_m]^T$ to generate the output *y*

$$y = f(\vec{x}) + \rho \tag{A.1}$$

where ρ is a Gaussian noise of zero mean and finite variance. However, the test does not make use of any particular function and simply consists of a near neighbor statistics derived from the data (\vec{x}_i, y_i) ($1 \leq i \leq N$). The dataset (\vec{x}_i, y_i) ($1 \leq i \leq N$) is normalized so that the transformed data have zero mean and standard deviation equal to 0.5. If we consider a point (\vec{x}_i, y_i) of the dataset, the point (\vec{x}_j, y_j) of the dataset is the nearest neighbor of (\vec{x}_i, y_i) if $\|\vec{x}_j - \vec{x}_i\|$ is minimal. This first neighbor is denoted as $\vec{x}_{i,1}$ and its corresponding value of the output is $y_{i,1}$. Similarly we can define the *k*th nearest neighbors of (\vec{x}_i, y_i) and denote them as $\vec{x}_{i,k}$ and $y_{i,k}$. The near neighbor statistical analysis is based on evaluating the quantities

$$\delta_k \equiv \frac{1}{N} \sum_{i=1}^N \|\vec{x}_{i,k} - \vec{x}_i\|^2 + \tag{A.2}$$

$$\gamma_k \equiv \frac{1}{2N} \sum_{i=1}^N (y_{i,k} - y_i)^2 \tag{A.3}$$

for $1 \leq k \leq p$, where *p* is the maximum number of nearest neighbors considered. Then, a regression analysis of the *p* points (δ_k, γ_k) by ordinary least squares method yields the regression line

$$\gamma = A\delta + \Gamma \tag{A.4}$$

The intercept Γ of the regression line is known as the Gamma statistic and represents the estimated variance of the errors for any model built on the data, as generally described by Eq. (A.1). The slope *A* of the regression line provides a rough measure of the complexity of the smooth function *f*. The standard error

$$SE = \sqrt{\frac{1}{p-2} \sum_{k=1}^p (\gamma_k - \bar{\gamma})^2} \tag{A.5}$$

is an estimate of the confidence degree of the Gamma statistic.

The raw dataset and the normalized dataset yield different values of *A* and Γ . In order to have a scale-invariant measure, normally in the range [0, 1], the ratio

$$V_{ratio} = \Gamma / Var(y) \tag{A.6}$$

is calculated; with the normalization explained above the variance of *y* is $Var(y) = 0.25$. A small value of *V_{ratio}* indicates that the output *y* is highly likely to be derived from the input data \vec{x} by a smooth function *f*. In order to determine the best Gamma statistic (characterized by lower values of *V_{ratio}* and *SE*) the test has been carried out, as usual, for increasing maximum number of nearest neighbors, and the value $p = 12$ has been finally chosen. A model with low *V_{ratio}*, low *A* and low *SE* is considered the best scenario (Durrant, 2001). According to this, the third combination of input variables in Table A1 (T_{\max}, \bar{T}_{\min} , precipitation, H_0) has been selected. Comparison of the first and third combinations leads to the conclusion that the incorporation of the precipitation variable improves significantly the estimation of solar radiation.

References

Alcaraz, D., Paruelo, J., Cabello, J., 2006. Identification of current ecosystem functional types in the Iberian Peninsula. *Glob. Ecol. Biogeogr.* 15, 200–212.

Almorox, J., Hontoria, C., 2004. Global solar radiation estimation using sunshine duration in Spain. *Energy Convers. Manage.* 45, 1529–1535.

Almorox, J., Benito, M., Hontoria, C., 2005. Estimation of monthly Angström–Prescott equation coefficients from measured daily data in Toledo, Spain. *Renewable Energy* 30, 931–936.

Alpaydin, E., 2004. *Introduction to Machine Learning (Adaptive Computation and Machine Learning)*. MIT Press, pp. 445.

Alsamamra, H., Ruiz-Arias, J.A., Pozo-Vázquez, D., Tovar-Pescador, J., 2009. A comparative study of ordinary and residual kriging techniques for mapping global solar radiation over southern Spain. *Agric. Forest Meteorol.* 149, 1343–1357.

Ångström, A., 1924. Solar and terrestrial radiation. *Quart. J. Roy. Meteorol. Soc.* 50, 121–125.

Badescu, V., 2008. *Modeling solar radiation at the Earth’s surface. In: Recent Advances*. Springer, Berlin.

Bakirci, K., 2009. Correlations for estimation of daily global solar radiation with hours of bright sunshine in Turkey. *Energy* 34, 485–501.

Ball, R.A., Purcell, L.C., Carey, S.K., 2004. Evaluation of solar radiation prediction models in North America. *Agron. J.* 96, 391–397.

Bishop, C.M., 1996. *Neural Networks for Pattern Recognition*. Oxford University Press, USA.

Bristow, K.L., Campbell, G.S., 1984. On the relationship between incoming solar radiation and daily maximum and minimum temperature. *Agric. Forest Meteorol.* 31, 159–166.

Buytaert, W., Célleri, R., Bièvre, B., Cisneros, F., Wyseu, G., Deckers, J., Hofstede, R., 2006. Human impact on the hydrology of the Andean páramos. *Earth-Sci. Rev.* 79, 53–72.

Cawley, G.C., Talbot, N.L.C., 2003. Efficient leave-one-out cross validation of kernel fisher discriminant classifiers. *Pattern Recogn.* 36, 2585–2592.

Chuanyan, Z., Zhongren, N., Guodong, C., 2005. Methods for modelling of temporal and spatial distribution of air temperature at landscape scale in the southern Qilian mountains. *China Ecol. Model.* 189, 209–220.

- Durrant, P.J., 2001. winGamma™: A Non-linear Data Analysis and Modelling Tool with Applications to Flood Prediction. Ph.D. Thesis, Department of Computer Science, University of Wales, Cardiff, Wales UK.
- Fodor, N., Mika, J., 2011. Using analogies from soil science for estimating solar radiation. *Agric. Forest Meteorol.* 151, 78–86.
- Fortin, J., Anctil, F., Parent, L., Bolinder, M., 2008. Comparison of empirical daily surface incoming solar radiation models. *Agric. Forest Meteorol.* 148, 1332–1340.
- Gómez, J., Martín, J.D., Soria, E., Vila, J., Carrasco, J.L., Valle, S., 2006. Neural networks for analysing the relevance of input variables in the prediction of tropospheric ozone concentration. *Atmos. Environ.* 40, 6173–6180.
- Goodin, D.G., Hutchinson, J.M.S., Vanderlip, R.L., Knapp, M.C., 1999. Estimating solar irradiance for crop modeling using daily air temperature data. *Agron. J.* 91, 845–851.
- Goovaerts, P., 1997. *Geostatistics for Natural Resources Evaluation*. Oxford University Press, New York, NY, USA, pp. 483.
- Guyon, I., 2005. Kernel Ridge Regression. <<http://clopinet.com/isabelle/Projects/ETH/KernelRidge.pdf>>.
- Haykin, S., 1999. *Neural Networks: A Comprehensive Foundation*. Prentice Hall Upper, Saddle River, NJ, USA, 842pp.
- Hugues, J.P., Lettenmaier, D.P., 1981. Data requirements for kriging: estimation and network design. *Water Resour. Res.* 17, 1641–1650.
- Immerzeel, W.W., Rutten, M.M., Droogers, P., 2009. Spatial downscaling of TRMM precipitation using vegetative response on the Iberian Peninsula. *Remote Sens. Environ.* 113, 362–370.
- Iqbal, M., 1983. *An Introduction to Solar Radiation*. Academic Press, Canada, 390pp.
- Isaaks, E.H., Srivastava, R.M., 1989. *Applied Geostatistics*. Oxford University Press, New York, 522pp.
- Jain, L., Fanelli, A.M., 2000. *Recent Advances in Artificial Neural Networks: Design and Applications*. CRC Press, Inc, Boca Raton, FL, USA, PA.
- Jarvis, E.H., Stuart, N., 2001. A comparison among strategies for interpolating maximum and minimum air temperatures. *J. Appl. Meteorol.* 40, 1075–1084.
- Keerthi, S.S., Lin, C.J., 2003. Asymptotic behaviors of support vector machines with Gaussian Kernel. *Neural Comput.* 15 (7), 1667–1689.
- Kermani, B.G., Schiffman, S.S., Nagle, H.T., 2005. Performance of the Levenberg–Marquardt neural network training method in electronic nose applications. *Sens. Actuat., B* 110, 13–22.
- Liu, X., Mei, X., Li, Y., Zhang, Y., Wang, Q., Jensen, J.R., Porter, J.R., 2009a. Calibration of the Ångström–Prescott coefficients (a, b) under different time scales and their impacts in estimating global solar radiation in the Yellow River basin. *Agric. Forest Meteorol.* 149, 697–710.
- Liu, X., Mei, X., Li, Y., Wang, Q., Jensen, J.R., Zhang, Y., Porter, J.R., 2009b. Evaluation of temperature-based global solar radiation models in China. *Agric. Forest Meteorol.* 149, 1433–1446.
- López, G., Battles, F.J., Tovar-Pescador, J., 2005. Selection of input parameters to model direct solar irradiance by using artificial neural networks. *Energy* 30, 1675–1684.
- Lourakis, M.I.A., 2005. A Brief Description of the Levenberg–Marquardt Algorithm Implemented by Levmar <<http://www.ics.forth.gr/lourakis/levmar/levmar.pdf>>.
- Madsen, K., Nielsen, H.B., Tingleff, O., 2004. *Methods for Non-linear Least Squares Problems*, second ed. IMM, Technical University of Denmark, 50pp <<http://www2.imm.dtu.dk/pubdb/views/edocdownload.php/3215/pdf/imm3215.pdf>>.
- Martínez, B., Cassiraga, E., Camacho, F., Garcia-Haro, J., 2010. Geostatistics for mapping leaf area index over a cropland landscape: efficiency sampling assessment. *Remote Sens.* 2, 2584–2606.
- Maselli, F., Barbati, A., Chiesi, M., Chirici, G., Corona, P., 2006. Use of remotely sensed and ancillary data for estimating forest gross primary productivity in Italy. *Remote Sens. Environ.* 100, 563–575.
- Maselli, F., Papale, D., Puletti, N., Chirici, G., Corona, P., 2009. Combining remote sensing and ancillary data to monitor the gross productivity of water-limited forest ecosystems. *Remote Sens. Environ.* 113, 657–667.
- Mellit, A., Benghanem, M., Hadj Arab, A., Guessoum, A., 2005. A simplified model for generating sequences of global solar radiation data for isolated sites: using artificial neural network and a library of Markov transition matrices approach. *Sol. Energy* 79, 469–482.
- Menges, H.O., Ertekin, C., Sonmete, M.H., 2006. Evaluation of global solar radiation models for Konya, Turkey. *Energy Convers. Manage.* 47, 3149–3173.
- Paulescu, M., 2008. Solar irradiation via air temperature data. In: Badescu, V. (Ed.), *Modeling Solar Radiation at the Earth's Surface. Recent Advances*. Springer, Berlin, pp. 175–192.
- Pons, X., Ninyerola, M., 2008. Mapping a topographic global solar radiation model implementation in a GIS and refined with ground data. *Int. J. Climatol.* 28, 1821–1834.
- Prescott, J.A., 1940. Evaporation from a water surface in relation to solar radiation. *Trans. Roy. Soc. Sci. Aust.* 64, 114–125.
- Rehman, S., Mohandes, M., 2008. Artificial neural network estimation of global solar radiation using air temperature and relative humidity. *Energy Policy* 36, 571–576.
- Remesan, R., Shamim, M.A., Han, D., 2008. Model data selection using gamma test for daily solar radiation estimation. *Hydrol. Process.* 22, 4301–4309.
- Robaa, S.M., 2009. Validation of the existing models for estimating global solar radiation over Egypt. *Energy Convers. Manage.* 50, 184–193.
- Shawe-Taylor, J., Cristianini, N., 2004. *Kernel Methods for Pattern Analysis*. University Press, Cambridge.
- Smithson, P., Addison, K., Atkinson, K., 2002. *Fundamentals of the Physical Environment*, Third ed. Routledge Taylor and Francis Group, London and New York.
- Tabios, G.Q., Salas, J.D., 1985. A comparative analysis of techniques for spatial interpolation of precipitation. *Water Resour. Res.* 21, 365–380.
- Tymvios, F.S., Jacovides, C.P., Michaelides, S.C., Scouteli, C., 2005. Comparative study of Ångström's artificial neural networks' methodologies in estimating global solar radiation. *Sol. Energy* 78, 752–762.
- Veroustraete, F., Sabbe, H., Eerens, H., 2002. Estimation of carbon mass fluxes over Europe using the C-Fix model and Euroflux data. *Remote Sens. Environ.* 83, 376–399.
- Vicente-Serrano, S.M., Pérez-Cabello, F., Lasanta, T., 2008. Assessment of radiometric correction techniques in analyzing vegetation variability and change using time series of Landsat images. *Remote Sens. Environ.* 112, 3916–3934.
- Webster, R., Oliver, M.A., 2001. *Geostatistics for Environmental Scientist*. John Wiley and Sons, Chichester, England.
- Weiss, A., Hays, C.J., Hu, Q., Easterling, E., 2001. Incorporating bias error in calculating solar irradiance. Implications for crop yield simulations. *Agron. J.* 93, 1321–1326.
- Woodcock, C.E., Strahler, A.H., Jupp, D.L.B., 1988. The use of variograms in remote sensing: I. Scenes models and simulated images. *Remote Sens. Environ.* 25, 323–348.

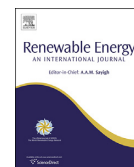
Apéndice **B**

Publication II: Validation of daily global solar irradiation images from MSG over Spain



Contents lists available at SciVerse ScienceDirect

Renewable Energy

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

Validation of daily global solar irradiation images from MSG over Spain

A. Moreno^a, M.A. Gilabert^{a,*}, F. Camacho^b, B. Martínez^a^aDepartament de Física de la Terra i Termodinàmica, Universitat de València, Dr. Moliner, 50. 46100 Burjassot, Valencia, Spain^bEarth Observation Laboratory (EOLAB), Parc Científic Universitat de València, Catedràtic José Beltrán, 2. 46980 Paterna, Valencia, Spain

ARTICLE INFO

Article history:

Received 20 November 2012

Accepted 8 May 2013

Available online

Keywords:

Daily irradiation

Validation

MSG

ABSTRACT

Daily irradiation images over Spain – area that embraces a highly heterogeneous landscape, climatic conditions and relief – are calculated from the down-welling surface short-wave radiation flux (DSSF) product derived from the MSG SEVIRI images. Their analysis and validation is carried out using two different station networks along the year 2008. The first network covers the peninsular Spain and Balearic islands. A denser one, covering the Catalanian territory and including many stations located in rugged terrain, is found useful to assess the elevation correction to be applied to the images. The statistics from the validation using the first network shows a relative mean bias of about 1%, a relative mean absolute difference of 6%, and a mean absolute difference of 1.0 MJ m^{-2} . The analysis of the second database shows that the elevation correction reduces the relative mean bias, for rugged terrains and for clear sky data, from 5% to 0.5%, whereas for the complete sampling the mean absolute difference of the derived daily irradiation images is 1.3 MJ m^{-2} . A downscaling of the DSSF product is also carried out, and a methodology to obtain topographically-corrected daily irradiation images, based on merging the DSSF with a digital elevation model, is proposed. These images satisfactorily map the surface solar radiation at 1-km spatial resolution even in rugged terrains.

© 2013 Elsevier Ltd. All rights reserved.

1. Introduction

The surface solar radiation is a crucial variable involved in a number of applications including weather and climate monitoring, environmental modeling, and solar energy availability assessment for heating and electrical power generation purposes [1–3]. The photosynthetic active radiation, which drives plant productivity [4] and vegetation distribution, can be also derived from the incoming solar radiation flux [5]. As the availability of ground stations recording solar radiation data has proven to be spatially inadequate for many applications, the use of spatially continuous databases mapping solar radiation is highly recommended. These images can be obtained by interpolating radiation measurements from a network of meteorological stations [6]. The interpolation accuracy depends on the network density: the larger the distance between ground stations, the larger the error [7,8]. Since the number of meteorological stations recording solar radiation is much lower than those recording temperature, precipitation, sunshine duration

and relative humidity data, different computational procedures have been developed to estimate solar radiation from the other available meteorological data [9–11]. Then, the interpolation accuracy increases and it is possible to generate reliable solar radiation maps [12].

Estimates of the surface solar radiation are also provided by geostationary satellites such as Meteosat and Meteosat Second Generation (MSG) [13,14], GMA, GOES and MTSat [15], which cover a wide area with adequate temporal frequency. The down-welling short-wave surface radiation flux (DSSF or irradiance, in W m^{-2}) quantifies the energy flux density (i.e., the integral of the spectral irradiance) in the wavelength range from $0.3 \mu\text{m}$ to $4.0 \mu\text{m}$ reaching the Earth's surface. The solar elevation and the cloud coverage are the main factors affecting the atmospheric transmittance, and hence the DSSF. Hence, the satellite observations can lead to an improvement of the radiation estimates because their higher temporal and spatial resolution allow for a precise location of clouds. Moreover, their radiation estimates present a similar quality over the entire observed area, with the exception of highly variable, mountainous areas [16–18]. For these complex terrains, the combination of satellite images and a digital elevation model (DEM) can increase the accuracy of the solar radiation maps [17,19].

* Corresponding author. Tel.: +34 963 543 118; fax: +34 963 543 385.
E-mail address: m.amparo.gilabert@uv.es (M.A. Gilabert).

There exist several methods to derive surface solar radiation from geostationary satellite observations. These methods are based on physical, empirical or hybrid models. As usual, the physical models require the resolution of radiation transfer equations and some input data often available, whereas the empirical models are based on regressions between satellite and ground data. The Heliosat method (in its different versions) [14,20,21] is a good example of an empirical approach, whereas the Land Surface Analysis Application Facility (LSA SAF) algorithm [13] makes use of a physical model of radiation transfer.

1.1. LSA SAF algorithm

MSG is the current generation of European geostationary meteorological satellites observing the Earth's full disc with the Spinning Enhanced Visible and Infrared Imager (SEVIRI) radiometer. SEVIRI is a 50 cm-diameter aperture, line-by-line scanning radiometer, which provides image data in four visible and near-infrared channels and eight infrared channels. A key feature of this imaging instrument is its continuous imaging of the Earth in 12 spectral channels, with a 15 min repeat cycle. The high temporal resolution of this sensor, which can provide 96 observations per day, has promoted the interest on its application to capture the diurnal cycle of the solar irradiance at the Earth's surface. The imaging sampling distance is 3 km at the sub-satellite point for standard channels, and a nominal 1-km resolution for the remaining high-resolution visible channel [22]. The LSA SAF algorithm [13] estimates the DSSF from the three short-wave SEVIRI channels (centered at 0.6, 0.8, 1.6 μm) at 3-km spatial resolution. This algorithm makes use of different parameterization procedures for clear and cloudy sky conditions because the DSSF is mostly determined by solar elevation and cloud coverage. Therefore, the cloud mask, developed by the SAF on Support to Nowcasting and Very Short-Range Forecasting [23], plays a crucial role in the algorithm. The DSSF product is generated with a temporal frequency of 30 min at the full spatial resolution of MSG SEVIRI instrument.

For clear sky conditions, the flux estimate is determined with a parameterization of the effective transmittance of the atmosphere [13] as a function of the concentration of air constituents (mainly water vapor, ozone, CO_2 , O_2 and aerosols), the surface albedo [24] and the spherical albedo of the atmosphere. In particular, the water vapor estimate is obtained from numerical weather forecasts of the European Centre for Medium-Range Weather Forecasts (ECMWF), and the ozone amount from the Total Ozone Mapping Spectrometer (TOMS). For cloudy sky conditions, a simple physical model of radiation transfer in the cloud–atmosphere–surface system is employed [13]. The effective transmittance is then depending also on the atmospheric transmittance between the surface and the cloud layer, the cloud transmittance, and the cloud albedo. The last two variables change on short time scales due to the temporal evolution of the clouds. The signal at the top of the atmosphere comprises contributions from Rayleigh scattering by the atmosphere above the cloud, radiation reflected by the cloud that is attenuated by the atmosphere above, and radiation reflected by the surface that is attenuated by the atmosphere and the cloud.

The estimated DSSF product derived from LSA SAF covers Europe, Africa, Middle East and parts of South America and it has been validated with *in situ* data from six European ground measurement stations in Portugal, France, Switzerland, United Kingdom and Estonia [13] throughout two years. The results showed a difference between instantaneous satellite estimates and ground measurements of about 40 and 110 W m^{-2} for clear and cloudy sky conditions, respectively. A more thorough validation of this product was

carried out by Journée and Bertrand [25] over Belgium, using measurements from 13 ground stations distributed all over the country. When considering instantaneous values, an average difference between estimates and measurements of about 70 W m^{-2} was found, although a tendency to overestimate the surface irradiance in overcast situations and to underestimate it in very clear sky situations was detected. Comparison between satellite retrievals and ground measurements was also performed on longer temporal horizons by considering integrated values over a given time period. It was shown that the accuracy of the product increases when considering daily averages rather than instantaneous values.

1.2. Objectives

The daily global irradiance H , expressed in MJ m^{-2} , is computed as the integral of DSSF images along the day. The main objective of this work is the analysis and validation of these daily irradiation images over Spain (except the Canary Islands), a region with highly heterogeneous landscape, climatic conditions and relief, which contains a great number of ground stations. A stratification of residuals in terms of spatial, temporal, missing data, topographic and model-related features is carried out to better identify the uncertainties sources. In addition, corrections procedures that improve the spatial resolution of the original DSSF product and take into account the topography (elevation effect) are also applied and assessed.

Compared to Belgium [25], Spain is much larger and rugged (the highest point in the Iberian Peninsula is at 3478 m, whereas in Belgium it is at 694 m). Moreover, different climates are found throughout the country: semiarid, Mediterranean, Atlantic and alpine climates, with precipitation ranging from 120 mm to 2000 mm. Belgium belongs practically to the Atlantic climatic zone; its climate is maritime temperate, with significant precipitation in all seasons. To carry out this validation two different datasets of daily solar global irradiation measurements during the year 2008 have been used. The first one comprises all available meteorological stations belonging to the AEMet (Spanish Meteorological Agency) network spread over the Spanish peninsular territory and the Balearic Islands. The other one comprises the stations spread only over Catalonia (NE Spain, with elevations ranging from sea level to about 3000 m), which belongs to the SMC (Catalonian Meteorological Service). The second dataset allows a precise analysis of the elevation effects because the Catalonian database has a denser measurement network in a more complex topography, whereas most of the stations from the first network are located in farmland and populated areas, usually throughout valleys and flat regions.

A previous study over Spain [12] showed that daily global irradiation images could be obtained by applying an artificial neural network to temperature and precipitation maps generated by ordinary kriging from *in situ* measurements. The mean absolute difference obtained was rather low, about 2.3 MJ m^{-2} , but the procedure could not be used in the near real time. The use of the daily irradiation images from DSSF can eliminate this drawback, which is a major advantage from an operational point of view.

2. Data

2.1. Ground data

The global irradiance is defined as the solar irradiance received from a 2π steradians solid angle on a horizontal surface. It includes irradiance received directly from the solar disk and the diffuse irradiance scattered by the atmosphere. Radiometric stations in the two networks are equipped for the measurement of global and

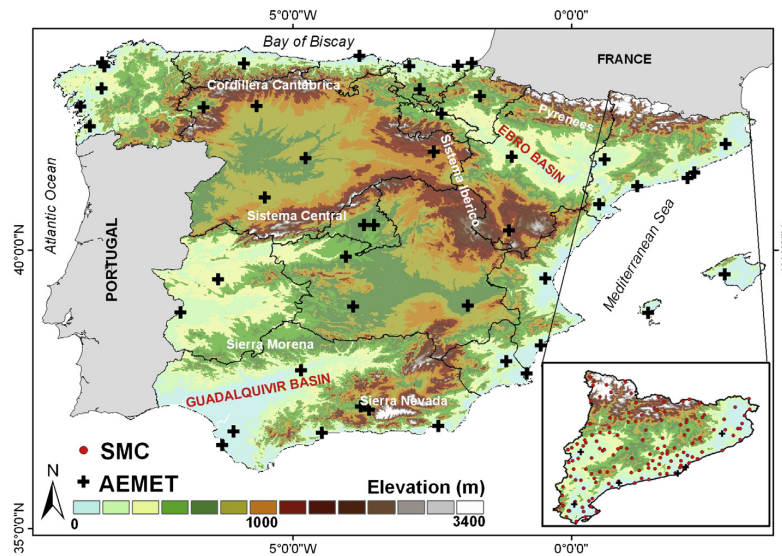


Fig. 1. Orography of study area with the location of the ground stations from the AEMet network (✚) and the SMC network (●).

diffuse irradiance with Kipp–Zonen pyranometers, models CM-11 and CM-21, with a spectral range of 305–2800 nm and an uncertainty of $\pm 2\%$. The experimental dataset includes daily global irradiation (H) measured at the surface from two networks (see Fig. 1) in 24 uniformly distributed dates along the year 2008. The first one (AEMet) includes around 45 ground stations spread over Spain. The second network (SMC) is much denser, and comprises 140 stations over Catalonia, 40 of them located in hilly areas near the Pyrenees. The actual number of available stations depends on the date.

The year 2008 was rather wet, with a precipitation above the average record during the last decade. It allows us to carry out the validation in a scenario not particularly favorable because the cloud coverage affects significantly the DSSF, and hence to obtain realistic uncertainty estimates.

2.2. Satellite data

The Satellite Application Facility (SAF) network is supported by the European Organization for the Exploitation of Meteorological Satellites (Eumetsat). The LSA SAF (<http://landsaf.meteo.pt/>) has been used to access and download the images of DSSF at the Earth's surface derived from the MSG images. This SAF follows a near real-time strategy with an emphasis on the land surfaces. Around 1300 files were processed to calculate the irradiation images as described in the next section.

2.3. Digital elevation model (DEM)

The DEM of Spain was extracted from a 3 arc second (approx. 90 m) globally available DEM from the Shuttle Radar Topography Mission (SRTM) (<http://edcftp.cr.usgs.gov/pub/data/>). A pre-processing step consisted of replacing bad values (e.g., occurring in radars shadows) using a surface fitting (Delaunay triangulation) technique. Data were then projected and resampled into the final (1 km) grid.

3. Methodology

Daily irradiation H images were generated from instantaneous DSSF satellite data as the sum of $[0.5 (DSSF_{t_1} + DSSF_{t_2}) (t_2 - t_1)]$,

for all pairs (t_1, t_2) with valid data from 0 to 23:30 UTC. In case of missing data, the same procedure was applied to a wider $(t_2 - t_1)$ interval [26]. In addition to this process of image integration, two extra layers were created attending to the quality flag information of each time step (every 30 min). The first one is the percentage of use of the cloud-condition algorithm in the integration process, which refers to the fraction (in %) of the diurnal estimates that use the cloudy model. The second one is the fraction (in %) of the diurnal estimates used to compose each pixel in the daily irradiation image that are missing.

A three-step methodology is proposed in this work. The first step consists of a validation of the estimated H product using the AEMet network data. In the second step, the SMC database over the Catalanian territory allows us to further analyze the topographic effects and to assess an improved elevation correction of the daily irradiation images. Finally, a procedure to increase the spatial resolution of the DSSF product (and hence of H), which is based on

Table 1

Statistical indices used to measure the quality of the daily irradiation satellite estimates $H_e(i)$ when compared to measured data $H_m(i)$; n is the number of data pairs.

Index	Definition
Mean bias difference	$MBD = \frac{1}{n} \sum_{i=1}^n (H_e(i) - H_m(i))$
Mean absolute difference	$MAD = \frac{1}{n} \sum_{i=1}^n H_e(i) - H_m(i) $
Root mean squared difference	$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (H_e(i) - H_m(i))^2}$
Relative mean bias difference	$rMBD = \frac{MBD}{H_m}$
Relative mean absolute difference	$rMAD = \frac{MAD}{H_m}$
Relative root mean square difference	$rRMSD = \frac{RMSE}{H_m}$
Coefficient of correlation	$R = \frac{\sum_{i=1}^n (H_e(i) - \bar{H}_e)(H_m(i) - \bar{H}_m)}{\sqrt{\sum_{i=1}^n (H_e(i) - \bar{H}_e)^2 \sum_{i=1}^n (H_m(i) - \bar{H}_m)^2}}$

spatial disaggregation of the irradiance by means of a DEM that also accounts for the coastline correction, is suggested.

3.1. Step 1: Validation of the daily irradiation images from the DSSF product over Spain using the AEMet dataset

The impact of clouds, missing data in the daily series and elevation on the irradiation estimation is analyzed using the statistical indices shown in Table 1.

3.2. Step 2: Validation of the daily irradiation images from the DSSF product over Catalonia using the SMC dataset. Assessment of the elevation correction

The atmospheric optical depth varies with elevation. Hence, the atmospheric transmittance is directly affected by it. The elevation

correction applied to improve the DSSF product, and thus the H images, has been assessed using the SMC data. The correction procedure proposed by Ruiz-Arias et al. [17] estimates the global irradiance from the top-of-atmosphere irradiance using an atmospheric transmittance $\tau(z)$ that depends on the altitude z (in km), as shown by equations (1) and (2)

$$\tau(z) = \tau(z_0) \cdot \beta^{z_0-z}, \quad z < 2 \text{ km} \tag{1}$$

$$\tau(z) = \tau(z_0) \cdot \beta_1^{z_0-2} \beta_2^{2-z}, \quad z \geq 2 \text{ km} \tag{2}$$

where $\tau(z_0)$ is the atmospheric transmittance at sea level, and $\beta = \beta_2 = 1.20$ and $\beta_1 = 1.30$ (1.20 if the Linke atmospheric turbidity is greater than 5) are empirical parameters. These equations apply to

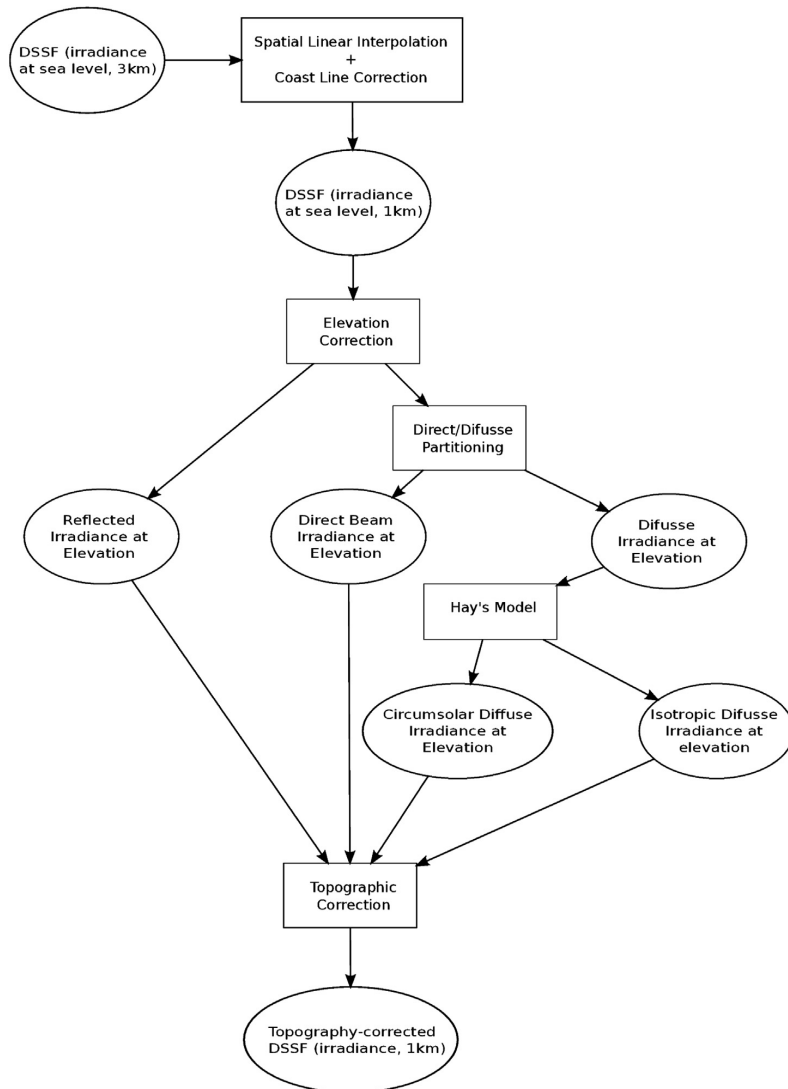


Fig. 2. Processing chain from the original DSSF product (3 km spatial resolution) to the final topographically-corrected irradiance images (1 km spatial resolution).

cloudless conditions, and hence the elevation correction was only carried out for air-mass-corrected clearness indices above 0.65 [27].

3.3. Step 3: Downscaling and topographic correction of the DSSF product

It has been shown that the disaggregation methods (which combine DEMs with satellite-derived radiation maps) produce increasingly detailed maps and improve their accuracy, especially in rugged terrains [17,19,28]. In this work, the downscaling of the DSSF product involves a linear spatial interpolation, a coastline correction and a topographic correction that includes the elevation correction (see Step 2) and slope and aspect effects by means of a DEM. Fig. 2 shows a flow chart describing the methodology in Step 3. The DSSF data are, in the native geostationary projection, centered at 0° longitude and with a sampling distance of 3 km at the sub-satellite point. First, these data have been re-projected and spatially smoothed by means of a linear interpolation to obtain

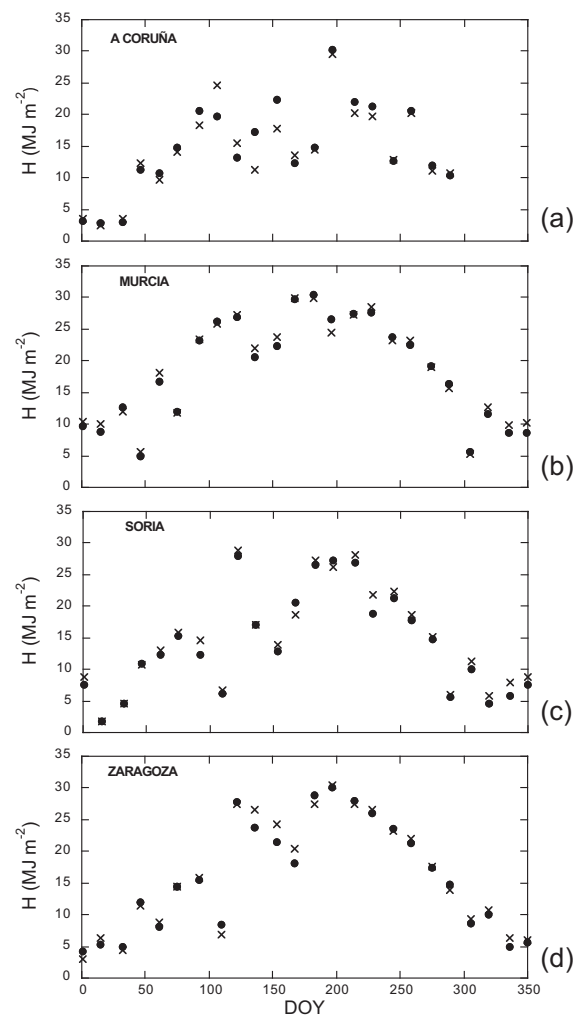


Fig. 3. Measured (×) and estimated (•) irradiation values as a function of the day-of-the-year (DOY) for four different sites (see Table 2).

1-km lat/lon global irradiance images. This procedure removes the sharp borders or discontinuities between adjacent pixels. However, when no surrounding data are available (e.g., in the coastline), the results are not satisfactory. To avoid this problem a different and fast technique to interpolate the images has been developed. It consists of (i) the detection of the coastal pixels by means of a Sobel filter applied to the water/no water mask supplied by the DSSF product, and (ii) the convolution (weighted sum) as a function of the neighboring pixel distance, when data are available. The product is finally refined with the aid of a water/no water mask from the DTM.

Secondly, a topographic correction has been applied. It includes the elevation correction described in Step 2 and further corrections due to the slope and aspect, which determine the angle between the solar rays and the normal to the surface in each pixel. The DEM provides the elevation, slope and aspect for each pixel in the image. In rugged terrains, these parameters affect significantly the global irradiance of the pixel [17,28]. The elevation determines the optical path depth and thus the atmospheric transmittance. Even when a homogeneous atmosphere is assumed, the solar irradiance on a tilted surface depends on its slope and it is more complex than over a horizontal surface due to the effect of configuration factors (the sun's elevation and azimuth) and to diffuse anisotropy over the sky dome [18,29–32]. The global irradiance on a tilted plane is composed of the direct and diffuse irradiances and the ground reflection. It has to be determined for each pixel because the Sun-target orientation-sensor geometry varies from one pixel to another.

Modeling the beam irradiance is trivial once the angle between the solar rays and the normal to the surface is known from the DEM parameters. However, modeling the diffuse irradiance on a tilted surface is more complicated. In this work, the diffuse irradiance on a horizontal surface has been estimated from the global irradiance using a diffuse fraction k_d (diffuse irradiance/global irradiance) obtained from a model based on a sigmoid function that uses an air-mass-corrected clearness index k_t and the relative optical mass m as predictors [27,33]

$$k_d(k_t, m) = 0.944 - 1.538e^A,$$

$$A = 2.808 - 4.759k_t + 2.276k_t^2 + 0.125m + 0.013m^2 \quad (3)$$

This model can estimate the hourly mean of diffuse irradiance under all sky conditions, it presents a good performance when compared against four other regressive models, and it has been validated over Europe and USA using ground data from different sites, including some Spanish stations [33].

Once the diffuse irradiance on a horizontal surface is known, the diffuse irradiance on a tilted plane is calculated using the Hay's model [34]. According to this anisotropic model, the diffuse irradiance on an

Table 2

Statistics of the comparison between ground and estimated irradiation for four sites: A Coruña (lat. 43°18'15", lon. -8°22'42"; elevation 97 m); Murcia (lat. 38°00'10", lon. -1°10'10"; elev. 62 m); Soria (lat. 41°46'10", lon. -2°28'00"; elev. 1082 m); Zaragoza (lat. 41°40'44", lon. -1°04'18"; elev. 260 m). MBD, MAD and RMSD are given in MJ m⁻².

	A Coruña	Murcia	Soria	Zaragoza
MBD	-0.5	-0.3	-0.7	-0.3
MAD	1.5	0.7	1.0	0.9
RMSD	2.2	0.9	1.2	1.2
rMBD	-0.03	-0.02	-0.04	-0.02
rMAD	0.10	0.04	0.06	0.06
rRMSD	0.14	0.06	0.07	0.07
R	0.958	0.995	0.992	0.991

inclined plane is the addition of two components: a circumsolar one for the near-the-sun sky dome and an isotropic-distributed one for the rest of the sky dome. Their relative weight depends on an anisotropy index, defined as the direct beam transmittance [34].

Additionally, the reflected radiation from the surrounding terrain incident on the tilted surface has been estimated considering a mean ground reflectance equal to 0.2. The total irradiance on a tilted surface is finally obtained as the sum of three contributions: the direct irradiance, the diffuse irradiance from the sky dome, and the diffuse irradiance reflected from the surroundings.

A direct validation of these estimates is not possible because the pyranometers are located on horizontal surfaces (regardless of underlying slopes), and hence ground irradiation data only allow for the validation of the daily irradiation from the DSSF product corrected from the elevation effect. Nevertheless, other authors have validated the topographic correction with reasonable results [17,35].

4. Results and discussion

4.1. Step 1: Validation of the H images from the DSSF product over Spain

4.1.1. Validation procedure at station level

Fig. 3 shows the temporal profile (throughout the available dates of year 2008) of *in situ* measured and estimated *H*. Four different stations are shown: one located at the NW, in the most humid region (A Coruña), another one from a semiarid region of the SE (Murcia), and two more located in an intermediate area, which are very close to each other but differ in elevation. The statistics

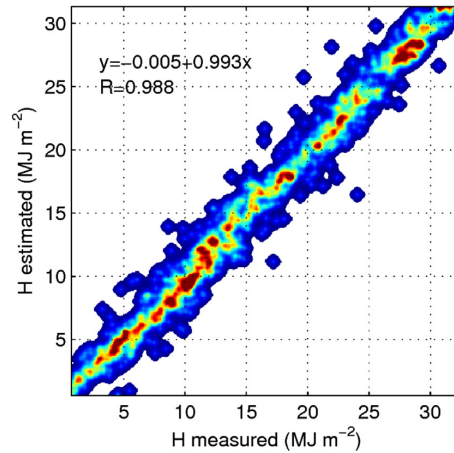


Fig. 5. Correlation between estimated and measured irradiation for all stations.

obtained from the comparison of measured and estimated daily irradiation data for these four stations are shown in Table 2. As observed, the accuracy of the irradiation estimates is higher in the semiarid area (Fig. 3(b)), with considerable lower precipitation, which highlights the influence of the cloud cover in the irradiation estimation from satellite. Nevertheless, even for the most humid region (Fig. 3(a)), the errors are rather low, being below 10%.

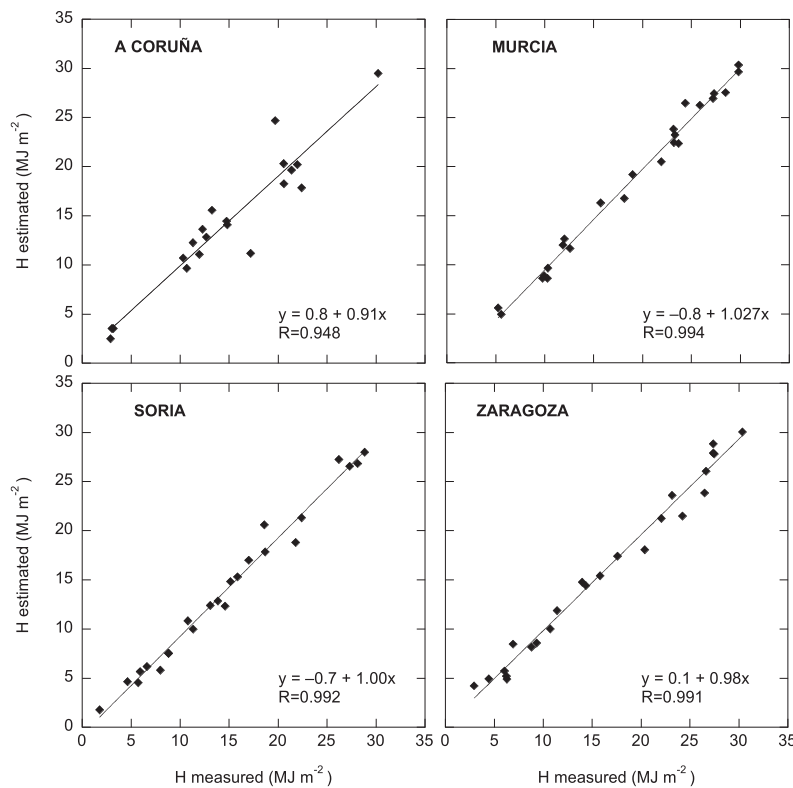


Fig. 4. Correlation between estimated and measured irradiation values for four stations.

Table 3

Statistics obtained from the comparison of measured and estimated daily irradiation for the whole dataset. The best and worst stations are also shown. MBD, MAD and RMSD are given in MJ m^{-2} .

	MBD	MAD	RMSD	rMBD	rMAD	rRMSD	R
All data	-0.12	1.0	1.3	-0.01	0.06	0.08	0.99
Best result	-0.04	0.7	0.8	0.00	0.04	0.05	0.99
Worst result	1.1	1.4	1.9	0.07	0.09	0.12	0.98

Fig. 3(a) shows the capability of satellite estimates to track daily irradiation temporal profile in most of the cases and even in the presence of clouds (minima observed in different dates). The comparison of Fig. 3(c) and (d) evidences the minor influence of elevation in the estimates because their uncertainties are similar in spite of the large difference in elevation of these two neighboring sites.

Fig. 4 shows that the correlation coefficient between estimated and measured H is rather high, although a bit lower for the humid site (A Coruña), where the observed dispersion is higher.

The correlation between estimated and measured H , for the data available from all the stations, is shown in Fig. 5. The data are mainly distributed along the 1:1 line (red color indicates higher density of points): the linear fit has a slope practically equal to one

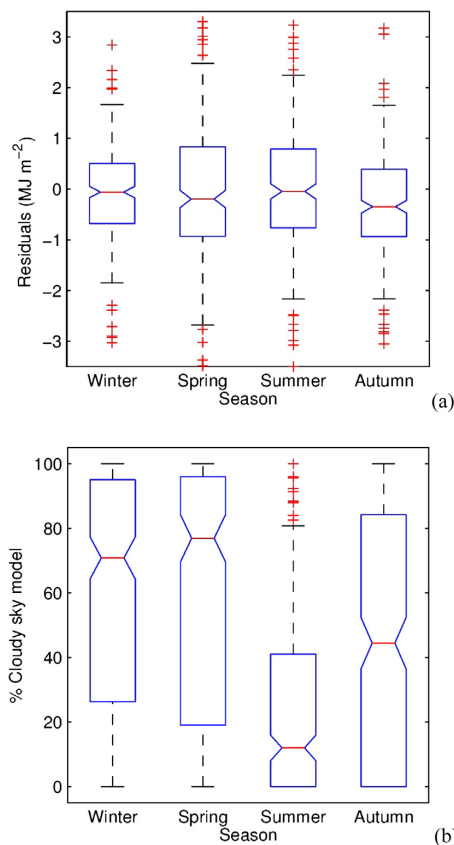


Fig. 6. (a) Seasonal analysis of the residuals ($H_{\text{estimated}} - H_{\text{measured}}$). (b) Percentage of use of the cloudy sky model.

and an intercept near 0, and a very high correlation coefficient is found ($R = 0.988$). Table 3 summarizes the statistics.

4.1.2. Temporal validation

The estimated H product has also been validated over Spain by analyzing the statistics of each date separately. To analyze the temporal dependence of uncertainties and to identify possible seasonal patterns, a box-plot graph of the residuals, i.e., $H_{\text{estimated}} - H_{\text{measured}}$, showing also the percentage of use of the cloudy sky model in each season, is given in Fig. 6. The more pronounced biases are found in spring and autumn, which are the seasons presenting a higher atmospheric instability and a higher precipitation. The median of the residuals is -0.06 , -0.20 , -0.05 , and -0.35 MJ m^{-2} for winter, spring, summer and autumn, respectively. In these periods, the satellite estimated values are obtained from the combination of the algorithms for clear and overcast conditions, respectively. This point is confirmed by Fig. 6(b), which shows the high variance of use of the cloudy sky model in spring and autumn. Underestimated values mainly occur during autumn, although it is not possible to determine whether the bias depends on the sky model used.

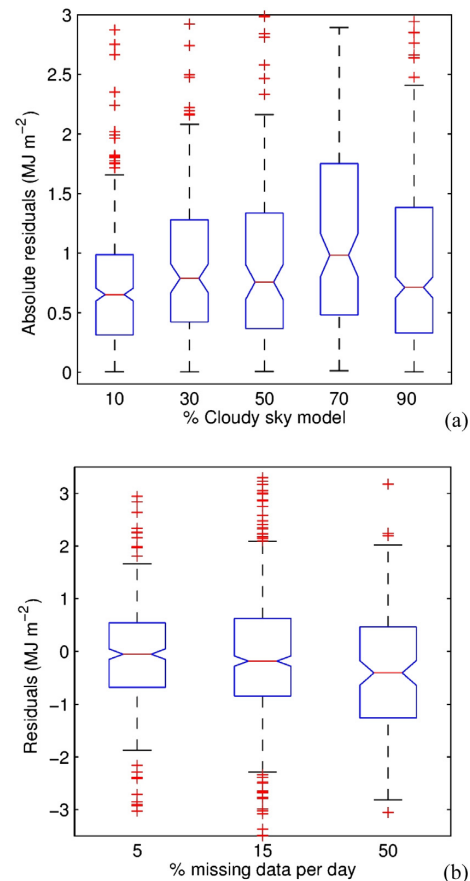


Fig. 7. (a) Impact of the percentage of the use of the cloudy sky model on the absolute residuals ($H_{\text{estimated}} - H_{\text{measured}}$). (b) Impact of the percentage of gaps (missing data) per day on the residuals.

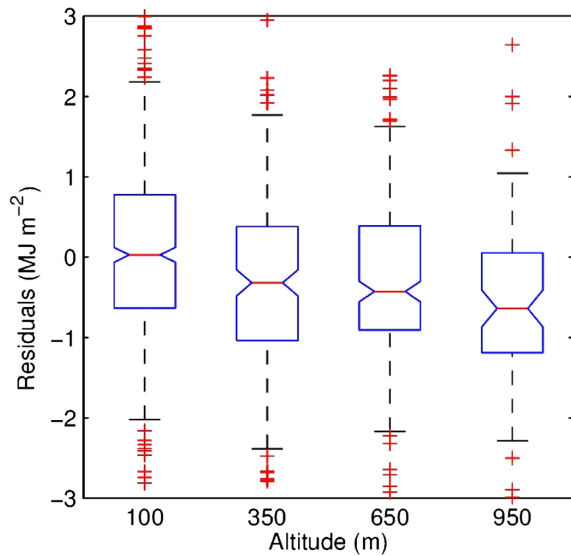


Fig. 8. Impact of the elevation on the residuals. The first box includes information on 23 stations (from 0 m to 200 m), the second 8 (from 200 m to 500 m), the third 11 (from 500 m to 800 m) and the fourth 3 (from 800 m to 1100 m).

4.1.3. Model and gap analysis

The daily irradiation has also been analyzed over Spain in terms of the percentage of use of the cloudy sky model, and in terms of the gap percentage in the daily composition of the DSSF product to obtain the daily irradiation. Fig. 7(a) illustrates that the combined use of cloudy and clear sky models produces higher uncertainties when estimating the daily irradiation. A larger box size indicates an increased residual variance. Fig. 7(b) shows that the estimate bias increases when increasing the number of gaps in the DSSF data series that is integrated each date to derive the global daily irradiation. When the number of gaps increases, the procedure tends to underestimate the irradiation. Thus, the median of the residuals increases in magnitude from approximately 0 for no missing data to -0.4 MJ m^{-2} .

4.1.4. Elevation effect

A DEM at 1-km spatial resolution has been used to evaluate the impact of elevation on the daily irradiation estimation. Fig. 8 shows that the magnitude of the MBD decreases with elevation: median values equal to 0.03, -0.32 , -0.43 , -0.65 MJ m^{-2} for ground stations with elevations ranging from 0 m to 200 m, 200 m to 500 m, 500 m to 800 m, and 800 m to 1100 m, respectively.

These results show that the satellite product tends to underestimate irradiation, probably due to changes in the vertical global

Table 4

Statistics obtained from the comparison of all ground and satellite daily irradiation data and averaged for all the SMC stations, before and after the elevation correction, and for mountain stations (altitude $\geq 500 \text{ m}$), under clear sky conditions. MBD, MAD and RMSD are given in MJ m^{-2} .

		MBD	MAD	RMSD	rMBD	rMAD	rRMSD	R
All stations	No correction	-0.20	1.3	1.8	-0.01	0.08	0.11	0.98
	Elevation correction	0.06	1.3	1.8	0.004	0.08	0.12	0.98
Mountain stations	No correction	-1.1	1.9	2.3	0.05	0.09	0.10	0.95
	Elevation correction	0.10	1.6	2.2	0.005	0.07	0.10	0.95

irradiance profile, especially in its beam irradiance part, as a consequence of the decrease of the optical path length when elevation increases.

4.2. Step 2: Validation of the H images from the DSSF product over Catalonia. Elevation correction

Table 4 shows the results of the validation of the H product using the SMC stations distributed over Catalonia, before and after the elevation correction described in the Methodology section. No relevant discrepancies are observed in the average values (for all stations) shown. However, when considering the mountain stations (elevation $\geq 500 \text{ m}$), the elevation correction is highlighted and the bias, initially around 5%, is reduced to approximately 0% after the correction. The scatterplots in Fig. 9 between measured and estimated daily irradiation illustrate that the elevation correction reduces the observed bias. For these mountains stations, in order to analyze exclusively the elevation effect, only the five clear sky dates were considered.

4.3. Step 3: Topographic correction

As shown in Fig. 2, a topographic correction has been applied to the original DSSF product. Irradiance images at 30 min temporal

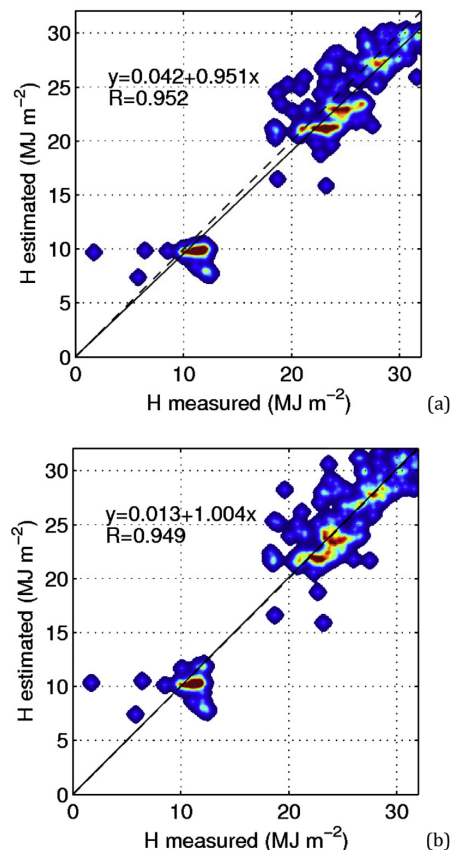


Fig. 9. Linear regression between measured and estimated irradiation for mountain stations (above 500 m) under clear sky conditions, before (a) and after (b) the elevation correction. The regression line (solid) and the 1:1 line (dashed) are shown.

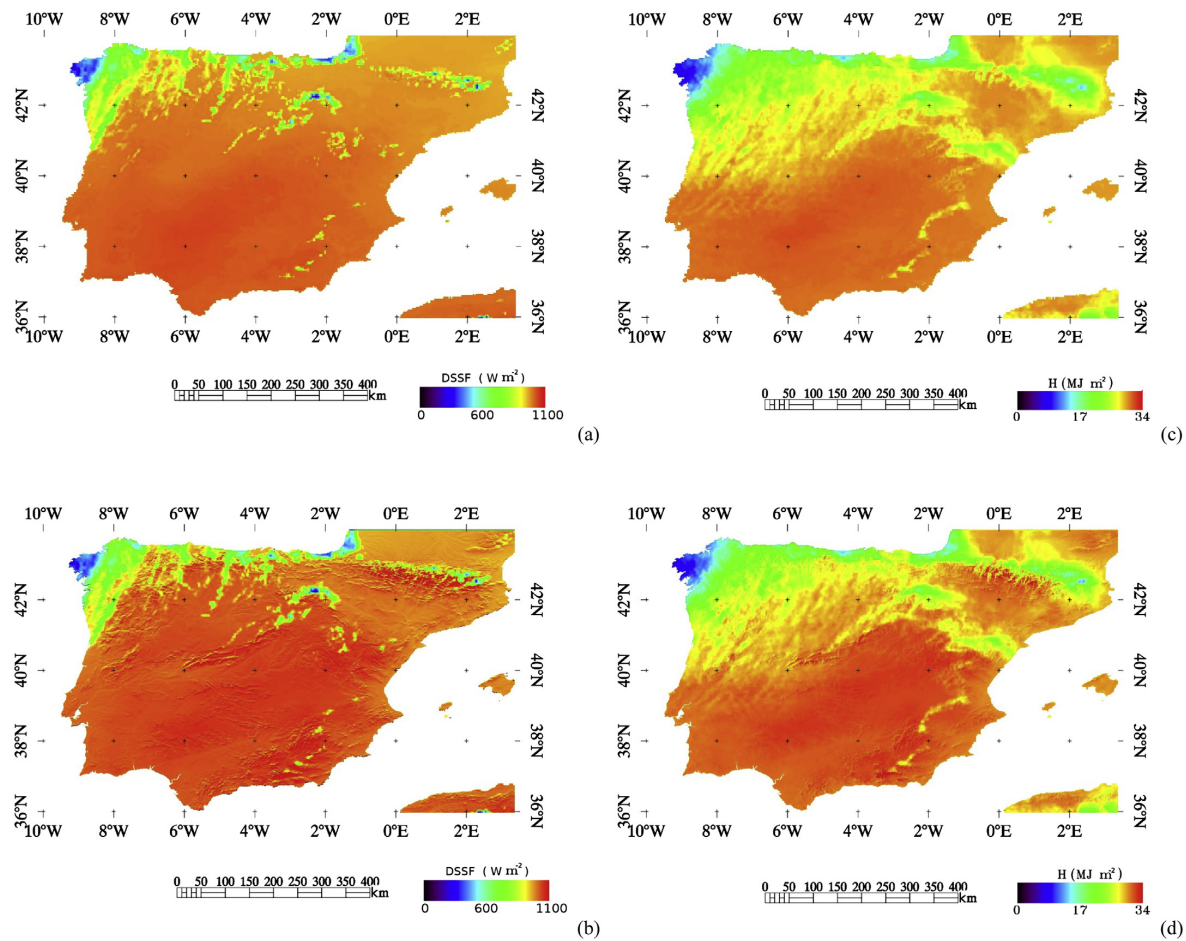


Fig. 10. Estimated irradiance image of July 1st 2008 at 12:00 UTC before (a) and after (b) the topographic correction, and daily irradiation image of this day before (c) and after (d) this correction.

resolution and daily irradiation images are then obtained with 1-km spatial resolution. As an example, Fig. 10 shows the irradiance image of July 1st 2008 at 12:00 UTC, before (a) and after (b) the topographic correction, and the daily irradiation image for that date, also before (c) and after (d) correction. An increase of the spatial variance is noted as a result of this correction, mainly due to effects related to illumination geometry involving the direct and diffuse radiation. This is most clearly observed in the irradiance image. The changing shadows along a day determine that the daily irradiation be less affected by the topography. Nevertheless, under cloud free conditions the topography hardly affects the irradiation, mainly in North and South facing slopes. After the correction, an increase (decrease) of daily irradiation is observed in slopes oriented to South (North).

5. Conclusions

The daily irradiation, computed as the integral of the SEVIRI MSG down-welling surface short-wave radiation flux (DSSF) derived applying the LSA SAF methodology [13], has been validated using *in situ* data of 24 dates uniformly distributed (i.e., every 15 days) along the year 2008. A first validation has been performed over Spain (except the Canary Islands) using 45 ground

stations of the AEMet network. A second one has used 140 stations of the SMC network covering Catalonia, many of them located in rugged terrain, in order to further analyze the influence of the topography in the irradiance estimation from satellite images. An elevation correction has been applied to the DSSF product and the resulting daily irradiation has been validated using the SMC network. Different error sources have been taken into account to discuss the results. Finally, a downscaling of the DSSF product is carried out jointly with a topographic correction (including elevation, slope and aspect effects) to obtain improved daily irradiation images.

The resulting statistics from the first validation show a bias of -0.12 MJ m^{-2} (rMBD of about 1%) and a mean absolute difference of 1.0 MJ m^{-2} (rMAD of 6%). These results improve the MAD of 2.3 MJ m^{-2} that was obtained in a previous paper using the same database to validate daily irradiation images derived from temperature and precipitation data using a neural network [12].

Regarding the influence of the cloud cover, Fig. 7a shows that the residuals are lower when considering clear skies (which also show the lower variance) and complete cloudy skies. However, in the presence of partially cloudy skies, the residuals increase. This is partially due to problems with the cloud mask used in the LSA SAF

procedure, and also to the higher complexity of the interaction of radiation with partially cloudy skies compared to clear or overcast skies. A further study could be recommended. In relation to the missing data influence, the residuals in Fig. 7b show that there is a tendency to underestimate the daily irradiation when the number of gaps in the data to produce this daily product is above 20%. Thus, the absolute value of the residual changes from around 0.2 (5% gaps) to 0.4 (50% gaps).

The residual analysis at different elevations shows that the daily irradiation images from the DSSF product tend to underestimate the daily irradiation when elevation increases. For example, the median of the residuals ranges between 0.0 and -0.6 when elevation ranges from 100 m to 1 km. This is due to the lack of elevation correction in the images and therefore to the atmospheric transmittance increase with altitude. These results justify the elevation correction applied to the original DSSF product, which has been validated using the Catalanian network. This correction produces daily irradiation estimates with lower biases, without affecting the linearity between measured and estimated values. Only data under clear sky conditions (the beam irradiance is many times the diffuse irradiance) have been analyzed in this part of the study, and special attention has been paid to mountain stations (elevation ≥ 500 m). The elevation correction leads to a ten-fold reduction of the bias, whereas the other statistical indices improve only slightly or remain unchanged. The linear correlation between estimated and measured daily irradiation shows a slope around one and an intercept more close to 0 than before the correction.

The daily irradiation from the elevation-corrected DSSF has been validated using ground data from the SMC database. For clear sky conditions, and for the complete sampling, the MAD is 1.3 MJ m^{-2} . When considering only the mountain stations, the MBD between estimated and measured daily irradiation is reduced from -1.1 MJ m^{-2} to 0.1 MJ m^{-2} (from 5% to 0.5% in relative values) after the elevation correction, whereas the MAD values are reduced from 1.9 MJ m^{-2} to 1.6 MJ m^{-2} (from 9% to 7% in relative values) before and after the correction, respectively.

Regarding the last step of the methodology, which includes a downscaling of the image, a coastline correction and a topographic correction, no quantitative assessment can be provided because the ground pyranometers are located on horizontal surfaces. Nevertheless, the qualitative analysis of the resulting images reveals the role of landscape complexity in the spatial distribution of radiation fluxes. After the topographic correction, it was shown that the spatial variability in irradiance was significantly increased with greater complexity (e.g., see the Pyrenees, the mountains in central Spain, and in the South-East area in Fig. 10). Thus, detailed features are found in these rugged terrains, showing the contrast between the illuminated (South facing) and shaded (North facing) slopes. In particular, in such a South-East area, Ruiz-Arias et al. [17] showed that the rMBD diminished from 2.3% to 0.4% after the terrain correction. The topographic correction is especially relevant when the spatial resolution increases (as in the downscaling procedure) since topography is one of the main sources of the spatial variability of the solar irradiance. Therefore, in a territory like our study area, with rugged terrain areas and with elevation values from 0 to around 3500 m, the topographic correction might lead to significant improvement of both DSSF and daily irradiation products, in terms of qualitatively bringing it closer to the expected spatial structure.

Finally, it is worth mentioning the small computation cost of the corrected irradiation images. The procedure described in this manuscript has been implemented by means of a highly efficient algorithm that allows to process all the corrections over 880,000 pixels per image in about 2 s (Intel(R) Core(TM) i5 CPU 650@3.20 GHz and 8 GB of RAM).

Acknowledgments

This work was supported by ÁRTEMIS (CGL2008–00381), RESET CLIMATE (CGL2012–35831) and LSA SAF (Eumetsat) projects. The ground data were kindly supplied by AEMet (Agencia Estatal de Meteorología, Spain) and SMC (*Servei Meteorològic de Catalunya*). We thank to Dr. J. Tamayo from AEMet for his assistance and to our LSA SAF colleagues for providing the DSSF products. The comments of the anonymous reviewers were also very helpful.

References

- Gueymard CA. Direct and indirect uncertainties in the prediction of tilted irradiance for solar engineering applications. *Solar Energy* 2009;83:432–44.
- Sayigh A. Worldwide progress in renewable energy. *Renewable Energy* 2009. <http://dx.doi.org/10.1016/j.renene.2008.12.025>.
- Loutzenhiser PG, Manz H, Felsmann C, Strachan PA, Frank T, Maxwell GM. Empirical validation of models to compute solar irradiance on inclined surfaces for building energy simulation. *Solar Energy* 2007;81:254–67.
- Monteith JL. Solar radiation and productivity in tropical ecosystems. *Journal of Applied Ecology* 1972;9:747–66.
- Iqbal M. An introduction to solar radiation. Toronto: Academic Press; 1983. 390 pp.
- Mubiru J, Banda EJK. Monthly average daily global solar irradiation maps for Uganda: a location in the equatorial region. *Renewable Energy* 2012;41:412–5.
- Hay JE, Hanson KJ. Evaluating the solar resource: a review of problems resulting from temporal, spatial and angular variations. *Solar Energy* 1985;34:151–61.
- Hugues JP, Lettenmaier DP. Data requirements for kriging: estimation and network design. *Water Resources Research* 1981;17:1641–50.
- Mellit A, Benganem M, Hadj Arab A, Guessoum A. A simplified model for generating sequences of global solar radiation data for isolated sites: using artificial neural network and a library of Markov transition matrices approach. *Solar Energy* 2005;79:469–82.
- Robaa SM. Validation of the existing models for estimating global solar radiation over Egypt. *Energy Conversion and Management* 2009;50:184–93.
- Tymvios FS, Jacovides CP, Michaelides SC, Scouteli C. Comparative study of Ångström's artificial neural networks' methodologies in estimating global solar radiation. *Solar Energy* 2005;78:752–62.
- Moreno A, Gilabert MA, Martínez B. Mapping daily global solar irradiation over Spain: a comparative study of selected approaches. *Solar Energy* 2011;85:2072–84.
- Geiger B, Meurey C, Lajas D, Franchistéguy L, Carrer D, Roujean J-L. Near real-time provision of downwelling shortwave radiation estimates derived from satellite observations. *Meteorological Applications* 2008;15:411–20.
- Moradi I, Mueller R, Alijani B, Kamali GA. Evaluation of the Heliosat-II method using daily irradiation data for four stations in Iran. *Solar Energy* 2009;83:150–6.
- Janjai S, Pankaew P, Laksanaboonsong J, Kitichantaropas P. Estimation of solar radiation over Cambodia from long-term satellite data. *Renewable Energy* 2009;36:1214–20.
- Manners J, Vosper SB, Roberts N. Radiative transfer over resolved topographic features for high-resolution weather prediction. *Quarterly Journal of the Royal Meteorological Society* 2012;138:720–33.
- Ruiz-Arias JA, Cebecauer T, Tovar-Pescador J, Sùri M. Spatial disaggregation of satellite-derived irradiance using a high-resolution digital elevation model. *Solar Energy* 2010;84:1644–57.
- Oliphant A, Spronken-Smith R. Spatial variability of surface radiation fluxes in mountainous terrain. *Journal of Applied Meteorology* 2003;42:113–28.
- Bosch JL, Batles FJ, Zarzalejo LF, López G. Solar resources estimation combining digital terrain models and satellite images techniques. *Renewable Energy* 2010;35:2853–61.
- Cano D, Monget JM, Albuissou M, Guillard H, Regas N, Wald L. A method for the determination of the global solar radiation from meteorological satellite data. *Solar Energy* 1986;37:31–9.
- Rigollier C, Lefevre M, Wald L. The method Heliosat-2 for deriving shortwave solar radiation from satellite images. *Solar Energy* 2004;77:159–69.
- Aminou DMA. MSG's SEVIRI instrument. *ESA Bulletin* 2002;111:15–7.
- NWC SAF product user manual for "Cloud Products", PGE01-02-03 of the SAFNW/MSG. Available at, <http://www.nwcsaf.org>; 2005. February 2012.
- Geiger B, Carrer D, Franchistéguy L, Roujean J-L, Meurey C. Land surface albedo derived on a daily basis from Meteosat Second Generation observations. *IEEE Transactions on Geoscience and Remote Sensing* 2008;46:3841–56.
- Journée M, Bertrand C. Improving the spatio-temporal distribution of surface solar radiation data by merging ground and satellite measurements. *Remote Sensing of Environment* 2010;114:2692–704.
- LSA SAF product user manual "Down-welling Surface Shortwave Flux (DSSF)", version 2.6. Available at, <http://landsaf.meteo.pt/algorithms.jsp?seltab=1&starttab=1>; July 2011. November 2011.
- Pérez R, Ineichen P, Seals R, Zelenka A. Making full use of the clearness index for parameterizing hourly insolation conditions. *Solar Energy* 1990;45:111–4.

- [28] Haurant P, Muselli M, Pillot B, Oberti P. Disaggregation of satellite derived irradiance maps: evaluation of the process and application to Corsica. *Solar Energy* 2012;86:3168–82.
- [29] Notton G, Paoli C, Ivanova L, Vasilea S, Nivet ML. Neural network approach to estimate 10-min solar global irradiation values on tilted planes. *Renewable Energy* 2013;50:576–84.
- [30] David M, Lauret P, Boland J. Evaluating tilted plane models for solar radiation using comprehensive testing procedures, at a southern hemisphere location. *Renewable Energy* 2013;51:124–31.
- [31] Conese C, Gilabert MA, Maselli F. Topographic normalization of TM scenes through the use of an atmospheric correction method and digital terrain models. *Photogrammetric Engineering and Remote Sensing* 1993;59:1745–53.
- [32] Dubayah R, Dozier J, Davis FW. Topographic distribution of clear-sky radiation over the Konza Prairie, Kansas. *Water Resources Research* 1990;26:679–90.
- [33] Ruiz-Arias JA, Alsamamra H, Tovar-Pescador J, Pozo-Vázquez D. Proposal of a regressive model for the hourly diffuse solar radiation under all sky conditions. *Energy Conversion and Management* 2010;51:881–93.
- [34] Hay JE. Calculation of monthly mean solar radiation on horizontal and inclined surfaces. *Solar Energy* 1979;23:301–7.
- [35] Dubayah R. Modeling topographic solar radiation using GOES data. *Journal of Applied Meteorology* 1997;36:141–54.

Apéndice **C**

Publication III: Assessment of MODIS imagery to track light-use efficiency in a water-limited Mediterranean pine forest



Contents lists available at SciVerse ScienceDirect

Remote Sensing of Environment

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

Assessment of MODIS imagery to track light-use efficiency in a water-limited Mediterranean pine forest

A. Moreno^a, F. Maselli^b, M.A. Gilabert^{a,*}, M. Chiesi^b, B. Martínez^a, G. Seufert^c

^a Departament de Física de la Terra i Termodinàmica, Facultat de Física, Universitat de València, Dr. Moliner, 50. 46100-Burjassot, Spain

^b IBIMET-CNR, Via Madonna del Piano, 10. 50019-Sesto Fiorentino (FI), Italy

^c Institute for Environmental Sustainability, JRC, Via Fermi, 2749-TP 050. 21027-Ispra (VA), Italy

ARTICLE INFO

Article history:

Received 20 April 2011

Received in revised form 29 March 2012

Accepted 1 April 2012

Available online 3 May 2012

Keywords:

PRI

MODIS

Light-use efficiency

Mediterranean forest

Water stress

ABSTRACT

Daily values of gross primary production (*GPP*) derived from an eddy-covariance flux tower have been used to analyze the information content of the MODIS Photochemical Reflectance Index (*PRI*) on the light-use efficiency (ϵ). The study has been conducted in a Mediterranean *Pinus pinaster* forest showing summer water stress. Advanced processing techniques have been used to analyze the effect of various external factors on ϵ and *PRI* temporal variations. The intra-annual correlation between these two variables has been found to be mostly attributable to concurrent variations in sun and view zenith angles. The *PRI* has been normalized from these angular effects (*NPRI*), and its ability to track ecosystem ϵ response to prolonged summer water limitations has been analyzed. The observed shift between ϵ and *NPRI* reveals that, for the study area and at MODIS spatial resolution, *NPRI* is informative on changes in pigments and canopy structure related to the vegetation response to prolonged water stress.

© 2012 Elsevier Inc. All rights reserved.

1. Introduction

1.1. Theoretical background

Monitoring forest ecosystem functioning on various spatial and temporal scales is essential to assess the forest role as carbon sink (IPCC, 2007; Waring & Running, 2007), to evaluate the forest capacity of producing woody biomass (Corona & Marchetti, 2007), and to guide forest management practices (FAO, 2005).

The gross primary production (*GPP*) is an essential parameter to estimate carbon fluxes that corresponds to total ecosystem photosynthesis. Several studies have demonstrated that remotely-sensed data can be combined with ancillary data for the estimation of gross forest production (Field et al., 1995; Maselli et al., 2006; Veroustraete et al., 2002). Forest *GPP* can be estimated by combining conventional meteorological measurements (temperature, radiation) with remotely-sensed estimates of vegetation conditions (Veroustraete et al., 2004; Maselli et al., 2009, 2010).

Monteith's (1972) approach provides the theoretical basis for most models of forest production based on remote sensing, such as CASA, MODIS-GPP, and C-Fix (Heinsch et al., 2003; Potter et al., 1993; Veroustraete et al., 2002). These models compute *GPP* from three variables: (i) the incoming photosynthetic active radiation (*PAR*), (ii) the fraction of photosynthetic radiation absorbed by vegetation (*fAPAR*),

which is related to vegetation canopy functioning and energy absorption capacity, and (iii) the light-use efficiency (ϵ) in the conversion of absorbed *PAR* into photosynthetically fixed CO_2 :

$$GPP = \epsilon fAPAR PAR. \quad (1)$$

While *PAR* is a meteorological variable, *fAPAR* and ϵ are descriptive of vegetation properties and can be obtained from remote sensing. The availability of MODIS is potentially useful to estimate *fAPAR* and ϵ with medium–low spatial resolution and high temporal frequency.

fAPAR is linked to canopy structure and condition (i.e. to green biomass) and has a clear seasonal evolution in deciduous species, while it shows limited variability in evergreen forest ecosystems. Multiple global *fAPAR* datasets are available from several initiatives: Boston University (MODIS) (Knyazikhin et al., 1999); MEDIAS-France (CYCLOPES) (Baret et al., 2007); Joint Research Center (JRC) (Gobron et al., 2004), the European Space Agency (GLOBECARBON) (Plummer et al., 2006), and LSA SAF (Land Surface Analysis Satellite Application Facility) (García-Haro et al., 2008). However, the discrepancies among them can be significant and there is no consensus about the more appropriate algorithm to develop this biophysical product. The comparison studies have indicated large differences in magnitude among the *fAPAR* products, although rather consistent seasonality (Camacho, 2007; Gobron et al., 2007; McCallum et al., 2010; Seixas et al., 2009; Weiss et al., 2007). A different approach is proposed by Zhang et al. (2005, 2009). These authors retrieve the *fAPAR* by the part of the canopy dominated by chlorophyll (*fAPAR_{chl}*) by means of a coupled canopy–leaf radiative transfer model and MODIS images. The *fAPAR_{chl}* may characterize better

* Corresponding author. Tel.: +34 96 3543118; fax: +34 96 3543385.
E-mail address: m.amparo.gilabert@uv.es (M.A. Gilabert).

the PAR absorbed by chlorophyll used for vegetation photosynthesis, and it is an alternative to be considered in future research.

The light-use efficiency ε has been shown to vary spatially between biomes, ecosystems, and plant species, and to vary temporally during the growing season, due to environmental and physiological limitations (Coops et al., 2010; Field et al., 1995; Garbulsy et al., 2010; Turner et al., 2003). ε responds more rapidly than $fAPAR$ to different environmental factors related to energy balance, water availability and nutrient levels (Connolly et al., 2009; Garbulsy et al., 2010). For operational applications, ε is generally modeled as the product of a maximum (ε_{max}) – which corresponds to optimal conditions and depends on the vegetation type – and a down-regulating factor that accounts for inter-annual variations due to environmental, short-term stresses (temperature stress, water deficit, etc.). The computation of this last factor frequently requires the use of meteorological data, which are not always available at the needed spatial and temporal scales. For example, the MODIS estimation of this factor is based on the daily minimum temperature and the daily average vapor pressure deficit (Heinsch et al., 2003). On the other hand, the modified C-Fix model (Veroustraete et al., 2002; Veroustraete et al., 2004) proposed by Maselli et al. (2009) for Mediterranean areas uses a water stress index, C_{ws} , to account for limited photosynthetic activity in case of short-term water stress. C_{ws} is obtained from a local water budget based on actual (AET) and potential (PET) evapotranspiration estimated over a two-month period:

$$C_{ws} = 0.5 + 0.5AET/PET. \quad (2)$$

1.2. The Photochemical Reflectance Index (PRI)

In theory, the canopy ε can be evaluated from remote sensing using short and thermal wavelengths (Damm et al., 2010; Sims et al., 2008). However, the capability of remote sensing to capture the full range of variability across biomes, particularly due to short-term stress, remains unclear. In recent years, the spatial and temporal variations of ε have been analyzed through the remote sensing of the xanthophyll plant pigment using a narrow-band spectral index, the Photochemical Reflectance Index (Gamon et al., 1992)

$$PRI = (R_{531} - R_{ref}) / (R_{531} + R_{ref}), \quad (3)$$

where R_{531} is the reflectance in a narrow band centered at 531 nm, which is related to xanthophyll biochemical reactions (de-epoxidation of violaxanthin via antheraxanthin to zeaxanthin) that take place to protect the photosynthetic apparatus (Gamon et al., 1990; Gamon et al., 1992). R_{ref} is the reflectance in a reference band insensitive to the short-term changes caused by xanthophyll interconversions (such as the 570 nm band). PRI is claimed to be a photosynthetic efficiency indicator across a wide range of conditions, species and functional types, because it also correlates with chlorophyll/carotenoid ratios (Filella et al., 2009) and can be used to monitor early water stress (Naumann et al., 2010; Sarlikioti et al., 2010; Suárez et al., 2009; Thenot et al., 2002). Different studies have used MODIS data to calculate PRI (Drolet et al., 2005; Drolet et al., 2008; Garbulsy et al., 2008) because of its band 11 centered at 531 nm and because of its daily revisiting cycle. Due to the lack of the 570 nm band to evaluate R_{ref} , bands 10 (483–493 nm), 12 (546–556 nm), 13 (662–672 nm) or 1 (620–670 nm) must be used as reference (Coops et al., 2010). At present, there is no consensus on the best wavelengths to calculate PRI because multiple scattering can affect the spectral response of the xanthophyll cycle feature at these spatial scales.

Concerning the potential of MODIS PRI to evaluate ε variability, research has been focused on both spatial and temporal variations (Coops et al., 2010; Drolet et al., 2008; Garbulsy et al., 2011). In general, no conclusive result has been obtained due to spatial (geographic)

and/or ecosystem variations. The statistical associations found are generally weak and variable (Drolet et al., 2008).

In relation to temporal changes, several studies show the successes of PRI to track ε variations at leaf and individual canopy scales. Canopy PRI is generally lower than the leaf PRI since it is the sum of contributions from all the scene components (including illuminated and shaded leaves, and soil background) weighted by the strength of their signal to the sensor (thus depending on the leaf area index and on the leaf angle distribution) (Barton & North, 2001). Strong statistical associations have also been found at an ecosystem (MODIS pixel) level between the annual evolution of ε and PRI (Drolet et al., 2008; Garbulsy et al., 2011; Goerner et al., 2009). For example, Garbulsy et al. (2008) have shown that the coefficient of determination (R^2) between PRI and ε for Mediterranean evergreen forests is about 0.5–0.7.

In contrast, no study has clearly demonstrated the PRI potential to track inter-annual ε changes. In summer ε decreases due to extreme water stress (Garbulsy et al., 2008; Goerner et al., 2011) and the correlation between PRI and ε remains unclear at MODIS pixel scale.

Unlike broad-band vegetation indices, which are mainly related to the red-edge reflectance of the vegetation, the PRI has a very low signal and a rather high noise level. It is worth to note that the reflectance change in the R_{531} band expected from xanthophylls conversion is only about 0.004, i.e., 7% of the total reflectance (Coops et al., 2010; Hall et al., 2008), whereas the noise level in this band is about 3% of the signal (Xiong et al., 2003). Therefore, PRI is difficult to recover in the presence of geometrical or atmospheric factors (Barton & North, 2001; Cheng et al., 2010; Coops et al., 2010; Drolet et al., 2005; Garbulsy et al., 2011) that can cause changes in spectral reflectance of similar or larger magnitudes. In particular, the sun/sensor geometry determines the fraction of shadowed and sunlit canopy that affects the reflectance in the PRI bands: a given canopy ε could have different PRI s depending on the amount of illuminated leaves seen by the sensor (Cheng et al., 2010; Hilker et al., 2008; 2010). Moreover, the differences in the area of observation (i.e., the pixel footprint) between the MODIS and the tower data within different satellite overpasses can cause significant uncertainties and noise in the PRI - ε relationship (Hilker et al., 2009).

As suggested by Mänd et al. (2010), the question remains how much of the variation in MODIS PRI is due to external factors, how much is due to changes in photosynthetic efficiency due to xanthophyll cycle, how much is due to changing pigment ratios (e.g. carotenoid and chlorophyll levels) and how much is due to changes in canopy structure.

The objective of this work is to assess the potential of MODIS data to track ε temporal variability. The analysis focuses on the capacity of the MODIS PRI to detect inter-annual ε changes due to summer water stress. To this aim, advanced processing techniques are applied to MODIS observations and to GPP data from a flux tower located in San Rossore (Central Italy) over five years (2001–2005).

2. Study area and data

2.1. Study area

The test site is a pine forest located within the Regional Park of San Rossore (43°43'41" N, 10°17'07" E). It is a protected area limited by the Tyrrhenian Sea on the West and the rivers Arno and Serchio on the South and North, respectively. Its sub-humid, Mediterranean climate shows an average annual temperature of 14.8 °C and an average rainfall of 900 mm (Rapetti & Vittorini, 1995). The water stress period usually lasts for about two months and is centered on the summer season. The soils of the area are prevalently sandy (D.R.E.Am, 2003).

2.2. Flux tower data

The study site belongs to the FLUXNET project (<http://www.fluxnet.ornl.gov/fluxnet/>), a network coordinating regional and global analyses

of observations from micrometeorological tower sites. The flux tower was installed by the end of 1998 at about 700 m from the seashore, in the center of a maritime pine plantation, established after a large fire about 60 years ago. The plantation extends for about 300 ha along the Tyrrhenian coast inside the 23,000 ha regional park and UNESCO Biosphere Reserve San Rossore. The described great size and homogeneity of the pine plantation around the tower guarantee quasi-optimal conditions for eddy covariance flux measurements. A footprint analysis indicates that 70% of the measured fluxes come from less than 200 m around the tower in unstable conditions during daytime and from less than 500 m in stable conditions during nighttime (Goeckede et al., 2008).

The average stand height is 18 m, the diameter at breast height of *Pinus pinaster* trees is 29 cm, the tree density is about 565 ha^{-1} (84% *P. pinaster*, 12% *Pinus pinea* and 4% *Quercus ilex*), the basal area of $39 \text{ m}^2 \text{ ha}^{-1}$, and the leaf area index is about 3 (Chiesi et al., 2005; Zenone et al., 2008), with practically no seasonal variation. The stand has no significant understorey.

The daily *NEE* (net ecosystem exchange) data in the five-year period 2001–2005 have been processed according to current state of the art methodology (Papale et al., 2006). Quality analysis has been performed according to the stationarity and integral turbulence tests (Foken & Wichura, 1996). Gap filling and flux partitioning have been applied to derive ecosystem respiration and *GPP*, according to Reichstein et al. (2005).

2.3. Meteorological data

Daily meteorological measurements (mean temperature, rainfall) were collected at the tower during the study period. These data, however, presented several gaps and could not be used in this work. Meteorological data were then derived from the interpolation, using the DAYMET algorithm (Thornton et al., 2000), of measurements taken at adjacent weather stations. Solar radiation was estimated using the MT-CLIM algorithm (Thornton et al., 1997). All these weather data were recalibrated using the procedure described by Rodolfi et al. (2007).

2.4. Satellite data

Table 1 shows MODIS products from the TERRA and AQUA platforms as downloaded from the Warehouse Inventory Search Tool, WIST (<http://wist.echo.nasa.gov>). The MODIS cloud mask (MOD35), at 250 and 1000 m resolutions, was used to select data with no clouds or optically thick aerosols.

3. Methodology

MODIS *fAPAR* and *PRI* products were processed as follows, and the corresponding time series from the tower pixel were extracted.

Table 1
MODIS products used in the study.

MODIS product	Variable
MOD15A2	<i>fAPAR</i> every 8 days
MCD43A1	BRDF (k_0 , k_1 and k_2)
MCD43A2	Quality flags (QF)
MOD35	Cloud mask
MOD021km	<i>PRI</i> (b11, b12) ^a
MOD03	Geolocation data

^a MOD02 contains calibrated and geolocated at-aperture radiances for 36 MODIS bands. Reflectance has been determined for bands 11 and 12.

3.1. *fAPAR*

As mentioned in the introductory section, there are different available operational *fAPAR* products. In particular, the MODIS product (MOD15A2) has shown an overestimation of the *fAPAR* in different ecosystems (Fensholt et al., 2004; Seixas et al., 2009), probably due to its land cover input (McCallum et al., 2010). As an alternative to the radiative transfer model inversion (Knyazikhin et al., 1999; Myneni et al., 2002), *fAPAR* can also be derived from MODIS using vegetation indices such as *NDVI* (Normalized Difference Vegetation Index) (Myneni & Williams, 1994) or *RDVI* (Renormalized Difference Vegetation Index) (Roujean & Bréon, 1995). The *RDVI* is actually being used to derive *fAPAR* from MSG-SEVIRI images in the context of LSA SAF (García-Haro et al., 2008). This *fAPAR* product has shown a rather good performance when validated with ground data (Camacho, 2007).

In this paper, *fAPAR* values were calculated as

$$fAPAR = 1.81RDVI - 0.21 \quad (4)$$

where *RDVI* is

$$RDVI = (R_{NIR} - R_R) / (R_{NIR} + R_R)^{1/2} \quad (5)$$

and R_{NIR} and R_R are the near infrared and red reflectances, respectively, for an optimal geometry ($\theta_s = 45^\circ$, $\theta_v = 60^\circ$, $\varphi = 0^\circ$) to reduce the influence of external factors (Roujean & Bréon, 1995). They have been obtained using the *BRDF* parameters (k_0 , k_1 and k_2) from the MCD43A1 product. These parameters are the spectrally dependent weights of the kernels associated to isotropic, geometric and volumetric scattering, respectively, in the kernel-driven *BRDF* model.

3.2. *PRI* and *NPRI*

For the tower pixel, *PRI* was calculated according to Eq. (3) using the MODIS bands 11 (526–536 nm) and 12 (546–556 nm), as suggested by Wu et al. (2010). This *PRI* shows good correlation with measured ε (Hilker et al., 2009).

The MOD021km dataset used to compute *PRI* is not corrected for atmospheric effects, illumination and observation geometries and clouds. However this is not a significant drawback in the study because bands 11 and 12 are close and suffer from similar atmospheric effects. In fact, higher correlations between *PRI* and ε have been found when using top-of-atmosphere reflectances instead of surface reflectances (Drolet et al., 2005; Drolet et al., 2008; Goerner et al., 2009).

As mentioned above, the correlation between *PRI* and ε can be highly influenced by the illumination and viewing geometries, which determine the intensity of light received by the leaf and the proportion of shadow in the scene. Thus, the reflectance anisotropy can affect the *PRI* signal even in the absence of significant changes in the vegetation activity. Moreover, noise in the *PRI* signal can also occur due to differences in the area of observation (i.e., the pixel footprint) between the MODIS and the tower data within different satellite overpasses (with varying view zenith angles). The MODIS observation footprint for the required bands is about 1 km^2 when the sensor view zenith angle is lower than 40° (Drolet et al., 2005; Goerner et al., 2011; Wolfe et al., 1998). For this reason, the scenes with a view zenith angle larger than 40° were rejected to eliminate pixels with a reduced spatial resolution, to reduce the effect of anisotropy in surface reflectance, and to minimize footprint problems.

In order to quantify the influence of illumination and view geometry on seasonal *PRI* variation, an elastic-net variable selection procedure (Zou & Hastie, 2005) was applied.

The elastic net is used here to find a relation (aka a model) between the MODIS *PRI* and the four geometric variables: (i) solar zenith angle,

SZA; (ii) view zenith angle, VZA; (iii) the cosine of the relative azimuth angle (between illumination and observation), $\cos RAA$; and (iv) its absolute value, $|\cos RAA|$. The $\cos RAA$ shows the differences between forward and backscatter illuminations, whereas $|\cos RAA|$ only varies when the sensor is not in the principal plane. Information related to geodetic coordinates, and solar and satellite zenith and azimuth angles for the tower pixel at each MODIS scene was extracted from the MOD03 product (see Table 1).

The elastic net combines the characteristics of lasso and ridge regression (Hoerl & Kennard, 1970; Tibshirani, 1996). Ridge regression is known to shrink the coefficients of correlated predictors towards each other. Lasso regression is somewhat indifferent to very correlated predictors, and will tend to pick one and ignore the rest. This elastic net technique often outperforms the lasso regression (Zou & Hastie, 2005), while enjoying a similar sparse representation. In addition, the elastic net technique encourages a grouping effect, where strongly correlated input variables tend to be in or out of the model together. The regularization is carried out using a linear combination (with tuning parameter α) of the penalties $|\beta|_1 = \sum_{j=1}^p |\beta_j|$ of ridge regression and $\|\beta\|^2 = \sum_{j=1}^p \beta_j^2$ of lasso regression and a global regularization parameter θ . In the above expressions, $p = 4$ is the number of input (geometric) variables. Thus, the expression to be minimized is

$$\|PRI_{modelled} - X\beta\|^2 + \theta [\alpha|\beta|_1 + (1-\alpha)\|\beta\|^2] \quad (6)$$

where $PRI_{modelled}$ is a vector of length n (the number of training samples) containing the PRI outputs, X is an $n \times p$ matrix with the input variables, and β is a vector of length p containing the regression coefficients to be optimized.

The differences between the measured PRI values and those calculated from the model explained above, $PRI_{modelled} = X\beta$, are the model residuals, and we refer to them as

$$NPRI = PRI_{measured} - PRI_{modelled} \quad (7)$$

3.3. Time series filtering and gap filling

The $fAPAR$ time series has been filtered using the local method LOESS (aka LOWESS) in order to remove undesirable day-to-day variability (noise) resulting from cloud, ozone, dust, and other aerosols that generally decrease the near-infrared reflectance and leads to spurious drops in the data (Goward et al., 1991; Holben, 1986). Off-nadir viewing and low sun zenith angles can also cause a similar effect (Gutman, 1991). Linear filters such as LOESS perform well for noise reduction, but can distort series where strong gradients are present such as the transition period from snow to snow-free conditions (Zhao & Fernandes, 2009). In order to minimize this problem, the input quality flag information was used to neglect cloudy or snow covered pixels.

Our method is based on two assumptions: (i) The $fAPAR$ is primarily related to vegetation activity. A $fAPAR$ time series follows a smooth annual cycle of growth and decline, which is considered as a low-frequency component of the signal. (ii) Clouds and poor atmospheric conditions usually lower $fAPAR$ values, and induce sudden drops in the series that are not compatible with the gradual process of vegetation activity. These drops are considered as a high-frequency component regarded as noise to be removed.

The LOESS method used involves fitting a low-degree polynomial to a subset of the time series (Atkeson et al., 1997) using a weighted least squares method, where the weights were computed as Eq. (8) considering the MODIS BRDF quality flag (MCD43A2) information (in particular, the quality flag information associated to the bands employed in $fAPAR$ estimation). A criterion to detect outliers resulting from clouds, shadows, etc. was first adopted and the weights

associated to these outliers were reduced in order to obtain an appropriate filtering of the signal.

$$w = 1/(1 + 0.5 \times QF) \quad (8)$$

A $QF = 254$ has been assigned to missing data. According to Eq. (8), the associated weight is around zero, giving the most of importance to the neighboring data, allowing LOESS to automatically interpolate this kind of data.

This method does not require fitting all data to the same polynomial. The width $m = 8$ of the smoothing window and the degree $d = 4$ of the smoothing polynomial have been chosen using a cross-validation procedure. These values allow to capture subtle and rapid changes and to fill the gaps in the time series.

The filtered time series, $fAPAR_{smooth}$, was used to adjust the weights of each point in the original $fAPAR$ time series. When $fAPAR < fAPAR_{smooth}$ the weight is decreased in a factor of 0.25. The points with higher $fAPAR$ values are more likely to reflect the actual vegetation cycle and can be regarded as acceptable points, so we gave them relatively higher weights.

Finally, the LOESS method was applied again to the original $fAPAR$ time series but considering the new re-adjusted weights. The results show that this methodology captures the upper envelope of the time-series, interpolate the missing data and remove most of the noise of the original unfiltered signal.

The PRI and the $NPRI$ series were also filtered using this LOESS method. As no QFs are available, the cloud mask was used to properly discard the cloudy pixels.

4. Results

GPP data from the tower (Fig. 1) show a typical seasonal evolution, with maxima in summer and minima in winter. The main contribution to GPP is the PAR (also presenting maxima in summer and minima in winter). The coefficient of determination between GPP ($\text{gC m}^{-2} \text{day}^{-1}$) and PAR ($\text{MJ m}^{-2} \text{day}^{-1}$) is $R^2 = 0.624$. When annual data are analyzed separately, this coefficient ranges from 0.53 in year 2003 (severely affected by drought) to 0.75 in year 2002.

Fig. 2 (a) shows the ratio $GPP/PAR = \varepsilon fAPAR$. The minima observed in summer are partly due to the vegetation water stress affecting this Mediterranean forest. The $fAPAR$ in Fig. 2 (b) presents low seasonality, as expected for this type of forest. The $\varepsilon = GPP/(fAPAR \cdot PAR)$ in Fig. 2 (c) shows minima in summer. They are more pronounced in 2001 and 2003 because these years were affected by a moderate and a severe drought, respectively (Reichstein et al., 2007). This is also shown by the water stress factor C_{ws} (Fig. 2 (d)), which indicates the presence and intensity of water stress periods.

The MODIS PRI time series in Fig. 2 (e) shows the seasonal variability with maxima in winter and minima in summer. The coefficient of determination between PRI and ε is $R^2 = 0.75$, mainly because of their similar seasonality. However the values of the PRI minima do not indicate clearly the periods of significant water stress effect (such as summer 2003).

As a result from the elastic net analysis, the main geometric variables to be considered to model the PRI signal are the solar and the view zenith angles. Fig. 3 shows the profile of the coefficient paths for the fitted model with the elastic net. All predictors have been standardized in order to obtain inter-comparable coefficients and evaluate their relative importance. Starting from a value of variance (captured by the model) equal to zero, penalties have their highest values and all coefficients have been shrunk to zero as an effect of the high penalty. As the global regularization parameter θ is relaxed (lower values in Eq. 6), the predictors that are most significant are firstly added to the model. This is an effect of the lasso penalty and results in variable selection (variables with coefficients of zero are omitted from the model). The profiles of the coefficients indicate that the first predictor (which

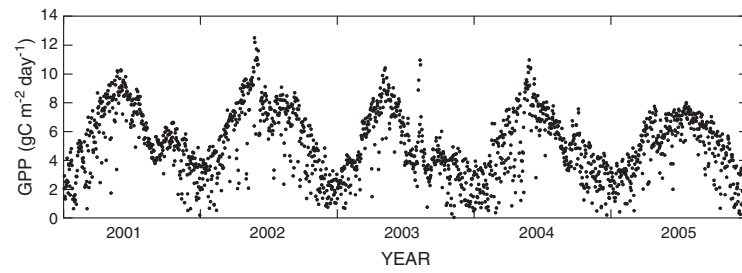


Fig. 1. GPP from the tower in San Rossore along five years.

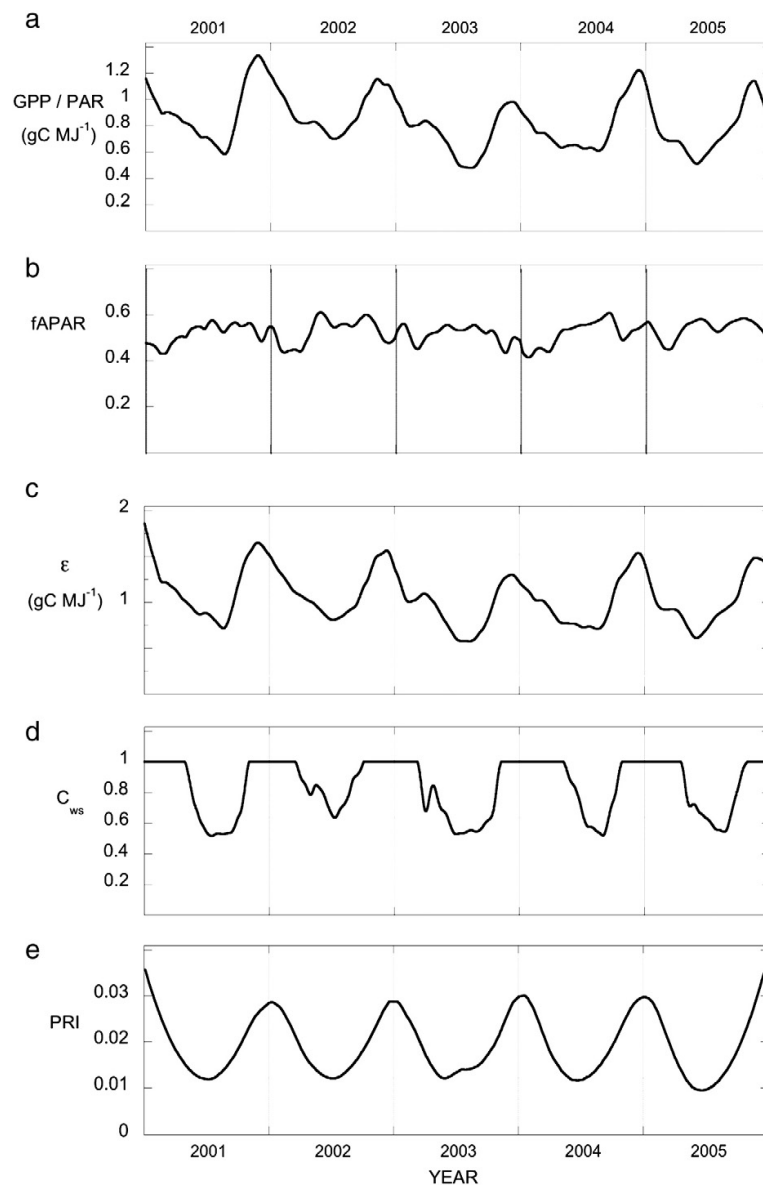


Fig. 2. For the tower pixel: (a) GPP/PAR ; (b) $fAPAR$; (c) ϵ ; (d) C_{ws} ; and (e) MODIS PRI . All the time series have been smoothed using the LOESS method.

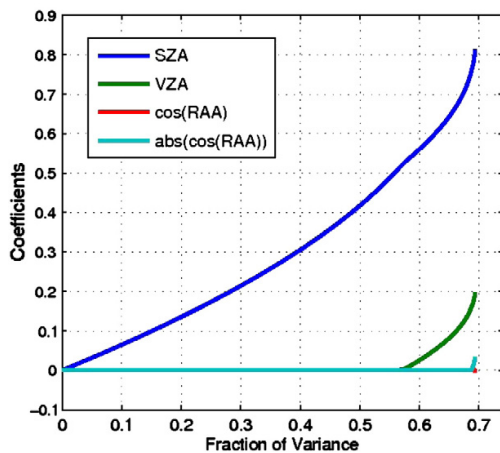


Fig. 3. Coefficients for the elastic-net estimates as a function of the variance explained by the model. SZA is the solar zenith angle, VZA is the view zenith angle and RAA is the sun-sensor relative azimuth angle.

captures most of *PRI* variance) included in the model is SZA. The second significant predictor is VZA, while the effect of RAA is negligible. According to these results, only SZA and VZA dependencies were currently considered.

The *PRI* modeling from the solar and view zenith angles has been carried out using both first- and higher-order polynomials. Only quadratic terms resulted in a marginal improvement on *PRI* modeling (the variances captured by the model were 69% and 75% for the linear and non-linear relationships, respectively). No significant improvement was observed when using third- or higher-order polynomials.

Fig. 4 shows the *PRI* modeled from sun and view zenith angles using the linear approach, and the measured *PRI*. The Jarque–Bera test (Jarque & Bera, 1987) shows that the residuals are normal with a 95% confidence interval.

To confirm that the sun and view zenith angles are the main drivers of the relationship between ϵ and *PRI*, an additional analysis has been carried out splitting the data in non-summer (high and medium SZA) and summer (low SZA) data (from 21 June to 23 September). The non-summer *PRI* signal modeled using the external factors (solar and view zenith angles) fits the original *PRI* signal with a high coefficient of determination ($R^2 = 0.78$). However, for the summer period there is practically no correlation (coefficient of determination $R^2 = 0.09$) between the original *PRI* signal and the modeled signal, which means that in this period there is additional information likely associated to the vegetation internal factors. In this last case, the residual is non-normal as shown by a Jarque–Bera test, indicating that the model cannot explain the observed *PRI* data.

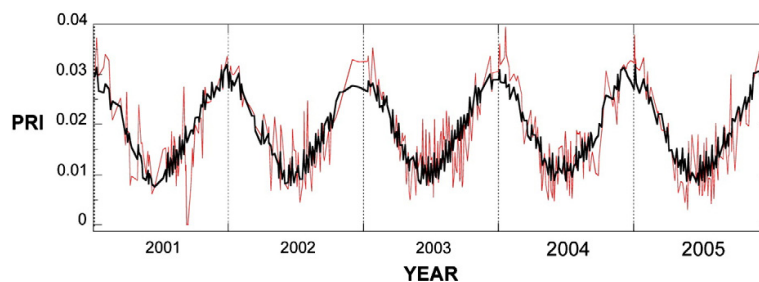


Fig. 4. Calculated *PRI* from the elastic-net analysis as a function of the sun and view zenith angles (black) (linear model) and measured *PRI* (red).

Once the *PRI* dependencies on solar and view zenith angles are quantified, a normalized *PRI* (*NPRI*) is obtained as the residuals computed after eliminating both dependencies in the *PRI* signal (Eq. 7). This *NPRI* is expected to be related to vegetation status and not to scene anisotropy due to geometry. Indeed, it identifies the inter-annual variations of the forest stand shown by the ϵ , and in particular the minima due to the droughts in 2001 and 2003 (Fig. 5). However, a time shift is observed between the ϵ and the *NPRI*. Fig. 5 (b) shows the *NPRI* calculated from the linear model, but the same shift is observed when using the quadratic model. This implies that the shift, which has a value of around two months, is not related to the model used in the normalization procedure.

In Figs. 2 and 5, both *PRI* and *NPRI* signals have been smoothed using the LOESS filtering. As a shift is observed also between them, we conclude that the shift between *NPRI* and ϵ is not a consequence of the smoothing procedure and should contain information related to the vegetation state. The application of a wavelet based smoothing procedure (Percival & Walden, 2000) confirms the existence of this shift, which is also present in the *PRI* series of neighboring pixels with similar characteristics.

In order to better quantify the shift between the different variables, a cross correlation analysis has been carried out. The correlation between different pairs of variables is calculated for a shift ranging from 0 to 1, 2, ..., N days. Table 2 shows the value of N corresponding to the highest correlation found between the different pairs of variables. Results are shown for the smoothed series (and interpolated to fill gaps). As mentioned above, this smoothing does not produce the shift (for example, the shift between ϵ and *PRI* was about a week, the same before and after the smoothing). The smoothing raises the correlation coefficients since at the same time it fills the gaps existing in particular days without *PRI* measurement and removes possible white noise on the signal.

Further analyses have been carried out to better confirm the existence of the shift between *NPRI* and ϵ time series. Firstly, besides the *RDVI fAPAR*, two more *fAPAR* products have been analyzed: the MODIS *fAPAR* product (MOD15A2), and a traditional *fAPAR* derived from a linear approach with the *NDVI* (Myneni & Williams, 1994) (in our case, the nadir *NDVI* from the k_0 (MCD43A1)). The three *fAPAR* profiles are generally similar apart from a scale factor, and result in rather similar ϵ profiles (apart the scale factor). The coefficients of determination between ϵ values and *PRI* are $R^2 = 0.75$, 0.74 and 0.64 for the *RDVI* ϵ , the MODIS ϵ and the *NDVI* ϵ , respectively. As mentioned before, these relatively large values result from their similar seasonality. The coefficients of determination lower to ca. $R^2 = 0.1$ when the *NPRI* is used instead of *PRI* due to the temporal shift. When the *NPRI* time series is shifted 65 days, the coefficient of determination increases in all cases to about $R^2 = 0.5$.

Finally, in order to check if the results are specific of the TERRA products (selected for the study due to their availability for the measurement period), *PRI* and *NPRI* data were also derived from AQUA products (shorter time series). Similar results (*NPRI* profile and

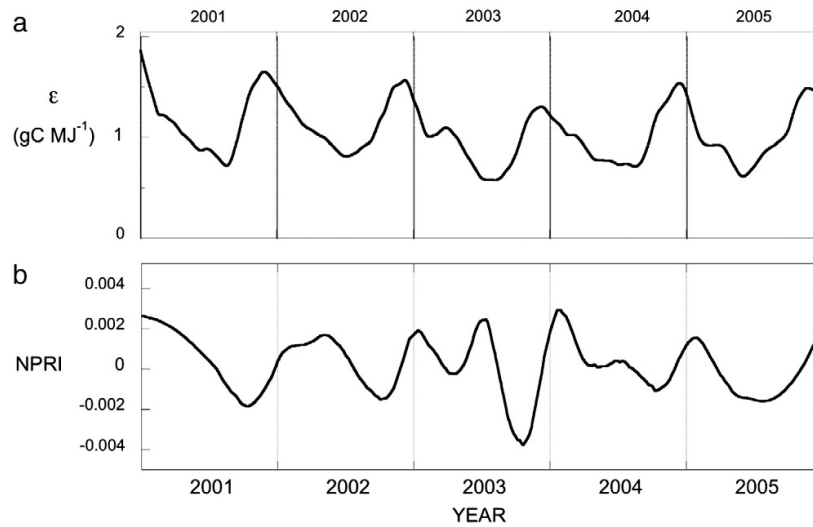


Fig. 5. (a) ϵ , and (b) $NPRI$ data from MODIS.

shift) were again obtained. These results also show that the observed shift is independent of the TERRA and AQUA footprint differences.

In summary, results have shown that the $NPRI$ shift is not associated neither with the smoothing procedure, nor the pixel selected as tower pixel, nor the $JAPAR$ used to derive ϵ , and nor TERRA images. On the other hand, the PRI signal is not as shifted as the $NPRI$ with respect to the ϵ since both PRI and ϵ present a seasonal trend highly affected by external factors. Therefore, the $NPRI$ shift appears after removing the effect of external factors (solar and view zenith angles) on PRI signal.

The presence of this shift would imply the existence of biophysical mechanisms different from those usually claimed to be at the origin of the relationship ϵ – PRI .

5. Discussion

The current experiment has investigated the relationship between ϵ and MODIS PRI concerning an evergreen coniferous ecosystem with typical Mediterranean features, i.e. whose seasonal growth is limited by water availability in summer. The dependence of PRI on external factors has been analyzed by the use of the elastic net approach, which is capable of overcoming most limitations of conventional regression methods (Zou & Hastie, 2005).

In this paper, we accepted that the ϵ at time of satellite overpass is representative for the whole day, as checked by Goerner et al. (2011). Intra-day ϵ variations can be notable, but the maximum effect on photosynthesis is linked to near midday ϵ , which is well represented by MODIS observations (see for ex. Drolet et al., 2005; Goerner et al., 2009; Goerner et al., 2011).

Table 2

Shift values after smoothing and interpolation (in days) between different pairs of variables and coefficient of correlation for two different adjustment models (linear and quadratic) used to normalize the PRI signal from external factors (solar and view angles).

Pair of variables		Shift (days) (coefficient of determination)
ϵ/PRI		8 ($R^2 = 0.69$)
$PRI/NPRI$	Linear model	50 ($R^2 = 0.28$)
	Quad. model	50 ($R^2 = 0.28$)
$\epsilon/NPRI$	Linear model	65 ($R^2 = 0.55$)
	Quad. model	80 ($R^2 = 0.44$)

The results indicate that the ϵ and PRI minima during summer would derive (at least partly) from concomitant decreases caused mainly by illumination conditions. The decrease of ϵ in summer is a common feature of most vegetation types due to several causes (Garbulsky et al., 2010; Maselli et al., 2009). First, the asymptotic dependence of photosynthesis on solar radiation intrinsically produces ϵ decreases for high radiation levels (near the summer solstice). This behavior is related to internal and external factors such as saturation of photosynthetic apparatus, soil nutrient limitation, and, particularly, water stress (Waring & Running, 2007). The maximum effect of water stress generally coincides with high potential evapotranspiration (and often low rainfall) in summer, which exacerbates the observed ϵ decrease.

A decrease of PRI with increasing illumination has been found in a number of investigations. PRI has been found to be negatively correlated with incident radiation and positively correlated with sun zenith angle and shadow fractions (Drolet et al., 2008; Goerner et al., 2009; Hall et al., 2008). The origin of these correlations is not completely understood. The change of the contribution of sunlit and shaded leaves to the overall optical signal received by the sensor, and the differences in atmospheric scattering and canopy shadowing at different wavelengths might be of relevance (Hilker et al., 2008). Recently, Hall et al. (2011) showed that using PRI to infer ϵ is confounded, among others, by the shadow fraction viewed by a sensor.

Our statistical analysis confirms that sun and view zenith angles account for a significant part of PRI seasonal evolution (75%–69%, with quadratic and linear model, respectively). This result does not imply that PRI cannot detect spectral changes due to the xanthophylls cycle. As previously noted, the signal associated with these changes is very low, and may be hidden by stronger reflectance variations. For example, in the current case of a Mediterranean forest, the variations of MODIS R_{531} associated with seasonal illumination changes have a range of about 0.1 (from 0.05 to 0.15).

The MODIS PRI can track the inter-annual ϵ variation when only view angles lower than 40° are considered and when the dependence of PRI on the solar and view zenith angles is eliminated or normalized ($NPRI$). A shift of about two months is observed between ϵ and $NPRI$ time series, which suggests that the $NPRI$ is not showing the short-term spectral changes due to the xanthophylls de-epoxidation.

The validity of this result relies on the effectiveness of the regression approach applied to model the dependence of PRI on solar and view angles and compute unbiased residuals. Such effectiveness has

been confirmed by further statistical analysis using different models (linear and quadratic). Moreover, the shift has been proven not to be a consequence of the smoothing procedures, and not to depend on the data source (MODIS TERRA/AQUA).

The eco-physiological interpretation of this shift is actually problematic, since the *NPRI* minimum is found when photosynthetic efficiency has apparently recovered. This ϵ recovering, however, could occur also independently of water stress cessation as a consequence of the *PAR* drop, which characterizes the period of the year around the autumn equinox. The previously mentioned asymptotic dependence of *GPP* on *PAR*, in fact, implies an automatic increase of ϵ with decreasing *PAR*, which could act even in the presence of persistent drought.

This hypothesis is supported by the analysis of the variables that contribute to ϵ calculation, i.e. *GPP*, *fAPAR* and *PAR*, for the two months around the equinox (from 21 August to 21 October). Mean *GPP* decreases slightly in this period (Fig. 1), while *fAPAR* is virtually stable (Fig. 2 (b)). The mean *PAR* halving (from about 8 to 4 MJ m⁻² day⁻¹) is strictly correlated with the corresponding ϵ rise ($R^2 = 0.61$). Water stress can be still active in October, particularly in the driest seasons of 2001 and 2003 (see C_{ws} in Fig. 2 (d)), when it can cause clear *NPRI* minima (Fig. 5).

These considerations can provide a plausible eco-physiological framework for the current experimental evidences. The mentioned nonlinear dependence of *GPP* on *PAR* can obscure the linkage between water stress and ϵ in late summer–early autumn. The persistence of this stress, however, can determine first leaf yellowing/wilting and eventually structural canopy damage. The former would be visible through *NPRI* decreases, while the latter could result in drops of more conventional vegetation indices related to *fAPAR*.

6. Conclusions

GPP from a flux tower within an evergreen forest in San Rossore (Tuscany) during five years (2001–2005) has been used to estimate the canopy ϵ via the *fAPAR* derived from MODIS images. The *fAPAR* has been derived from three different approaches: *RDVI fAPAR*, *MODIS fAPAR* and *NDVI fAPAR*. The corresponding ϵ values show similar temporal variation and differ significantly only in a scale factor. The inter- and intra-annual temporal patterns of MODIS *PRI* for this forest canopy have been compared to those of ϵ .

The experimental results obtained from this analysis lead to formulate the following conclusions for this Mediterranean evergreen forest and at MODIS spatial scale: (i) The coefficient of determination between *PRI* and ϵ is 0.69. (ii) The dependence of *PRI* on the sun and view zenith angles has been determined by using statistical methods. (iii) Most (around 70%) of the observed concurrent seasonal variability in ϵ and *PRI* is due to the effect of illumination and viewing conditions (first order effects), part to chlorophyll and canopy structure (second order effects), and the de-epoxidation of xanthophylls only introduce third order effects (which are, at MODIS pixel scale of the same magnitude order than the sensor noise). (iv) When the first-order variations are removed, also spectral changes due to xanthophylls (third order effects) seem to be removed, or at least, they cannot be detected in the residual *PRI* (*NPRI*). (v) The shift between ϵ and *NPRI* time series can be partly attributed to the *PAR* decrease around the autumn equinox. (vi) *NPRI* seems to be mostly related to the changes in pigments (chlorophyll and carotenoids) and canopy structure (leaf wilting), which characterize the vegetation response to particularly long and intense water stress. (vii) *NPRI* becomes informative on actual inter-annual ϵ variability associated with prolonged water stress.

Due to the limited study area, the above conclusions may have local validity, although they are thought to be of wider interest. Further studies are necessary to cast light on this potentially relevant issue regarding forest ecosystems with different structural and functional features.

Acknowledgments

This work was supported by the ÁRTEMIS Project from the Spanish Ministry of Science and Innovation (CGL2008-00381). We would also like to acknowledge the suggestions and advices from the anonymous referees, which helped us to improve the paper.

Appendix A. Supplementary data

Supplementary data associated with this article can be found in the online version, at <http://dx.doi.org/10.1016/j.rse.2012.04.003>. These data include Google map of the most important areas described in this article.

References

- Atkeson, C. G., Moore, A. W., & Schaal, S. (1997). Locally weighted learning. *Artificial Intelligence Review*, 11, 11–73.
- Baret, F., Hagolle, O., Geiger, B., Bicheron, P., Miras, B., Huc, M., et al. (2007). LAI, *fAPAR* and *fCover* CYCLOPES global products derived from VEGETATION. Part 1: Principles of the algorithm. *Remote Sensing of Environment*, 110, 275–286.
- Barton, C. V. M., & North, P. R. J. (2001). Remote sensing of canopy light use efficiency using the photochemical reflectance index. Model and sensitivity analysis. *Remote Sensing of Environment*, 78, 264–273.
- Camacho, F. (2007). Evaluation of the Land-SAF *fAPAR* prototype along one year of MSG BRDF data: Algorithm, Product description, and intercomparison against equivalent satellite products and ground-truth, LSA SAF Visiting Scientist report 60 pp., available at <http://landsaf.meteo.pt/documentsView.jsp>
- Cheng, Y. B., Middleton, E. M., Huemmrich, K. F., Zhang, Q., Campbell, P. K. E., Corp, L. A., et al. (2010). Utilizing in situ directional hyperspectral measurements to validate bio-indicator simulations for a corn crop canopy. *Ecological Informatics*, 5, 330–338.
- Chiesi, M., Maselli, F., Bindi, M., Fibbi, L., Cherubini, P., Arlotta, E., et al. (2005). Modelling carbon budget of Mediterranean forests using ground and remote sensing measurements. *Agricultural and Forest Meteorology*, 135, 22–34.
- Connolly, J., Roulet, N. T., Seaquist, J. W., Holden, N. M., Lafleur, P. M., Humphreys, E. R., et al. (2009). Using MODIS derived *fPAR* with ground based flux tower measurements to derive the light use efficiency for two Canadian peatlands. *Biogeosciences*, 6, 225–234.
- Coops, N. C., Hilker, T., Hall, F. G., Nichol, C. J., & Drolet, G. G. (2010). Estimation of light-use efficiency of terrestrial ecosystems from space: A status report. *BioScience*, 60, 788–797.
- Corona, P., & Marchetti, M. (2007). Outlining multi-purpose forest inventories to assess the ecosystem approach in forestry. *Plant Biosystems*, 141, 243–251.
- D.R.E.Am (2003). Note illustrative della carta forestale e della fruizione turistica. *Ente Parco Migliarino San Rossore, Massaciuccoli*. Firenze: S.E.L.C.A.
- Damm, A., Elbers, J., Erler, A., Gioli, B., Hamdi, K., Hutjes, R., et al. (2010). Remote sensing of sun-induced fluorescence to improve modeling of diurnal courses of gross primary production (GPP). *Global Change Biology*, 16, 171–186.
- Drolet, G. G., Huemmrich, K. F., Hall, F. G., Middleton, E. M., Black, T. A., Barr, A. G., et al. (2005). A MODIS-derived photochemical reflectance index to detect inter-annual variations in the photosynthetic light-use efficiency of a boreal deciduous forest. *Remote Sensing of Environment*, 98, 212–224.
- Drolet, G. G., Middleton, E. M., Huemmrich, K. F., Hall, F. G., Amiro, B. D., Barr, A. G., et al. (2008). Regional mapping of gross light-use efficiency using MODIS spectral indices. *Remote Sensing of Environment*, 112, 3064–3078.
- FAO (2005). Global Forest Resource Assessment. Progress towards sustainable forest management. available on <http://www.fao.org/docrep/008/a0400e/a0400e00.htm>
- Fensholt, R., Sandholt, I., & Rasmussen, M. S. (2004). Evaluation of MODIS LAI, *fAPAR* and the relation between *fAPAR* and NDVI in a semi-arid environment using in situ measurements. *Remote Sensing of Environment*, 91, 490–507.
- Field, C. B., Randerson, J. T., & Malmstrom, C. M. (1995). Global net primary production-combining ecology and remote-sensing. *Remote Sensing of Environment*, 51, 74–88.
- Filella, I., Porcar-Castell, A., Munne-Bosch, S., Back, J., Garbalsky, M. F., & Peñuelas, J. (2009). *PRI* assessment of long-term changes in carotenoids/chlorophyll ratio and short-term changes in de-epoxidation of the xanthophyll cycle. *International Journal of Remote Sensing*, 30, 4443–4455.
- Foken, T., & Wichura, B. (1996). Tools for quality assessment of surface based flux measurements. *Agricultural and Forest Meteorology*, 78, 83–105.
- Gamon, J. A., Field, C. B., Bilger, W., Björkman, O., Fredeen, A., & Peñuelas, J. (1990). Remote sensing of the xanthophyll cycle and chlorophyll fluorescence in sunflower leaves and canopies. *Oecologia*, 85, 1–7.
- Gamon, J. A., Peñuelas, J., & Field, C. B. (1992). A narrow-waveband spectral index that tracks diurnal changes in photosynthetic efficiency. *Remote Sensing of Environment*, 41, 35–44.
- Garbalsky, M. F., Peñuelas, J., Gamon, J., Inoue, Y., & Filella, I. (2011). The photochemical reflectance index (*PRI*) and the remote sensing of leaf, canopy and ecosystem radiation use efficiencies. A review and meta-analysis. *Remote Sensing of Environment*, 15, 281–297.
- Garbalsky, M. F., Peñuelas, J., Papale, D., Ardö, J., Goulden, M. L., Kiely, G., et al. (2010). Patterns and controls of the variability of radiation use efficiency and primary

- productivity across terrestrial ecosystems. *Global Ecology and Biogeography*, 19, 253–267.
- Garbulska, M. F., Peñuelas, J., Papale, D., & Filella, I. (2008). Remote estimation of carbon dioxide uptake by a Mediterranean forest. *Global Change Biology*, 14, 2860–2867.
- García-Haro, F. J., Camacho, F., & Meliá, J. (2008). *Product User Manual (PUM) of Land Surface Analysis Vegetation products (SAF/LAND/UV/PUM_VEGA/2.1)*. 53 pp.
- Gobron, N., Auzan, O., Pinty, B., Taberner, M., & Verstraete, M. (2004). MERIS level 2 land surface products algorithm theoretical basis document. *JRC Publication No. EUR 21387 EN*.
- Gobron, N., Pinty, B., Mélin, F., Taberner, M., Verstraete, M., Robustelli, M., et al. (2007). Evaluation of the MERIS/ENVISAT FAPAR product. *Advance in Space Research*, 39, 105–115.
- Goeckede, M., Foken, T., Aubinet, M., more than 30 authors, et al. (2008). Quality control of CarboEurope flux data – Part 1: Coupling footprint analyses with flux data quality assessment to evaluate sites in forest ecosystems. *Biogeosciences*, 5, 433–450.
- Goerner, A., Reichstein, M., & Rambal, S. (2009). Tracking seasonal drought effects on ecosystem light use efficiency with satellite-based PRI in a Mediterranean forest. *Remote Sensing of Environment*, 113, 1101–1111.
- Goerner, A., Reichstein, M., Tomelleri, E., Hanan, N., Rambal, S., Papale, D., et al. (2011). Remote sensing of ecosystem light use efficiency with MODIS-based PRI. *Biogeosciences*, 8, 189–202.
- Goward, S., Markham, B., & Dye, D. (1991). Normalized difference vegetation index measurements from the advanced very high resolution radiometer. *Remote Sensing of Environment*, 35, 257–277.
- Gutman, G. (1991). Vegetation indices from AVHRR: An update and future prospects. *Remote Sensing of Environment*, 35, 121–136.
- Hall, F. G., Hilker, T., & Coops, N. C. (2011). PHOTOSYNSTAT, photosynthesis from space: Theoretical foundations of a satellite concept and validation from tower and spaceborne data. *Remote Sensing of Environment*, 115, 1918–1925.
- Hall, F. G., Hilker, T., Coops, N. C., Lyapustin, A., Huemmrich, K. F., Middleton, E., et al. (2008). Multi-angle remote sensing of forest light use efficiency by observing PRI variation with shadow fraction. *Remote Sensing of Environment*, 112, 3201–3211.
- Heinsch, F. A., Reeves, M., Votava, P., Kang, S., Milesi, C., Zhao, M., et al. (2003). User's guide: GPP and NPP (MOD17A2/A3) products. *NASA MODIS land algorithm, version 2.0* (pp. 1–57).
- Hilker, T., Coops, N. C., Hall, F. G., Black, T. A., Wulder, M. A., Nesic, Z., et al. (2008). Separating physiological and directionally induced changes in PRI using BRDF models. *Remote Sensing of Environment*, 112, 2777–2788.
- Hilker, T., Hall, F. G., Coops, N. C., Lyapustin, A., Wang, Y., Nesic, Z., et al. (2010). Remote sensing of photosynthetic light-use efficiency across two forested biomes: Spatial scaling. *Remote Sensing of Environment*, 114, 2863–2874.
- Hilker, T., Lyapustin, A., Hall, F. G., Wang, Y., Coops, N. C., Drolet, G., et al. (2009). An assessment of photosynthetic light use efficiency from space: Modeling the atmospheric and directional impacts on PRI reflectance. *Remote Sensing of Environment*, 113, 2463–2475.
- Hoerl, A. E., & Kennard, R. W. (1970). Ridge regression: Applications to nonorthogonal problems. *Technometrics*, 12, 69–82.
- Holben, B. N. (1986). Characteristics of maximum-value composite images from temporal AVHRR data. *International Journal of Remote Sensing*, 7, 1417–1434.
- IPCC (2007). Climate change 2007. *IPCC Fourth Assessment Report*. Available from: <http://www.ipcc.ch/>
- Jarque, C. M., & Bera, A. K. (1987). A test for normality and regression residual. *International Statistical Review*, 55, 163–172.
- Knyazikhin, Y., Glassy, J., Privette, J. L., Tian, Y., Lott, A., Zhang, Y., et al. (1999). MODIS Leaf Area Index (LAI) and Fraction of Photosynthetically Active Radiation Absorbed by Vegetation (FPAR) Product (MOD15). *Algorithm theoretical basis document* (<http://eosps.gsfc.nasa.gov/atbd/mod15.html>)
- Mänd, P., Hallik, L., Peñuelas, J., Nilson, T., Duce, P., Emmet, B. A., et al. (2010). Responses of the reflectance indices PRI and NDVI to experimental warming and drought in European shrublands along a north–south climatic gradient. *Remote Sensing of Environment*, 114, 626–636.
- Maselli, F., Barbati, A., Chiesi, M., Chirici, G., & Corona, P. (2006). Use of remotely sensed and ancillary data for estimating forest gross primary productivity in Italy. *Remote Sensing of Environment*, 100, 563–575.
- Maselli, F., Chiesi, M., Moriondo, M., Fibbi, L., Bindi, M., & Running, S. W. (2010). Modelling of forest carbon fluxes through the integration of ground and satellite data. In F. Maselli, M. Menenti, & P. A. Brivio (Eds.), *Remote sensing optical observations of vegetation properties, 2010978-81-308-0421-7*.
- Maselli, F., Papale, D., Puletti, N., Chirici, G., & Corona, P. (2009). Combining remote sensing and ancillary data to monitor the gross productivity of water-limited forest ecosystems. *Remote Sensing of Environment*, 113, 657–667.
- McCallum, I., Wagner, W., Schmittius, C., Shvidenko, A., Obersteiner, M., Fritz, S., et al. (2010). Comparison of four global FAPAR datasets over Northern Eurasia for the year 2000. *Remote Sensing of Environment*, 114, 941–949.
- Monteith, J. L. (1972). Solar radiation and productivity in tropical ecosystems. *Journal of Applied Ecology*, 9, 747–766.
- Myneni, R. B., Hoffman, S., Knyazikhin, Y., Privette, J. L., Glassy, J., Tian, Y., et al. (2002). Global products of vegetation leaf area and fraction absorbed PAR from year one of MODIS data. *Remote Sensing of Environment*, 83, 214–231.
- Myneni, R. B., & Williams, D. L. (1994). On the relationship between fAPAR and NDVI. *Remote Sensing of Environment*, 49, 200–211.
- Naumann, J. C., Bisset, S. N., Young, D. Y., Edwards, J., & Anderson, J. E. (2010). Diurnal patterns of photosynthesis, chlorophyll fluorescence, and PRI to evaluate water stress in the invasive species, *Elaeagnus umbellata* Thunb. *Trees*, 24, 237–245.
- Papale, D., Reichstein, M., Aubinet, M., Canfora, E., Bernhofer, C., Kutsch, W., et al. (2006). Towards a standardized processing of Net Ecosystem Exchange measured with eddy covariance technique: Algorithms and uncertainty estimation. *Biogeosciences*, 3, 571–583.
- Percival, D. B., & Walden, A. T. (2000). *Wavelet methods for time series analysis: Cambridge University Press* 594 pp.
- Plummer, S., Arino, O., Simon, M., & Steffen, W. (2006). Establishing an Earth observation product service for the terrestrial carbon community: The GlobCarbon Initiative. *Mitigation and Adaptation Strategies for Global Change*, 11, 97–111.
- Potter, C. S., Randerson, J. T., Field, C. B., Matson, P. A., Vitousek, P. M., Mooney, H. A., et al. (1993). Terrestrial ecosystem production: A process model based on global satellite and surface data. *Global Biogeochemical Cycles*, 7(4), 811–841.
- Rapetti, F., & Vittorini, S. (1995). *Carta climatica della Toscana*. Pisa (Italy): Pacini Editore.
- Reichstein, M., Ciais, P., Papale, D., Valentini, R., Running, S., Viovy, N., et al. (2007). A combined eddy covariance, remote sensing and modeling view on the 2003 European summer heatwave. *Global Change Biology*, 13, 634–651.
- Reichstein, M., Falge, E., Baldocchi, D., Papale, D., Aubinet, M., Berbigier, P., et al. (2005). On the separation of net ecosystem exchange into assimilation and ecosystem respiration: Review and improved algorithm. *Global Change Biology*, 11, 1424–1439.
- Rodolfi, A., Chiesi, M., Tagliaferri, G., Cherubini, P., & Maselli, F. (2007). Assessment of forest GPP variations in Central Italy by the analysis of meteorological, satellite and dendrochronological data. *Canadian Journal of Forest Research*, 37, 1944–1953.
- Roujean, J. L., & Bréon, F. M. (1995). Estimating PAR absorbed by vegetation from Bidirectional Reflectance Measurements. *Remote Sensing of Environment*, 51, 375–384.
- Sarikioti, V., Driever, S. M., & Marcelis, L. F. M. (2010). Photochemical reflectance index as a mean of monitoring early water stress. *Annals of Applied Biology*, 157, 81–89.
- Seixas, J., Carvalhais, N., Nunes, C., & Benali, A. (2009). Comparative analysis of MODIS-FAPAR and MERIS-MGVI datasets: Potential impacts on ecosystem modeling. *Remote Sensing of Environment*, 113, 2547–2559.
- Sims, D. A., Rahman, A. F., Cordova, V. D., El-Masri, B. Z., Baldocchi, D. D., Bolstad, P. V., et al. (2008). A new model of gross primary productivity for North American ecosystems based solely on the enhanced vegetation index and land surface temperature from MODIS. *Remote Sensing of Environment*, 112, 1633–1646.
- Suárez, L., Zarco-Tejada, P. J., Berni, J. A. J., González-Dugo, V., & Fereres, E. (2009). Modelling PRI for water stress detection using radiative transfer models. *Remote Sensing of Environment*, 113, 730–744.
- Thenot, F., Méthy, M., & Winkel, T. (2002). The Photochemical Reflectance Index (PRI) as a water-stress index. *International Journal of Remote Sensing*, 23, 5135–5139.
- Thornton, P. E., Hasenauer, H., & White, M. A. (2000). Simultaneous estimation of daily solar radiation and humidity from observed temperature and precipitation: An application over complex terrain in Austria. *Agricultural and Forest Meteorology*, 104, 255–271.
- Thornton, P. E., Running, S. W., & White, M. A. (1997). Generating surfaces of daily meteorological variables over large regions of complex terrain. *Journal of Hydrology*, 190, 214–251.
- Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society, Series B*, 58, 267–288.
- Turner, D. P., Urbanski, S., Bremer, D., Wofsy, S. C., Meyers, T., Gower, S. T., et al. (2003). A cross-biome comparison of daily light use efficiency for gross primary production. *Global Change Biology*, 9, 383–395.
- Veroustraete, F., Sabbe, H., & Eeren, H. (2002). Estimation of carbon mass fluxes over Europe using the C-Fix model and Euroflux data. *Remote Sensing of Environment*, 83, 376–399.
- Veroustraete, F., Sabbe, H., Rasse, D. P., & Bertels, L. (2004). Carbon mass fluxes of forests in Belgium determined with low resolution optical sensors. *International Journal of Remote Sensing*, 25, 769–792.
- Waring, H. R., & Running, S. W. (2007). *Forest ecosystems. Analysis at multiple scales* (3rd edition). San Diego: Academic Press.
- Weiss, M., Baret, F., Garrigues, S., & Lacaze, R. (2007). LAI and FAPAR CYCLOPES global products derived from VEGETATION. Part 2: Validation and intercomparison with MODIS Collection 4 products. *Remote Sensing of Environment*, 110, 317–331.
- Wolfe, R. E., Roy, D. P., & Vermote, E. F. (1998). MODIS land data storage, gridding, and compositing methodology: Level 2 grid. *IEEE Transactions on Geoscience and Remote Sensing*, 36, 1324–1338.
- Wu, C., Niu, Z., Tang, Q., & Huang, W. (2010). Revised photochemical reflectance index (PRI) for predicting light use efficiency of wheat in a growth cycle: Validation and comparison. *International Journal of Remote Sensing*, 31, 2911–2924.
- Xiong, X., Sun, J., Esposito, J., Guenter, B., & Barnes, W. (2003). MODIS reflective solar bands calibration algorithm and on-orbit performance. *Proceedings of SPIE Optical Remote Sensing of the Atmosphere and Clouds III*, 4891. (pp. 95–104).
- Zenone, T., Morelli, G., Teobaldelli, M., Fischinger, F., Matteucci, M., Sordini, M., et al. (2008). Preliminary use of ground-penetrating radar and electrical resistivity tomography to study tree roots in pine forests and poplar plantations. *Functional Plant Biology*, 35, 1047–1058.
- Zhang, Q., Middleton, E. M., Margolis, H. A., Drolet, G. G., Barr, A. A., & Black, T. A. (2009). Can a satellite-derived estimate of the fraction of PAR absorbed by chlorophyll (FAPAR_{chl}) improve predictions of light-use efficiency and ecosystem photosynthesis for a boreal aspen forest? *Remote Sensing of Environment*, 113, 880–888.
- Zhang, Q., Xiao, X., Braswell, B., Linder, E., Baret, F., & Moore, B., III (2005). Estimating light absorption by chlorophyll, leaf and canopy in a deciduous broadleaf forest using MODIS data and a radiative transfer model. *Remote Sensing of Environment*, 99, 357–371.
- Zhao, H., & Fernandes, R. (2009). Daily snow cover estimation from Advanced Very High Resolution Radiometer Polar Pathfinder data over Northern Hemisphere land surfaces during 1982–2004. 14 pp. *Journal of Geophysical Research*, 114(D05113). <http://dx.doi.org/10.1029/2008JD101272>.
- Zou, H., & Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 67, 301–320.

Apéndice **D**

Publication IV: Monitoring water stress in Mediterranean semi-natural vegetation with satellite and meteorological data



Contents lists available at ScienceDirect

International Journal of Applied Earth Observation and Geoinformation

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

Monitoring water stress in Mediterranean semi-natural vegetation with satellite and meteorological data

A. Moreno^a, F. Maselli^b, M. Chiesi^b, L. Genesio^b, F. Vaccari^b, G. Seufert^c, M.A. Gilabert^{a,*}^a Departament de Física de la Terra i Termodinàmica, Facultat de Física, Universitat de València, Dr. Moliner 50, 46100 Burjassot, Spain^b IBIMET-CNR, Via Madonna del Piano 10, 50019 Sesto Fiorentino (FI), Italy^c Institute for Environmental Sustainability, JRC, Via Fermi 2749-TP 050, 21027 Ispra (VA), Italy

ARTICLE INFO

Article history:

Received 9 April 2013

Accepted 9 August 2013

Keywords:

Light use efficiency

Water stress

Satellite data

Semi-natural vegetation

ABSTRACT

In arid and semi-arid environments, the characterization of the inter-annual variations of the light use efficiency ϵ due to water stress still relies mostly on meteorological data. Thus the GPP estimation based on procedures exclusively driven by remote sensing data has not found yet a widespread use. In this work, the potential to characterize the water stress in semi-natural vegetation of three spectral indices (NDWI, SIWSI and NDI7) – from MODIS broad spectral bands – has been analyzed in comparison to a meteorological factor (C_{ws}). The study comprises 70 sites (belonging to 7 different ecosystems) uniformly distributed over Tuscany, and three eddy covariance tower sites. An operational methodology, which combines meteorological and MODIS data, to characterize the inter-annual variations of ϵ due to summer water stress is proposed. Its main advantage is that it relies on existing series of meteorological data characterizing each site and allows calculating a typical C_{ws} profile that can be “updated” (C_{ws}^*) for the actual conditions using MODIS spectral indices. The results confirm that the modified C_{ws}^* can be used as a proxy of water stress that does not require concurrent information on meteorological data.

© 2013 Elsevier B.V. All rights reserved.

1. Introduction

In Mediterranean areas, the coincidence of the dry season with the period of maximum solar irradiation and temperature creates particular problems for annual vegetation growth, which is strongly limited by summer water availability (Bolle et al., 2006). Water stress affects the vegetation production by a reduction of the leaf area, the stomatal conductance and the CO₂ uptake and hence of the photosynthesis, and by a slowdown of root elevation and development (Verstraeten et al., 2006). The gross primary production (GPP), which corresponds to total ecosystem photosynthesis, is an essential parameter to characterize most relevant ecosystem processes. It can be modeled using the Monteith's approach (1972) as the product of the APAR (viz. the photosynthetic active radiation, PAR, absorbed by the canopy) and ϵ (the light use efficiency in the conversion of APAR into photosynthetically fixed CO₂). The APAR in turn can be obtained as the product of the f_{APAR} (fraction of PAR that is absorbed) and the incident PAR, whereas ϵ can be estimated from a maximum value (ϵ_{max}) corresponding to the optimal ecosystem functioning and depending on the plant functional type (Garbulsky et al., 2009; Heinsch et al., 2003), which is down-regulated through

the use of several dimensionless scalars (varying from 0 to 1) that quantify different stress conditions.

Relying on the use of eddy covariance (EC) flux tower data several authors have shown that ϵ of Mediterranean forests is strongly reduced during dry summer spells (Allard et al., 2008; Garbulsky et al., 2008; Moreno et al., 2012) and presents an inter-annual variability depending on the intensity and duration of the summer water stress period. Therefore, as the water stress is the main factor controlling the inter-annual variations of ϵ in Mediterranean ecosystems, it is crucial to define appropriate scalars to account for its temporal variability (i.e., the factors that modify the maximum ϵ defined under non-stress conditions). These ‘water stress scalars’ can be derived from both meteorological and remotely sensed data. For example, the MODIS approach uses the daily average vapor pressure deficit, which generally tends to underestimate the effect of water stress (Hansen et al., 2000; Running et al., 2004). The CASA water scalar is derived from a local water budget that takes into account the actual and potential evapotranspiration (AET and PET, respectively) (Field et al., 1995; Potter et al., 1993). Maselli et al. (2009a) uses the same theoretical foundation to introduce an index, C_{ws} , which is aimed to capture the short-term effect of water stress. Significant correlations have been found between inter-annual variations of ϵ from flux tower stations and C_{ws} . Its main drawback is the need of meteorological data since the network of weather stations can be sparse or non-existent. This justifies

* Corresponding author. Tel.: +34 963543118; fax: +34 963543385.
E-mail address: m.amparo.gilabert@uv.es (M.A. Gilabert).

Table 1

Main characteristics of the 70 sites selected in Tuscany (10 per forest type, FT). The FT number follows a gradient of decreasing xeric condition. For each FT, the minimum and maximum values of elevation, mean annual temperature and annual precipitation over the 10 sites are provided.

	FT	Elevation (m)	Mean annual temperature (°C)	Annual precipitation (mm)
1	Mediterranean macchia	67–436	13.9–15.9	590–739
2	Holm oak	121–482	13.6–15.4	701–923
3	Plane/hilly conifers	2–576	12.2–15.7	642–1182
4	Deciduous oaks	165–915	10.5–15.0	703–1551
5	Chestnut	420–888	10.7–12.6	949–2013
6	Beech	904–1540	6.6–10.7	994–1709
7	Mountain conifers	934–1364	6.6–10.0	1213–2067

the search for spectral indices sensitive to the short-term effect of water stress on vegetation.

Several attempts have been made to estimate the decrease of ϵ due to water stress through the use of remotely sensed spectral indices (Coops et al., 2010). These studies, however, have provided diverse results, since the spectral responses of plants to varying eco-physiological conditions are often complex and delayed (Maselli et al., 2009a). Moreover, and especially in water-limited ecosystems, many spectral indices are sensitive not only to leaf water content but also to LAI and pigment content, which often vary simultaneously (Danson and Bowyer, 2004). A relatively novel spectral index is the Photochemical Reflectance Index (PRI) (Gamon et al., 1990, 1992), based on the remote sensing of the xanthophyll plant pigment effect on the reflectance in the narrow band centered at 531 nm. PRI is claimed to be a photosynthetic efficiency indicator across a wide range of conditions, species and functional types, because it also correlates with chlorophyll/carotenoid ratios (Filella et al., 2009) and can be used to monitor early water stress (Naumann et al., 2010; Sarlikioti et al., 2010; Suárez et al., 2009; Thenot et al., 2002). However, the PRI efficiency at ecosystem level (MODIS pixel) is still elusive and no conclusive results have been obtained yet to track the inter-annual variability of ϵ as a consequence of the water stress. PRI is a narrow band spectral index and it has a very low signal and a rather high noise level. The reflectance change in the 531 band expected from xanthophylls conversion in presence of water stress is, at MODIS pixel scale, of the same magnitude order than the sensor noise (Coops et al., 2010; Hall et al., 2008). However, PRI is more affected by external factors such as illumination and view geometry (Drolet et al., 2008). In particular, Moreno et al. (2012) showed that in an evergreen coniferous ecosystem whose seasonal growth is limited by water availability in summer, most (around 70%) of the correlation between ϵ and PRI was explained by the PRI inter-annual variability due to the illumination and viewing conditions.

To avoid reflectance changes of the same magnitude order of sensor noise, some other spectral indices using broader bands can be used. In particular, as water stress produces a decrease of plant water and chlorophyll content, both affecting the vegetation reflectance, this stress can be theoretically assessed from spectral indices related to leaf water content (Danson et al., 1992; Zarco-Tejada et al., 2003). The water content decrease is noticeable during early stages mainly in the 0.9–2.5 μm (comprising the near infrared, NIR, and the short wave infrared, SWIR, spectral regions), mainly in bands centered at 1.45, 1.94, and 2.50 μm (Carter, 1991). However, the decrease of chlorophyll content (Goerner et al., 2011) – noticeable in visible (VIS) wavebands – requires more prolonged periods of drought. Several authors have demonstrated linkages between reflectance and different water content related indices using ground data (Ceccato et al., 2002; Fensholt and Sandholt, 2003; Clevers, 2008) as well as radiative transfer models at leaf and canopy level (Tarantola et al., 2001; Ceccato et al., 2002; Zarco-Tejada et al., 2003; Trombetti et al., 2008). Most of these indices are obtained from the normalized difference between the NIR

reflectance and a SWIR reflectance; for example the Normalized Difference Water Index (NDWI) (Gao, 1996), the Shortwave Infrared Water Stress Index (SIWSI) (Fensholt and Sandholt, 2003), and the Normalized Difference Water Index 7 (NDI7) (Rubio et al., 2006).

In this work, the potential to characterize the water stress in natural vegetation of several ‘water content indices’ – from MODIS broader spectral bands – has been analyzed. The MODIS bands used to obtain these indices present a higher spatial resolution (500 m instead of 1000 m) and a double width (of about 20–50 nm) as compared with the bands used to calculate PRI. Moreover, they are atmospherically corrected, and the BRDF parameters are known (in particular the k_0). Therefore these indices present a lower noise level, a better spatial resolution, and a higher processing level (corrected from atmospheric effects as well as from angular since they are calculated from nadir reflectance).

The study comprises 70 sites (belonging to 7 different ecosystems) uniformly distributed over Tuscany (Central Italy). Since the dependence on ground meteorological data cannot be completely avoided at present, our main goal is to propose an operational methodology for the characterization of the inter-annual variations of ϵ due to summer water stress which combines meteorological and MODIS data. The main advantage of this methodology is that it relies on existing series of meteorological data characterizing each site, which are ‘‘updated’’ for the actual conditions using MODIS spectral indices.

2. Materials and methods

2.1. Study area

Tuscany (9°–12° E, 42°–44° N) has extremely heterogeneous morphological and land cover features. Its climate ranges from Mediterranean to temperate warm or cool following the altitudinal and latitudinal gradients and the distance from the sea (Rapetti and Vittorini, 1995). Forests cover about half of the region and are mostly placed in the inner hilly and mountainous areas. The dominant forest species are various oaks, both evergreen (*Quercus ilex* L.) and deciduous (*Quercus pubescens* Willd. and *Quercus cerris* L.), Mediterranean pines (*Pinus pinaster* Ait., *Pinus pinea* L.), chestnut (*Castanea sativa* Mill.), beech (*Fagus sylvatica* L.) and spruce (*Abies alba* Mill.). Mediterranean macchia is widespread in the most arid zones near the coast and in the islands.

2.2. Data

2.2.1. Ancillary data

The spatial distribution of Tuscany forests was derived from the map of Arrigoni et al. (1998). The 18 classes reported by this map were grouped, following eco-physiological criteria, into the seven forest types (FTs) shown in Table 1. Within each FT ten sites of at least 1 km² were selected as representative of relevant eco-climatic conditions (Maselli et al., 2009b) paying attention to avoid discontinuities or boundary areas. The spatial distribution of the 70 sites

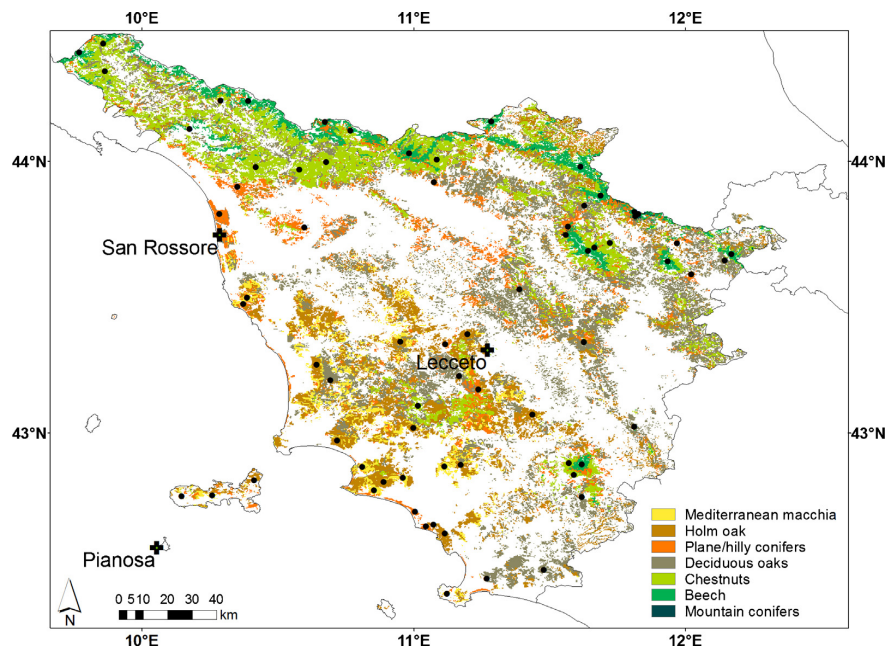


Fig. 1. Spatial distribution of the selected 70 sites (dots) belonging to 7 forest types (see legend), and locations of the three EC towers: San Rossore, Lecceto and Pianosa (crosses) over a land-use map.

is shown in Fig. 1 over the aforementioned land-use map (Arrigoni et al., 1998).

2.2.2. Meteorological data

Daily minimum and maximum temperatures and precipitation were derived from the measurements of existing weather stations spread all over Tuscany for the years 2001–2010. In particular, daily maximum and minimum temperatures and daily total precipitation were collected from 139 and 179 stations, respectively.

These daily meteorological data were extended to each test site by applying the DAYMET interpolation procedure (Thornton et al., 1997). PAR was computed as the 46.4% of the daily global solar radiation (Iqbal, 1983), which was estimated using the MT-CLIM model (Thornton et al., 2000).

2.2.3. Flux Tower data

Forest carbon fluxes were measured by the eddy covariance technique using three towers belonging to the FLUXNET international network (<http://fluxnet.ornl.gov/fluxnet/index.cfm>). Table 2 summarizes the main environmental features of the three tower sites, together with relevant measurement periods. A more complete description of these sites is provided in Chiesi et al. (2011).

Fig. 2 shows the thermo-pluviometric diagrams of the three sites computed from interpolated data over 10 years (2001–2010). The dry season – considered as the period when temperature in °C exceeds twice rainfall in mm (Emberger et al., 1963) – lasts about

2 months at San Rossore, 3 months at Lecceto and 4 months at Pianosa, being July the driest month in all cases.

The data of net carbon ecosystem exchange (*NEE*) measured by the three towers were processed according to current state of the art methodology (Papale et al., 2006). Quality analysis was performed according to the stationarity and integral turbulence tests (Foken and Wichura, 1996). Gap filling and flux partitioning were applied to derive *GPP* following the procedure described in Appendix A of Reichstein et al. (2005).

Some of the *GPP* values obtained were noisy and incorrect, particularly for the two towers of Lecceto and Pianosa. These data series were therefore cleaned using more restrictive criteria, i.e. removing all daily values that were unrealistic (negative and corresponding to excessively high ϵ) or dubious (derived from incomplete half-hour measurements).

2.2.4. Satellite data

Moderate Resolution Imaging Spectroradiometer (MODIS) images collected by both TERRA and AQUA sensors were obtained from the NASA Land Processes Distributed Active Archive Center (LP DAAC) (<https://lpdaac.usgs.gov/>) as 10-degree tiles in a sinusoidal projection. The present work uses the Nadir BRDF Adjusted Reflectance (NBAR) from the MODIS MCD43 product for estimating the spectral indices. As deriving BRDF requires merging multiple looks at each pixel, the necessary products were provided every 16 days, with 500 m spatial resolution. All images of the tile

Table 2
Main environmental features of the Tuscany forest sites where the three eddy covariance towers are located.

Study site	Geographical position	Topographic situation and elevation (m)	FT	Measurement period
San Rossore	43.73° N 10.28° E	Flat coastal area (10 m)	3	2000–2005
Lecceto	43.30° N 11.27° E	Hilly inner area (320 m)	2	2006–2008
Pianosa	42.58° N 10.05° E	Flat island (2 m)	1	2007–2008

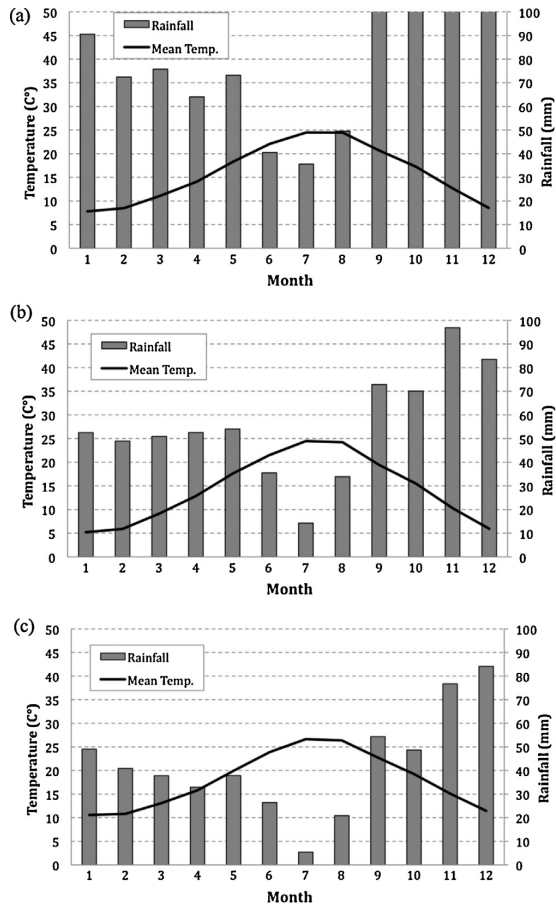


Fig. 2. Thermo-pluviometric diagrams of the three tower sites ((a) San Rossore; (b) Lecceto; (c) Pianosa) computed over 10 years (2001–2010).

corresponding to Central Italy were downloaded from the aforementioned NASA website for the period 2001–2010.

2.3. Methods

2.3.1. Time series

All the spectral values for all the variables listed below were extracted from the 500-m pixel containing each EC tower or centered at each of the 70 forest sites. Daily values for all these variables were obtained by means of an interpolation of the original series, which were slightly smoothed to avoid noisy and spurious data. A loess filter (grade 1) was used in all the cases, except for the f_{APAR} series. In this case, the MODIS BRDF parameters quality flag (MCD43A2) information was taken into account in order to compute the weights associated in the loess filter to remove noise. Such a noise results from cloud, ozone, dust, and other aerosols that generally decrease NIR reflectance and leads to spurious drops in the data (see Moreno et al., 2012 for more details). After applying the loess filter, the upper envelope of the f_{APAR} series is obtained.

The spectral indices selected as ‘water content indices’ (here denoted as ‘water content spectral indices’, WCSIs) have been obtained using the nadir reflectance in NIR band 2 (centered at 850 nm) and in the SWIR bands 5 (1240 nm), 6 (1635 nm) and 7

(2130 nm), according to Gao (1996), Fensholt and Sandholt (2003), and Rubio et al. (2006), Trombetti et al. (2008), respectively:

$$NDWI = \frac{b_2 - b_5}{b_2 + b_5} \quad (1.a)$$

$$SIWSI = \frac{b_2 - b_6}{b_2 + b_6} \quad (1.b)$$

$$NDI7 = \frac{b_2 - b_7}{b_2 + b_7} \quad (1.c)$$

In the expressions above, b_i indicates the nadir reflectance in MODIS spectral band i .

Moreover, the Renormalized Difference Vegetation Index

$$RDVI = \frac{b_2 - b_1}{(b_2 + b_1)^{1/2}} \quad (2)$$

was computed from NIR and red reflectances, b_2 and b_1 , respectively. In order to reduce external factors such as the canopy background (Roujean and Bréon, 1995), for an optimal geometry ($\theta_s = 45^\circ$, $\theta_v = 60^\circ$, $\varphi = 0^\circ$) in the principal plane, a $RDVI_{opt}$ was calculated with the NIR and red reflectances obtained using k_0 , k_1 and k_2 from the BRDF product (MCD43A1).

The f_{APAR} was then computed as

$$f_{APAR} = 1.81RDVI_{opt} - 0.21 \quad (3)$$

For the three eddy covariance tower sites shown in Table 2, daily ε was derived from the tower GPP, the PAR and the f_{APAR} using the Monteith’s approach:

$$\varepsilon = \frac{GPP}{f_{APAR}PAR} \quad (4)$$

The water stress index C_{ws} was then calculated from interpolated meteorological data as (Maselli et al., 2009a):

$$C_{ws} = 0.5 \left[1 + \left(\frac{AET}{PET} \right) \right] \quad (5)$$

where AET and PET are actual and potential evapotranspiration, respectively. For the sake of simplicity, AET is assumed to be equal to the precipitation when this is lower than PET . Consequently, C_{ws} can vary between 0.5 (when strong water shortage reduces photosynthesis to half of its potential value) to 1 (when there is no water shortage and photosynthesis reduction). Both AET and PET are computed over two months (preceding the day of GPP estimation). This time interval allows the estimation of the short-term effect of water stress without the need for computing a complete site water balance including all input and output terms and soil retention processes.

2.3.2. Steps of the methodology

The main goal of our analysis is to evaluate the capability of the spectral indices given in Eqs. (1.a)–(1.c) to track the inter-annual variations of ε due to water stress. As the number of eddy-covariance towers in the region is low, an indirect approach is adopted to test the WCSIs.

First, light use efficiency data from the three tower sites (three Tuscan forest sites where the EC towers are located) are analyzed to check the usefulness of the C_{ws} to track the inter-annual variations of ε related to water stress. Statistical correlations between them are carried out. If they confirm that C_{ws} accounts for the reduction of ε during summer water stress, then the WCSI on the other 70 sites is analyzed.

Our assumption is that if ε is correlated with C_{ws} , and C_{ws} is highly correlated with WCSI, then ε is correlated with WCSI (i.e., the transitive property is satisfied):

$$\varepsilon \leftrightarrow C_{ws} \text{ and } C_{ws} \leftrightarrow WCSI \Rightarrow \varepsilon \leftrightarrow WCSI \quad (6)$$

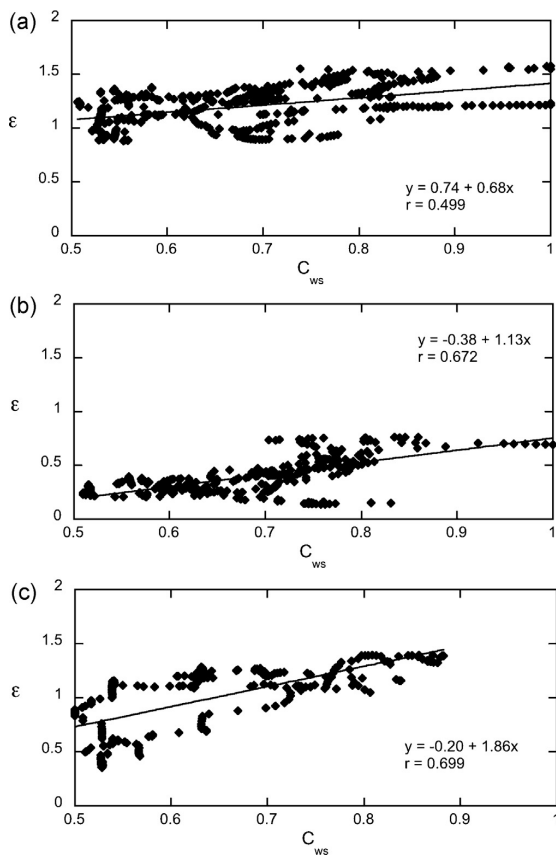


Fig. 3. Correlation between the light use efficiency and the water stress index C_{ws} for the three tower sites: (a) San Rossore, (b) Lecceto and (c) Pianosa.

Thus, the WSCI can be used to account for ε summer variations.

Finally, a procedure will be proposed and tested which uses the water content spectral indices to characterize forest water status and estimate GPP .

3. Results and discussion

3.1. EC Tower sites

A Pearson's correlation analysis was applied to test the significance of the inter-annual correlation between ε and C_{ws} . Only the late spring–summer period, compressed between the DOY (day of year, or Julian date) 121 and 243 (i.e., 1 May–31 August), was taken into account. As shown by the thermo-pluviometric diagrams of the three tower sites (Fig. 2) this 4-month period assures the inclusion of the drought period around this three study sites. Fig. 3 shows the correlations between the light use efficiency derived using the tower data and the f_{APAR} from MODIS (according to Eq. (3)) and the C_{ws} . The three correlations are statistically significant ($p < 0.01$), but their values are not high (from 0.499 to 0.699). The relatively high dispersion of these relationships can be partly attributed to the numerous error sources that may affect both the ground meteorological data (e.g. measurement and interpolation uncertainty, particularly concerning rainfall, which is most strongly variable in space and time) and the ε estimates (mainly inaccuracy in the tower GPP data and in the remotely sensed f_{APAR} estimates).

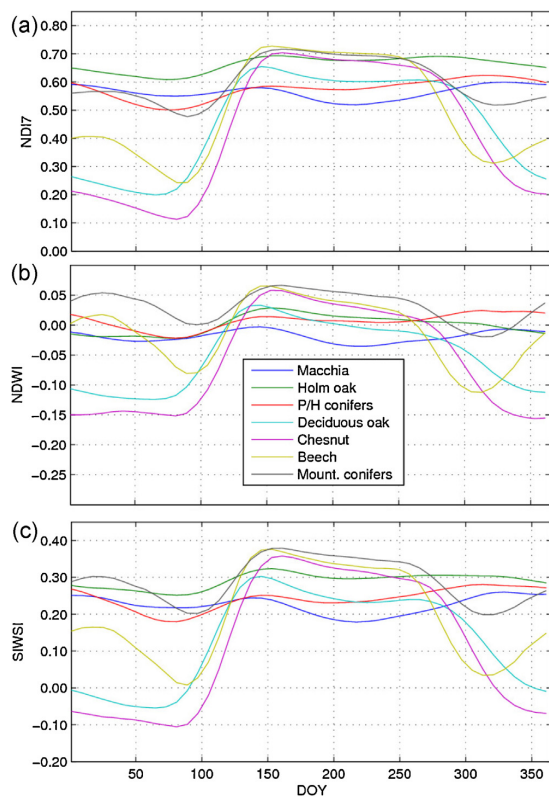


Fig. 4. Typical profile (intra-annual variation) of the WCSIs: (a) ND17, (b) NDWI, and (c) SIWSI. The curves correspond to different FTs as shown in the legend.

In general, correlation values increase as the site aridity increases. The San Rossore site, the tallest and densest forest, shows the weakest correlation, probably due to resistance of the pine trees to water stress. The influence of soil depth on the water holding capacity is another factor to be taken into account. The soil depth in San Rossore is actually higher than that of the other sites (1 m versus 30–40 cm). This fact probably enhances the lower dependence of ε on water stress due to the more humid climate. Intermediate correlation is found for Lecceto, which is a site with semi-arid climate, covered by evergreen Mediterranean oaks. The highest correlation is found in Pianosa, where the predominant cover is Mediterranean macchia, very sensitive to water stress because of the short root system, which does not allow this vegetation to obtain water from deeper and wetter soil.

These results confirm those of Maselli et al. (2009a), who showed that C_{ws} is a good indicator of summer ε inter-annual variations due to water stress in dry Mediterranean forest ecosystems. On the basis on these observations, we consider that C_{ws} can be used to assess the information content of the WSCI on summer water stress.

3.2. 70 sites

Fig. 4 shows the typical intra-annual variation profiles of the WCSIs for each vegetation class. These profiles were obtained averaging the 10-year data series and the 10 sites corresponding to each FT.

The profiles are descriptive of vegetation phenology and can be divided into two groups: one for the deciduous FTs (deciduous

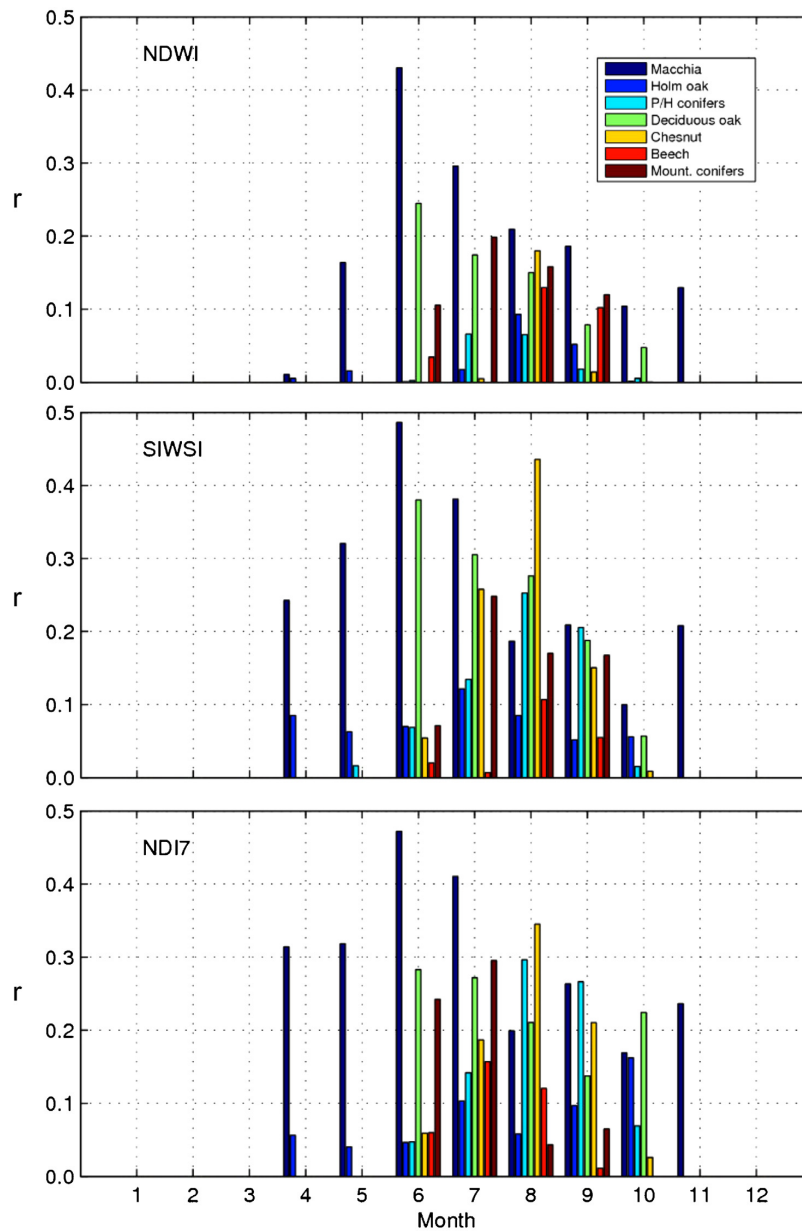


Fig. 5. Coefficient r of the inter-annual correlation between each one of the three WCSIs and C_{ws} calculated for every month and every FT.

oaks, chestnut and beech) and another one for the rest of them (macchia, Holm oak, plain/hilly conifers and mountain conifers). The two groups show the highest differences during winter, where the fractional cover and the photosynthetic activity of the deciduous classes is very low. Most of the profiles show a weak fall around DOY 200 (middle of July) that can be attributed to both water content and pigment content losses during summer drought. Macchia is the vegetation class that shows more intensively this effect.

The WCSIs are supposed to be affected by the water stress that inhibits the photosynthetic activity of the plant. Fig. 5 shows the

monthly inter-annual correlation coefficients (r) between these spectral indices and C_{ws} .

These correlations are low to moderate, and r is not statistically significant ($p > 0.05$). The highest values are reached during the late spring–summer period and mainly when considering *SIWSI* and *NDI7*. The correlations are lower when using the *NDWI*. The lowest values are found for holm oak and mountain conifers, and the highest correspond to macchia, deciduous oak and chestnut. As a consequence, the direct use of these spectral indices does not seem to be recommended during the whole period of time with any type of vegetation. These behaviors might be attributed to

Table 3

Mean value (and standard deviation) of the correlation coefficient r between each WCSI and C_{ws} for each FT and considering only the late spring–summer period (DOY 121–243). FT codes: (1) Macchia, (2) Holm oak, (3) Plane/Hilly conifers, (4) Deciduous oaks, (5) Chestnut, (6) Beech, and (7) Mountain conifers.

	FT						
	1	2	3	4	5	6	7
NDWI	0.49 (0.24)	0.16 (0.28)	0.09 (0.26)	0.21 (0.32)	−0.16 (0.35)	−0.22 (0.25)	0.20 (0.30)
SIWSI	0.54 (0.23)	0.23 (0.25)	0.29 (0.24)	0.30 (0.30)	0.03 (0.32)	−0.10 (0.27)	0.30 (0.29)
NDI7	0.52 (0.22)	0.21 (0.29)	0.26 (0.33)	0.23 (0.28)	−0.15 (0.32)	−0.10 (0.40)	0.25 (0.24)

different factors: the existence of deciduous and evergreen species; leaf area index (LAI) changes that also affect the SWIR reflectance; varying background (from bare soil to snow cover) effects on canopy reflectance, mainly in mountain conifers, which are located at high altitude; etc. As discussed below, it is also possible that the WCSIs respond to the water stress later than the C_{ws} does.

In order to quantify the redundancy among the information supplied by the WCSIs, the inter-annual correlations for the different forest types have been analyzed (Fig. 6). It is observed that *SIWSI* and *NDI7* are the most redundant, whereas *NDWI* seems to offer additional information mainly for macchia, holm oak, plain/hilly conifers during winter, chestnut, beech and mountain conifers during summer.

In order to identify the WCSI most sensitive to water stress, the inter-annual correlations between the mean values for each of them and C_{ws} were calculated for each site, considering only the late spring–summer period. The 3×70 correlation values are represented in Fig. 7. A considerable variability is observed among the sites belonging to the same vegetation class. Thus, the mean value (and standard deviation) for each class, considering only the summer period, is given in Table 3.

The correlation coefficient between *SIWSI* and C_{ws} is higher than for the other two WCSIs, which in some cases even show negative values. The correlations between any WCSI and C_{ws} are the lowest for FTs 5 (Chestnut) and 6 (Beech). As mentioned above, the intra-FT variance is considerable. This may be partly attributed to the possible compositional heterogeneity of the FTs at the MODIS pixel scale (500 m).

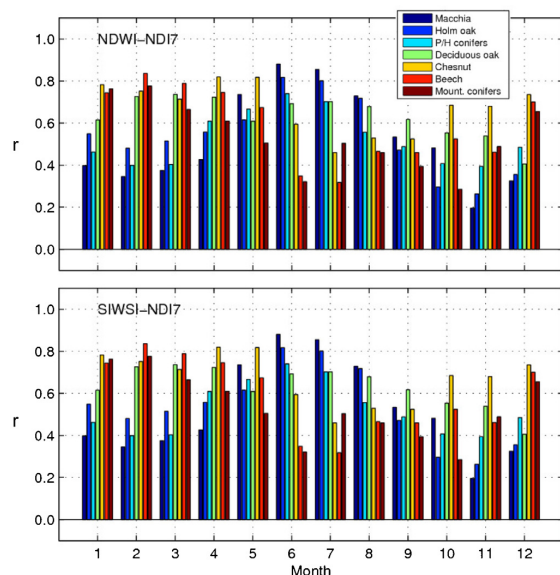


Fig. 6. Inter-annual correlations between pairs of WCSIs calculated using monthly averages for each vegetation class.

In most cases, the correlations improve when the time period is reduced to July and August (DOY 182–243, see Fig. 2), the two months mainly affected by the drought. In this case, C_{ws} was also calculated for a two-month period, but previous to that in which the spectral indices were obtained. A two-month moving window (which was shifted at 1-day steps) was considered to find out the date when the influence of water budget was more significant in each of the spectral indices. For FTs 1 and 2 (Macchia and Holm oak) the water budget calculated in April–May is the most influential on the spectral indices calculated in July–August, whereas the other FTs are more influenced by the value of C_{ws} during May–June or even June–July (in the case of FT 5, Chestnut). These inter-annual correlations basically confirm an observation reported in Maselli et al. (2013, submitted). The dependence of summer *NDVI* on spring meteorology (water balance) is directly related to the aridity of the different ecosystems (more than to the forest types). Arid (xeric) ecosystems (macchia, holm oak, etc.) show much higher correlation values than the more humid ones, whose inter-annual variations may be mostly attributed to other factors (temperature, radiation, timing, etc.). The results obtained in this work with more spectral indices are in agreement with previous results.

3.3. Operational approach to combine meteorological and spectral data

The results obtained in the previous section indicate that, although some of the correlations between WCSIs and C_{ws} are significant for some FTs, the selection of an appropriate spectral index to track the inter-annual variations of the light use efficiency for all FTs is challenging. The evaluation of the effect of water stress on ϵ apparently requires meteorological data in addition to remote sensing data (at least, in the case of the indices considered). Nevertheless, we can take advantage of our previous findings using a hybrid methodology that combines the climatic information of C_{ws} and the seasonal WCSI variation.

Firstly, by averaging all the years where meteorological data are available, a typical, mean intra-annual profile $\overline{C_{ws}}$ and its standard deviation C_{ws} are calculated. Secondly, a spectral index anomaly ΔSI is defined as

$$\Delta SI_{ji} = \frac{SI_{ji} - \overline{SI}_i}{std(SI)_i} \quad (7)$$

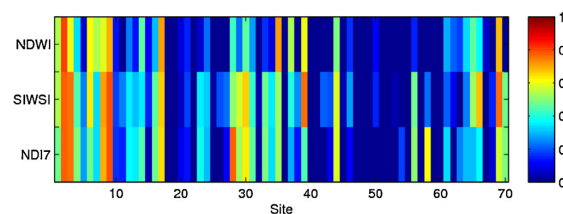


Fig. 7. Correlation coefficient values between each WCSIs and C_{ws} for each one of the 70 sites.

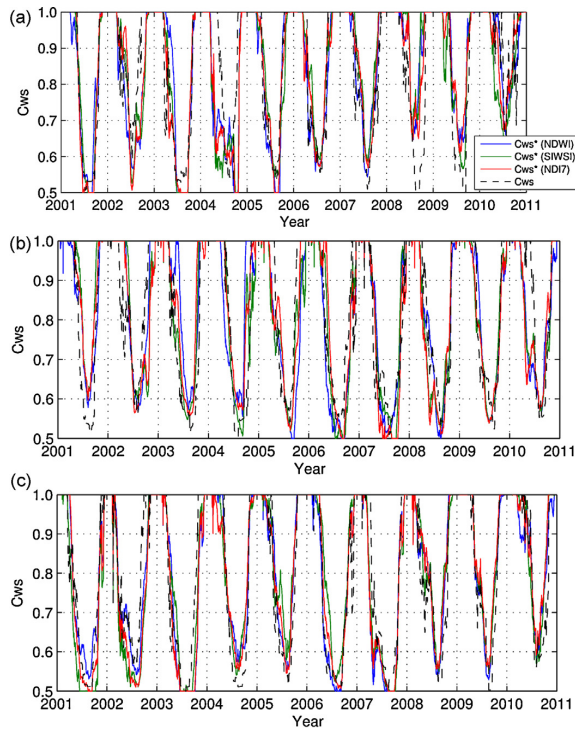


Fig. 8. Original C_{ws} time series and its three modifications using Eq. (8) and each one of the three WCSIs for the tower sites: (a) San Rossore, (b) Lecceto, and (c) Pianosa.

where the subscripts i and j refer to the DOY and the year, respectively.

Finally, the index anomaly is used to calculate a modified profile for C_{ws} , which is denoted here by C_{ws}^* :

$$(C_{ws}^*)_{ji} = (\Delta SI_{ji})std(C_{ws})_i + \overline{(C_{ws})}_i \quad (8)$$

Therefore, taking into account the SI derived from satellite data, we could update C_{ws} for each year, even when no meteorological data were available.

To validate this procedure, the modified C_{ws} time series for the three tower sites using the three WCSIs ($NDWI$, $SIWSI$ and $NDI7$) were obtained. They are shown in Fig. 8 (blue, green and red lines, respectively), where the original C_{ws} time series is also represented (black, dashed line).

As shown in Fig. 8, the inter-annual variations captured by the spectral indices allow us to track the inter-annual C_{ws} variability, which is a consequence of the water budget registered in each year. The global correlation coefficients between the original C_{ws} and the modified C_{ws}^* data using the remotely sensed spectral indices are shown in Table 4 for late spring–summer period (with their corresponding root mean square errors, RMSE). These values have

Table 4
Correlation coefficients between the original C_{ws} and its modified values C_{ws}^* using each of the WCSIs for the three tower sites, and RMSE values. All the correlations are statistically significant ($p < 0.01$).

	San Rossore		Lecceto		Pianosa	
	r	RMSE	r	RMSE	r	RMSE
C_{ws}^* (NDWI)	0.77	0.10	0.76	0.12	0.82	0.06
C_{ws}^* (SIWSI)	0.68	0.12	0.80	0.10	0.76	0.09
C_{ws}^* (NDI7)	0.72	0.11	0.76	0.11	0.84	0.07

Table 5
Correlation coefficients (and RMSE, in $g\ m^{-2}\ day^{-1}$) between the tower GPP and the GPP estimated at the corresponding tower sites using the improved C-Fix model driven by different water stress factors. All the correlations are statistically significant ($p < 0.01$).

	San Rossore		Lecceto		Pianosa	
	r	RMSE	r	RMSE	r	RMSE
C_{ws}	0.80	0.9	0.41	1.1	0.64	0.8
C_{ws}^* (NDWI)	0.80	0.9	0.40	0.9	0.65	0.8
C_{ws}^* (SIWSI)	0.80	0.9	0.40	0.9	0.64	0.8
C_{ws}^* (NDI7)	0.80	0.9	0.39	1.0	0.65	0.8

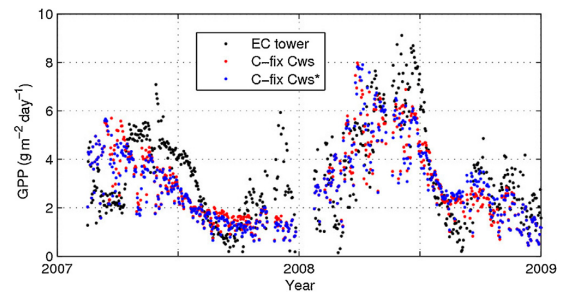


Fig. 9. GPP data series of Pianosa measured by the EC tower and estimated using the improved C-Fix model with the original C_{ws} and the $SIWSI$ -modified C_{ws}^* .

been obtained using a leave-one-out cross-validation procedure for the ten years examined. The rather high correlations found and confirm the validity of modified C_{ws}^* as a proxy of water stress that does not require concurrent information on meteorological data.

3.4. Methodology validation using the three flux tower sites

A final validation test of the methodology proposed to combine meteorological and spectral data makes use of both C_{ws} and C_{ws}^* to drive modified C-Fix (Maselli et al., 2009a). C-Fix is a Monteith type model that was proposed by Veroustraete et al. (2002) to predict forest GPP and improved by Maselli et al. (2009a) to adapt it to water-limited Mediterranean environments through the insertion of the C_{ws} . The model GPP (using the different water stress factors) was tested against the tower GPP for the three tower sites. Table 5 summarizes the correlation coefficients and the RMSE, which are similar for the three tower sites. These results confirm that when the actual C_{ws} is not available, it can be estimated as a value C_{ws}^* calculated using Eq. (8) from the typical C_{ws} profile of the site and the MODIS spectral index.

As an example, Fig. 9 shows three GPP profiles for Pianosa, the most arid site, and then the most sensitive to the water stress effect on the light use efficiency. They correspond to the GPP tower measurements, and the GPP calculated using the improved C-Fix model with C_{ws} and the $SIWSI$ -modified C_{ws}^* . The overall performance of improved C-Fix is rather good and the GPP modeled using C_{ws} and C_{ws}^* show a similar profile.

4. Conclusions

The estimation of vegetation GPP from remote sensing data is an old problem that, so far, has found only a partial solution. Some approaches are purely statistical and require calibration for each examined site/ecosystem (Yang et al., 2007), which presents several drawbacks from an operational point of view. Most methods rely on Monteith's approach, which combines estimates of solar radiation, f_{APAR} and the light use efficiency, ϵ . Both PAR and f_{APAR} can be derived from remotely sensed data (Moreno et al., 2013;

Martínez et al., 2013). Nevertheless, the estimation of the last term, which is variable depending on environmental constraints mainly related to thermal and water stresses, relies basically on meteorological data. The attempts to characterize ε using only remote sensing information have given controversial results (Coops et al., 2010).

In particular, in arid and semi-arid environments the inter-annual variations of ε are mainly due to water stress. This stress has been modeled by different factors, generally based on meteorological data (such the vapor pressure deficit, the aforementioned C_{ws}). The spectral index *PRI* (Gamon et al., 1990, 1992), sensitive to the xanthophyll inter-conversion cycle that takes place in the presence of water stress, was a promising choice to quantify ε from spectral data. The performance of this index, however, is good when measured *in situ* at leaf and plant levels, but becomes problematic when using MODIS data at ecosystem level (Moreno et al., 2012). In this context, the use of spectral indices indicative of leaf water content has been rarely considered.

We have shown that a two-step approach yields promising results. First, the information content of a water stress factor (C_{ws}) on the response of ε to the water stress has been assessed for three flux tower sites in Central Italy. Second, a statistical analysis has evidenced the inter-annual correlations between C_{ws} and three leaf water content spectral indices (*NDWI*, *SIWSI* and *NDI7*) for seven forest types.

We have shown that these three water content indices are moderately informative on C_{ws} inter-annual variations, particularly in the most arid areas. *SIWSI* presents the best results in most of the cases. The correlations between WCSIs and C_{ws} are rather low for Chestnut and Beech, which are generally placed in most humid areas. Thus, replacing C_{ws} by such indices is complicated due to the varying leaf phenology and seasonal evolution of some forest types. These results seem to indicate that the water stress dependence of ε cannot be estimated only from this type of spectral indices.

To solve this problem, an operational methodology that optimally combines meteorological and spectral data has been presented and validated. This methodology uses historical series of meteorological data that are updated to characterize actual vegetation conditions by means of WCSIs and provides good accuracy for the three test sites examined. In fact, the correlation coefficients between the original C_{ws} and the modified C_{ws}^* data using the water content spectral indices are around 0.7–0.8. This confirms that modified C_{ws}^* can be used as a proxy of water stress without the need for concurrent meteorological information.

It must be recognized, however, that the methodology proposed is only a partial solution that circumvents the problem for operational monitoring applications. Further efforts to track the temporal variations of ε due to water stress using only remotely sensed data are required.

Acknowledgments

This research was supported by the Spanish Ministry of Economy and Competitiveness (Project RESET CLIMATE CGL2012-35831) and by LSA SAF (EUMETSAT). Special thanks are due to the anonymous referees for their suggestions, which greatly improved the quality of the paper.

References

- Allard, V., Ourcival, J.M., Rambal, S., Joffre, R., Rocheteau, A., 2008. Seasonal and annual variation of carbon exchange in an evergreen Mediterranean forest in southern France. *Global Change Biol.* 14, 714–725. <http://dx.doi.org/10.1111/j.1365-2486.2008.01539.x>.
- Arrigoni, P.V., Raffaelli, M., Rizzotto, M., Selvi, F., Vicini, D., Lombardi, L., Foggi, B., Melillo, C., Benesperi, R., Ferretti, G., Benucci, S., Turrini, S., di Tommaso, P.L., Signorini, M., Bargelli, E., Miniati, U., Farioli, C., de Dominicis, V., Casini, S., Chiarucci, A., Tomei, P.E., Ansaldo, M., Maccioni, S., Guazzi, E., Zocco Pisana, L., Cenerini, A., Dell'Olmo, L., Menicagli, E., 1998. *La vegetazione forestale. Serie Boschi e Macchie di Toscana, Regione Toscana, Giunta regionale.*
- Bolle, H.J., Eckardt, M., Koslowsky, D., Maselli, F., Meliá Miralles, J., Menenti, M., Olesen, F.S., Petkov, L., Rasool, I., Van de Griend, A., 2006. *Mediterranean land surface processes assessed from space. In: Regional Climate Studies, vol. XXVIII. Springer Series.*
- Carter, G.A., 1991. Primary and secondary effects of water content on the spectral reflectance of leaves. *Am. J. Bot.* 78, 916–924.
- Ceccato, P., Gobron, N., Flasse, S., Pinty, B., Tarantola, S., 2002. Designing a spectral index to estimate vegetation water content from remote sensing data: part 1. Theoretical approach. *Remote Sens. Environ.* 82, 188–197.
- Chiesi, M., Fibbi, L., Genesio, L., Gioli, B., Magno, R., Maselli, F., Moriondo, M., Vaccari, F., 2011. Integration of ground and satellite data to model Mediterranean forest processes. *Int. J. Appl. Earth Observ. Geoinform.*, 504–515.
- Clevers, J.G.P.W., 2008. The derivation of a simplified reflectance model for the estimation of leaf area index. *Remote Sens. Environ.* 25, 53–69.
- Coops, N.C., Hilker, T., Hall, F.G., Nichol, C.J., Drolet, G.G., 2010. Estimation of light use efficiency of terrestrial ecosystems from space: a status report. *Bioscience* 60, 788–797.
- Danson, F.M., Steven, M.D., Malthus, T.J., Clark, J.A., 1992. High-spectral resolution data for determining leaf water content. *Int. J. Remote Sens.* 13, 461–470.
- Danson, F.M., Bowyer, P., 2004. Estimating live fuel moisture content from remotely sensed reflectance. *Remote Sens. Environ.* 92, 309–321.
- Drolet, G.G., Middleton, E.M., Huemmrich, K.F., Hall, F.G., Amiro, B.D., Barr, A.G., Black, T.A., McCaughey, J.H., Margolis, H.A., 2008. Regional mapping of gross light-use efficiency using MODIS spectral indices. *Remote Sens. Environ.* 112, 3064–3078.
- Emberger, L., Gaussen, H., Kassas, M., De Philippis, A., 1963. *Carte bioclimatique de la zone Méditerranéenne*. 1. UNESCO/FAO, Paris/Rome, pp. 5000000.
- Fensholt, R., Sandholt, I., 2003. Derivation of a shortwave infrared water stress index from MODIS near- and shortwave infrared data in a semiarid environment. *Remote Sens. Environ.* 87, 111–121.
- Field, C.B., Randerson, J.T., Malmstrom, C.M., 1995. Global net primary production: combining ecology and remote sensing. *Remote Sens. Environ.* 51, 74–88.
- Filella, I., Porcar-Castell, A., Munne-Bosch, S., Back, J., Garbulsky, M.F., Peñuelas, J., 2009. PRI assessment of long-term changes in carotenoids/chlorophyll ratio and short-term changes in de-epoxidation of the xanthophylls cycle. *Int. J. Remote Sens.* 30, 4443–4455.
- Foken, T., Wichura, B., 1996. Tools for quality assessment of surface based flux measurements. *Agric. For. Manage.* 78, 83–105.
- Gamon, J.A., Field, C.B., Bilger, W., Björkman, O., Fredeen, A., Peñuelas, J., 1990. Remote sensing of the xanthophyll cycle and chlorophyll fluorescence in sunflower leaves and canopies. *Oecologia* 85, 1–7.
- Gamon, J.A., Peñuelas, J., Field, C.B., 1992. A narrow-waveband spectral index that tracks diurnal changes in photosynthetic efficiency. *Remote Sens. Environ.* 41, 35–44.
- Gao, B., 1996. NDWI—a normalized difference water index for remote sensing of vegetation liquid water from space. *Remote Sens. Environ.* 58, 257–266.
- Garbulsky, M.F., Peñuelas, J., Papale, D., Filella, I., 2008. Remote estimation of carbon dioxide uptake by a Mediterranean forest. *Global Change Biol.* 14, 2860–2867.
- Garbulsky, M., Peñuelas, J., Papale, D., Ardó, J., Goulden, M.L., Kiely, G., Richardson, A.D., Rotenberg, E., Veenendaal, E.M., Filella, I., 2009. Patterns and controls of the variability of radiation use efficiency and primary productivity across terrestrial ecosystems. *Global Ecol. Biogeogr.* 19, 253–267. <http://dx.doi.org/10.1111/j.1466-8238.2009.00504.x>.
- Goerner, A., Reichstein, M., Tomelleri, E., Hanan, N., Rambal, S., Papale, D., et al., 2011. Remote sensing of ecosystem light use efficiency with MODIS-based PRI. *Biogeosciences* 8, 189–202.
- Hall, F.G., Hilker, T., Coops, N.C., Lyapustin, A., Huemmrich, K.F., Middleton, E., et al., 2008. Multi-angle remote sensing of forest light use efficiency by observing PRI variation with shadow fraction. *Remote Sens. Environ.* 112, 3201–3211.
- Hansen, M.C., DeFries, R.S., Townshend, J.R.G., Sohlberg, R., 2000. Global land cover classification at 1 km spatial resolution using a classification tree approach. *Int. J. Remote Sens.* 21, 1331–1364.
- Heinsch, F.A., Reeves, M., Votava, P., Kang, S., Milesi, C., Zhao, M., Glassy, J., Jolly, W.M., Loehman, R., Bowker, C.F., Kimball, J.S., Nemani, R.R., Running, S.W., 2003. User's Guide GPP and NPP (MOD17A2/A3) Products NASA MODIS Land Algorithm, Version 2.0 December 2, 2003 www.ntsg.umd.edu/modis/
- Iqbal, M., 1983. *An Introduction to Solar Radiation*. Academic Press, New York, USA.
- Martínez, B., Camacho, F., Verger, A., García-Haro, F.J., Gilabert, M.A., 2013. Inter-comparison and quality assessment of MERIS, MODIS and SEVIRI FAPAR products over the Iberian Peninsula. *Int. J. Appl. Earth Observ. Geoinform.* 21, 463–476.
- Maselli, F., Papale, D., Puletti, N., Chirici, G., Corona, P., 2009a. Combining remote sensing and ancillary data to monitor the gross productivity of water-limited forest ecosystems. *Remote Sens. Environ.* 113, 657–667.
- Maselli, F., Chiesi, M., Moriondo, M., Fibbi, L., Bindi, M., Running, S.W., 2009b. Modelling the forest carbon budget of a Mediterranean region through the integration of ground and satellite data. *Ecol. Model.* 220, 330–342.
- Maselli, F., Cherubini, P., Chiesi, M., Gilabert, M.A., Moreno, A., Teobaldelli, M., Tognetti, R., 2013. Start of the dry season as a main determinant of inter-annual Mediterranean forest production variations (submitted for publication).
- Monteith, J.L., 1972. Solar radiation and productivity in tropical ecosystems. *J. Appl. Ecol.* 9, 747–766.

- Moreno, A., Maselli, F., Gilabert, M.A., Chiesi, M., Martínez, B., Seufert, G., 2012. Assessment of MODIS imagery to track light-use efficiency in a water limited Mediterranean pine forest. *Remote Sens. Environ.* 123, 359–367.
- Moreno, A., Gilabert, M.A., Camacho, F., Martínez, B., 2013. Validation of daily global solar irradiation images from MSG over Spain. *Renew. Energy* 60, 332–342. <http://dx.doi.org/10.1016/j.renene.2013.05.019>.
- Naumann, J.C., Bisset, S.N., Young, D.Y., Edwards, J., Anderson, J.E., 2010. Diurnal patterns of photosynthesis, chlorophyll fluorescence, and PRI to evaluate water stress in the invasive species, *Elaeagnus umbellata* Thumb. *Trees* 24, 237–245.
- Papale, D., Reichstein, M., Aubinet, M., Canfora, E., Bernhofer, C., Kutsch, W., Longdoz, B., Rambal, S., Valentini, R., Vesala, T., Yakir, D., 2006. Towards a standardized processing of Net Ecosystem Exchange measured with eddy covariance technique: algorithms and uncertainty estimation. *Biogeosciences* 3, 571–583.
- Potter, C.S., Randerson, J.T., Field, C.B., Matson, P.A., Vitousek, P.M., Mooney, H.A., Klooster, S.A., 1993. Terrestrial ecosystem production: a process model based on global satellite and surface data. *Global Biogeochem. Cycles* 7 (4), 811–841.
- Rapetti, F., Vittorini, S., 1995. *Carta climatica della Toscana*. Pacini Editore, Pisa, Italy.
- Reichstein, M., Falge, E., Baldocchi, D., Papale, D., Aubinet, M., Berbigier, P., Bernhofer, C., Buchmann, N., Gilmanov, T., Granier, A., Grünwald, T., Havránková, K., Ilvesniemi, H., Janous, D., Knohl, A., Laurila, T., Lohila, A., Loustau, D., Matteucci, G., Meyers, T., Miglietta, F., Ourcival, J., Pumpanen, J., Rambal, S., Rotenberg, E., Sanz, M., Tenhunen, J., Seufert, G., Vaccari, F., Vesala, T., Yakir, D., Valentini, R., 2005. On the separation of net ecosystem exchange into assimilation and ecosystem respiration: review and improved algorithm. *Global Change Biol.* 11, 1424–1439.
- Roujean, J.L., Bréon, F.M., 1995. Estimating PAR absorbed by vegetation from bidirectional reflectance measurements. *Remote Sens. Environ.* 51, 375–384.
- Rubio, M.A., Riaño, D., Cheng, Y.B., Ustin, S.L., 2006. Estimation of canopy water content from MODIS using artificial neural networks trained with radiative transfer models. In: 6th EMS/6th ECAC Ljubljana, Slovenia, European Meteorological Society.
- Running, S.W., Nemani, R.R., Heinsch, F.A., Zhao, M., Reeves, M., Hashimoto, H., 2004. A continuous satellite-derived measure of global terrestrial primary production. *Bioscience* 54, 547–560.
- Sarlikioti, V., Driever, S.M., Marcellis, L.F.M., 2010. Photochemical reflectance index as a mean of monitoring early water stress. *Ann. Appl. Biol.* 157, 81–89.
- Suárez, L., Zarco-Tejada, P.J., Berni, J.A.J., González-Dugo, V., Fereres, E., 2009. Modelling PRI for water stress detection using radiative transfer models. *Remote Sens. Environ.* 113, 730–744.
- Tarantola, S., Ceccato, P., Gre, J., 2001. Detecting vegetation leaf water content using reflectance in the optical domain. *Remote Sens. Environ.* 77, 22–33.
- Thenot, F., Méthy, M., Winkel, T., 2002. The Photochemical Reflectance Index (PRI) as a water-stress index. *Int. J. Remote Sens.* 23, 5135–5139.
- Thornton, P.E., Running, S.W., White, M.A., 1997. Generating surfaces of daily meteorological variables over large regions of complex terrain. *J. Hydrol.* 190, 214–251.
- Thornton, P.E., Hasenauer, H., White, M.A., 2000. Simultaneous estimation of daily solar radiation and humidity from observed temperature and precipitation: an application over complex terrain in Austria. *Agric. For. Meteorol.* 104, 255–271.
- Trombetti, M., Riano, D., Rubio, M., Cheng, Y., Ustin, S., 2008. Multi-temporal vegetation canopy water content retrieval and interpretation using artificial neural networks for the USA. *Remote Sens. Environ.* 112, 203–215. <http://dx.doi.org/10.1016/j.rse.2007.04.013>.
- Veroustraete, F., Sabbe, H., Eerens, H., 2002. Estimation of carbon mass fluxes over Europe using the C-Fix model and Euroflux data. *Remote Sens. Environ.* 83, 376–399.
- Verstraeten, W., Veroustraete, F., Feyen, J., 2006. On temperature and water limitation of net ecosystem productivity: implementation in the C-Fix model. *Ecol. Model.* 199, 4–22. <http://dx.doi.org/10.1016/j.ecolmodel.2006.06.008>.
- Yang, F., Ichii, K., White, M.A., Hashimoto, H., Michaelis, A.R., Votava, P., Zhu, A.-X., Huete, A., Running, S.W., Nemani, R.R., 2007. Developing a continental-scale measure of gross primary production by combining MODIS and AmeriFlux data through Support Vector Machine approach. *Remote Sens. Environ.* 110, 109–122.
- Zarco-Tejada, P., Rueda, C., Ustin, S., 2003. Water content estimation in vegetation with MODIS reflectance data and model inversion methods. *Remote Sens. Environ.* 85, 109–124. [http://dx.doi.org/10.1016/S0034-4257\(02\)00197-9](http://dx.doi.org/10.1016/S0034-4257(02)00197-9).

Apéndice **E**

Error estimation

- Mean Bias Error (MBE).

$$MBE = \frac{1}{n} \sum_{i=1}^n e_i \quad (\text{E.1})$$

with

$$e_i = e_{si} - e_{ti} \quad (\text{E.2})$$

where e_{si} is the estimated value, e_{ti} is the measured/target value and n is the number of values.

- Root Mean Squared Error (RMSE).

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (e_i)^2} \quad (\text{E.3})$$

- Mean Absolute Error (MAE).

$$MAE = \frac{1}{n} \sum_{i=1}^n |e_i| \quad (\text{E.4})$$

- Relative Mean Bias Error (rMBE).

$$rMBE = \frac{MBE}{\bar{e}_t} \quad (\text{E.5})$$

where \bar{e}_t is the average of the measured values (e_{ti}).

- Relative Mean Absolute Error (rMAE).

$$rMAE = \frac{MAE}{\bar{e}_t} \quad (\text{E.6})$$

- Relative Root Mean Squared Error (rRMSE).

$$rRMSE = \frac{RMSE}{\bar{e}_t} \quad (\text{E.7})$$

- Coefficient of determination (R^2).

$$R^2 = 1 - \frac{\sum_{i=1}^n (e_{ti} - \bar{e}_t)^2}{\sum_{i=1}^n (e_{si} - \bar{e}_t)^2} \quad (\text{E.8})$$

The MBE is a signed measure of error that indicates whether the estimates of the model are biased. The RMSE is more sensitive than other measures to the occasional large error: the squaring process gives disproportionate weight to very large errors. If an occasional large error is not a problem in model selection, then MAE may be a more relevant criterion. R^2 is often described as the proportion of “variance” explained by the model.