Geosci. Model Dev. Discuss., 5, 3573–3620, 2012
www.geosci-model-dev-discuss.net/5/3573/2012/
doi:10.5194/gmdd-5-3573-2012

Geoscientific
Model Development
Discussions

ECOCLIMAP-II/Europe: a twofold database of ecosystems and surface parameters at 1-km resolution based on satellite information for use in land surface, meteorological and climate models

S. Faroux[1], A. T. Kaptué Tchuenté[1,*], J.-L. Roujean[1], V. Masson[1], E. Martin[1], and P. Le Moigne[1]

[1]CNRM-GAME (Météo-France, CNRS), URA1357, 42 avenue Gaspard Coriolis, 31057 Toulouse CEDEX, France
[*]now at: Geographic Information Science Center of Excellence (GIScCE), 1021 Medary Ave, Wecota Hall, Box 506B, Brookings SD 57007, USA

Received: 17 September 2012 – Accepted: 10 October 2012 – Published: 7 November 2012

Correspondence to: S. Faroux (stephanie.faroux@meteo.fr)

Published by Copernicus Publications on behalf of the European Geosciences Union.

Abstract

The overall objective of the present study is to introduce the new ECOCLIMAP-II database for Europe, which is an upgrade for this region of the former initiative, ECOCLIMAP-I, already implemented at global scale. The ECOCLIMAP programme is a dual database at 1-km resolution that includes an ecosystem classification and a coherent set of land surface parameters that are primarily mandatory in meteorological modelling (notably leaf area index and albedo). Hence, the aim of this innovative physiography is to enhance the quality of initialisation and impose some surface attributes within the scope of weather forecasting and climate related studies. The strategy for implementing ECOCLIMAP-II is to depart from prevalent land cover products such as CLC2000 (Corine Land Cover) and GLC2000 (Global Land Cover) by splitting existing classes into new classes that possess a better regional character by virtue of the climatic environment (latitude, proximity to the sea, topography). The leaf area index (LAI) from MODIS and NDVI from SPOT/Vegetation yield the two proxy variables that were considered here in order to perform a multi-year trimmed analysis between 1999 and 2005 using the K-means method. Further, meteorological applications require each land cover type to appear as a partition of fractions of 4 main surface types or tiles (nature, water bodies, sea, urban areas) and, inside the nature tile, fractions of 12 Plant Functional Types (PFTs) representing generic vegetation types – principally broadleaf forest, needleleaf forest, C3 and C4 crops, grassland and bare land – as incorporated by the SVAT model ISBA developed at Météo France. This landscape division also forms the cornerstone of a validation exercise. The new ECOCLIMAP-II can be verified with auxiliary land cover products at very fine and coarse resolutions by means of versatile land occupation nomenclatures.

1 Introduction

Land cover regulates the surface energy budget and hydrological cycle, which are essential inputs for climate and weather prediction models. It is strictly defined as the observed physical layer that covers the surface of the Earth, including natural and planted vegetation, and man-made constructions. Actually, land cover is one of the most crucial properties of the Earth system for many areas of benefit to society (GEOSS, 2005). Information on land cover is essential for the protection of environment quality and biotic diversity worldwide (Sutherland et al., 2009). It is also of primary importance for sustainable management of natural resources (EEA, 2005) and human needs (Vitousek et al., 1997). In climate modelling, originally 1-degree global land cover databases were derived that combined pre-existing land cover maps and other atlases (Matthews, 1983; Olson et al., 1983; Wilson and Henderson-Sellers, 1985). Clearly, the coarse resolution of the grid mesh of a climate model led to mixing of vegetation species while focusing on broad scale natural ecosystems. This meant that climate modellers needed to reclassify pre-existing information in order to accurately model the land surface processes on the basis of a mosaic of individual ecosystems that were homogeneous from the functional point of view. The conversion of land covers – of ecosystems – into a suitable number of Plant Functional Types (PFTs) is a matter of great concern as PFTs allow vegetation models to capture most variations of defined plant traits that seem to be better represented by state variables than by fixed parameter values (Gitay and Noble, 1997; Kattge et al., 2011).

During recent decades, the advent of satellite observations has fostered the development of land cover products compatible with landscape units. In this respect, vegetation indices that combine spectral measurements in the visible and near infrared spectral wavebands have been widely used to discriminate vegetation species. The Normalized Difference Vegetation Index (NDVI), defined as the difference between the near infrared andred reflectance divided by the sum of the two, is undoubtedly the most widely used of the many indices available as it responds clearly to change in

GMDD
5, 3573–3620, 2012

the amount of green biomass (Tucker, 1979; Hill and Donald, 2003), chlorophyll content (Dawson et al., 2003), fire (Telesca and Lasaponara, 2006) and climate variability (Gong and Shi, 2003). A pioneering global study using satellite-based information has identified vegetation species with the objective of creating a coherent worldwide 8-km land cover map (Defries et al., 1995). Since the beginning of the 2000s, with the advent of a new generation of on-board sensors having increased radiometric and geometric resolutions, numerous land cover maps have been developed in the framework of national and international initiatives. At the scale of the European Union member states, there is the CORINE mapping initiative, which has produced two classifications at 100-m resolution: one for the year 2000, based on single-satellite images (Landsat 7 ETM), and the other one for the year 2006, based on images from two satellites (SPOT-4 or IRS LISS).

The other most popular land cover maps use the International Geosphere-Biosphere Programme Data and Information System (IGBP DISCover) (Loveland et al., 2000), University of Maryland (UMD) (Hansen et al., 2000), Moderate Resolution Imaging Spectroradiometer (MODIS) (Friedl et al., 2002), ECOCLIMAP-I database (Masson et al., 2003), Global Land Cover (GLC2000) (Bartholomé and Belward, 2005) and GlobCover (Bicheron et al., 2006), which were produced using data from NOAA/AVHRR, MODIS, SPOT/Vegetation and Envisat/MERIS at a spatial resolution of few hundred metres to 1 km. In addition to land cover classifications, the ECOCLIMAP product provides sets of surface parameters that are primarily useful in meteorology: notably surface albedo and leaf area index (LAI). However, stratification of land surface in Europe is permanently under investigation as knowledge of the geographic extent and dynamics of land cover is still incomplete, and broad levels of disagreement exist among current land cover maps due to the high level of landscape fragmentation at mid-latitudes (Herold et al., 2008; Fritz and See, 2008).

The aim of the present study is to update ECOCLIMAP-I at the European continental scale. This database was specifically designed to answer the needs of the meteorological community in investigating natural and managed ecosystems in connection with

GMDD
5, 3573–3620, 2012
ECOCLIMAP-II/Europe
S. Faroux et al.

weather forecasting and climate change modelling (Masson et al., 2003). The rationale for building the new classification map (ECOCLIMAP-II/Europe hereafter) is to better discriminate the land cover classes over Europe than is done by the existing continental maps such as ECOCLIMAP-I, GLC2000, and MODIS products. The latter products refer to single annual cycles of satellite data and may capture some undesirable anomalies that could be avoided with a multi-annual time series analysis of consistent remote sensing observations, as has been clearly demonstrated over Africa (Kaptué et al., 2011a, b). Hence, ECOCLIMAP-II/Europe will take advantage of the improvements provided by SPOT/Vegetation (acquired during the 7-yr period 1999–2005) in regard to radiometry, calibration monitoring, atmospheric correction, and normalization of surface directional effects compared to the NOAA/AVHRR datasets (acquired between April 1992 and March 1993) that were used to produce ECOCLIMAP-I.

In this paper we describe the methods and datasets used to produce ECOCLIMAP-II/Europe. Section 2 recalls the main characteristics of the ECOCLIMAP database and gives some technical information. In Sect. 3, we detail the satellite information used as input for the classification (SPOT/VGT NDVI) and the aggregation tool (MODIS LAI), the existing land cover products considered below, and also the elements of validation. The method of K-means employed for implementing ECOCLIMAP-II/Europe is thoroughly described in Sect. 4, along with the strategy for maintaining a minimum number of clusters and the technique of aggregation, i.e. limiting the division of land cover classes into a reduced number of PFT. The ECOCLIMAP-II/Europe classification with 273 land cover units is validated in Sect. 5 by comparison of landscape stratification at fine and coarse resolution scales, and with agricultural statistics over France. The last part, Sect. 6, summarizes the study and presents some perspectives for exploiting the results.

GMDD
5, 3573–3620, 2012

2 Principles of the ECOCLIMAP database

The objectives of the ECOCLIMAP classification are first to perform a stratification of the landscape into land cover units. ECOCLIMAP-I consisted of a global land cover map of 215 ecosystems at 1/120° resolution., with a data set of surface parameters associated with each land cover type, in tabular form: albedo, leaf area index (LAI), fraction of vegetation cover, fraction of photosynthetically active radiation (FPAR), roughness length, minimum stomatal resistance, and root zone. This set of surface parameters is highly suitable for initializing Soil-Vegetation-Atmosphere-Transfer (SVAT) models (Boone et al., 2009). As in other models investigating vegetation dynamics (Bonan et al., 2003; Rodell et al., 2004; Krinner et al., 2005), landscape scenes in ECOCLIMAP are organized in PFT patches. Sets of surface parameters are assigned mostly at the level of the PFTs representing generic land surface types, called tiles. The tile approach has been widely employed (Avissar and Pielke, 1989; Molod and Salmun, 2002). It consists of assigning surface parameters (albedo, LAI, and emissivity, for instance) to parts of the grid mesh (land, sea, inland water and built-up areas) within which these parameters vary as little as possible. The exercise in fact tries to describe each land cover as a combination of possible fractions of PFTs. Hence, the spatial distribution of the vegetation within a given cover is crucial as it ascertains the subsequent aggregation of the energy, water, and carbon fluxes, which are calculated separately for each tile. Average fluxes over the entire grid cell are returned to the atmospheric model and are used as the lower boundary condition. In the SVAT (Soil Vegetation Atmosphere Transfer) model ISBA (Interactions Surface Biosphere Atmosphere) used at Météo France (Noilhan and Mahfouf, 1996), the content of each ecosystem is formulated as a linear combination of 4 main surface types or tiles: *sea*, *inland water* bodies, *human built up* areas and *natural land* areas. The *natural land* tile is composed of 12 plant functional types (PFT): bare soil, bare rock, permanent snow and ice, deciduous broadleaf forest, evergreen broadleaf forest, needleleaf forest, C3 crops, C4 crops, irrigated crops, C3 herbaceous, C4 herbaceous, wetlands. C3 crops are winter crops (wheat, barley),

GMDD
5, 3573–3620, 2012
ECOCLIMAP-II/Europe
S. Faroux et al.

C4 crops are summer crops (maize, sorghum). Requirements coming from land surface modelling establish the set of parameters that are needed. For each land cover, ECOCLIMAP-I defined an annual profile of 10-day-averaged LAI values, a root depth, a total soil depth and a height for tree stands. Other surface parameters (notably fraction of vegetation, vegetation albedo, and roughness length) were only assigned per PFT, regardless of the cover. They were assigned values that were constant or calculated with formulas relying on LAI, soil depth or tree height. Worth noting here is that the 10-day periods of ECOCLIMAP-I were defined according to the legal calendar: from the 1st to the 10th; from the 11th to the 20th; and from the 21st to the end of each month.

Because of the need for the highest possible resolution over land, some initiatives for updating the ECOCLIMAP-I database have already been implemented. For instance, Han et al. (2005) updated the land cover over France. An example is the improvement of the description of biomes for south-western France, with which winter and summer crops could be separated, thereby leading to relevant detailed simulations of the atmospheric carbon dioxide in the CarboEurope Regional Experiment Strategy (CERES) (Sarrat et al., 2007). More recently, Kaptué et al. (2010) developed a new ecosystem classification within the ECOCLIMAP-II programme, with 37 distinct types over West Africa. This database was developed over the AMMA (African Monsoon Multidisciplinary Analysis) zone to provide upgraded information on the land surface properties of the West Africa region. In this study, GLC2000 classes were split using ECOCLIMAP-I classes. Then the MODIS LAI temporal profiles were used to group together the classes obtained.

3 Datasets

The update of ECOCLIMAP over Europe – defined here as the region between longitudes 11° W and 62° E and latitudes 25° N and 75° N (Fig. 1) – was performed in three steps: gathering input data sets, delineating the definition and characteristics of land cover types, and, finally, consolidating the database through validation exercises. The characteristics of information sources used either as input or for validation are indicated in Table 1.

3.1 Pre-existing land cover map products

The starting point was pre-existing land cover maps, which were divided up to form more regionalized ecosystems. The first one was the most successful Corine Land Cover map for the year 2000 (CLC2000), produced by the European Environment Agency and covering the 25 European Union member states (EEA, 2005b). This map is available for the EU at a spatial resolution of 100 m in the Lambert Azimuthal Equal Area (LAEA) projection. It mirrors land cover in Europe for the years 1999 to 2001. Based on photo-interpretation of high-resolution satellite imagery from SPOT (Satellite Pour l'Observation de la Terre) and Landsat and embedding other sources of data (aerial photographs, topographic and thematic maps), national land cover maps were produced by each EU member state. CLC2000 consists of 44 classes with a fine breakdown into categories obtained by merging the consistent national products into one dataset. CLC2000 can be deemed the reference map for the European part of the domain.

The second land cover map used in this study was the Global Land cover 2000 (GLC2000) database produced by the Joint Research Centre (JRC) (Bartholomé and Belward, 2005). It yields a global product derived from the analysis of 14 months (November 1999 to December 2000) of daily global data acquired by the SPOT/Vegetation (VGT) sensor at a spatial resolution of $1/112^{\circ}$ (1 km). The product was developed on the basis of regional classifications made with the aid of regional expertise. Using a bottom-up approach of 19 regional windows, the regional legends were based on the FAO (Food and Agricultural Organization) classification scheme (Di Gregorio and Jansen, 2000), which consists of 22 land cover classes for the whole world. Also, GLC2000 proposes a mosaic of five regional maps for Europe, including main land units with more detailed categories than the global one.

GMDD
5, 3573–3620, 2012

These two land cover maps were combined, so that GLC2000 filled areas not covered by CLC2000. In this respect, the legend was simplified into 14 categories of surface types: crop, needleleaf forest, broadleaf forest, herbaceous forest, rock, water body, bare soil, mixed forest, crop/natural vegetation mosaic, wetland, urban area, forest/other vegetation mosaic, irrigated crop, snow and ice. The resulting map, called C14, is shown in Fig. 1.

3.2 NDVI data from SPOT/Vegetation

The overall objective was to build up a consistent map product at the continental scale privileging satellite information. The orbital configuration combined with the viewing geometry of the Vegetation (VGT) sensor, which has been a passenger onboard SPOT-4 since 1998 and SPOT-5 since 2002, ensures daily global Earth coverage. Based on the Maximum Value Composite (MVC) (Holben, 1986), S10 10-day composite products of the Normalized Difference Vegetation Index (NDVI) are produced at $1/112^{\circ}$ spatial resolution in a Plate-Carrée projection (WGS84 ellipsoid) (Hagolle et al., 2004; Maisongrande et al., 2004). The compositing period was defined according to the legal calendar: from the 1st to the 10th; from the 11th to the 20th; and from the 21st until the last day of the month. The choice of the compositing period was a trade-off between the expected frequency of changes in vegetation and the minimum length of time necessary to produce cloud-free images. The period investigated spans seven years, from 1 January 1999 to 31 December 2005. This seven-year-long archive captures the mean annual vegetation cycle on a nearly climatic scale but can also be used to depict the inter-annual variability. S10 data composites also provide per-pixel cloud condition information allowing most cloud contamination in the NDVI signal to be removed. If less than 4 unsuitable NDVI values occur successively, a linear interpolation is applied to fill the gaps. Otherwise, the gaps are kept as missing data. To fill in the gaps caused by cloud contamination, a 4-degree polynomial function is used. This approach is similar to that used by Mayaux et al. (2004) and Kaptué et al. (2010) in previous studies.

GMDD
5, 3573–3620, 2012
ECOCLIMAP-II/Europe
S. Faroux et al.

3.3 LAI data from MODIS

The Leaf Area Index (LAI) is defined as the ratio of one-sided green foliage area per unit of horizontal ground area in broadleaf canopies, or the projected needle-leaf area per unit of ground area in conifer canopies, and is given in $m^2\,m^{-2}$ (Yang et al., 2006). In the original version, ECOCLIMAP-I, the seasonality of LAI was scaled on the annual dynamics of NDVI obtained from the Earth observing satellite system NOAA-AVHRR. The maximum value of LAI corresponded to the annual maximum green vegetation. In the new version, ECOCLIMAP-II, we consider collection 5 of MODIS LAI, the algorithm of which employs a look-up-table (LUT) approach using the MODIS 8-biome land cover classification with the radiative transfer approach of Myneni et al. (1999). MODIS LAI is available at a spatial resolution of $1/120^{\circ}$ in an Integerized Sinusoidal Grid (ISG) and at a temporal resolution of 8 days. It was re-projected on a Plate-Carrée grid to match VGT NDVI. MODIS LAI was also linearly interpolated for the sake of synchronicity with the ECOCLIMAP 10-day temporal resolution. The data were smoothed following the same procedure as described for the VGT NDVI. Unclassified and missing data, including urban areas, wetlands, snow, bare soil and water bodies were excluded from the procedure.

3.4 Climatic data sets

A climate database is used in order to avoid grouping classes pertaining to different climates. Two climate maps were actually used. The first, proposed by Koeppe and De Long (1958) is global with 16 climate classes. The second, produced by the FIRS project (EC, 1995), covers Europe and suggests 23 classes. It is leveraged by geo-factors such as climate, soil and topography. A combined map covering all areas of interest was built by assigning values of the FIRS classes to Koeppe and De Long's (1958) classes, which extended the former out of their area of definition (Fig. 2). This revealed conspicuous agreement between the two climate maps although slight

GMDD
5, 3573–3620, 2012
ECOCLIMAP-II/Europe
S. Faroux et al.

evidence of the limits between the two original maps was visible in southern and eastern parts of Europe.

3.5 Validation data

A multi-scale land cover verification was performed. In particular, multi-scale ancillary information was considered in order to consolidate the aggregation strategy, i.e. the fractional distribution of PFT within the land cover types. Statistics from the French Ministry of Agriculture and Forests can be found in the AGRESTE (http://agreste.agriculture.gouv.fr/) database. This database is composed of annual land use expressed in hectares for each French department, distinguishing different types of crops and non-agricultural areas. The available years are from 1999 to 2005, which are common with the available NDVI data sets. Another means of verification was the 1° global map of percentage of C4 vegetation produced within the framework of the International Satellite Land Surface Climatology Project (ISLSCP-II) Initiative II Data Collection (Still et al., 2003). We also considered a classification product prepared by CESBIO (Centre d'Etudes Spatiales de la BIOsphere) at 20-m resolution for an area of 3600 km^2 located near the city of Toulouse (France) where crop and forest types are notably encountered. The method was a supervised maximum likelihood combining multi-date (6 dates) and multi-spectral (21 bands) data from FORMOSAT and SPOT imagery. The resulting land cover map was updated each year from 2002 to 2005. Only the 4-yr average was considered for the purposes of this study, which concerned the percentage of FORMOSAT classes in ECOCLIMAP-II 1-km pixels. Note that all land cover maps were finally re-projected on a Plate Carré grid with the WGS84 geoid system.

4 Implementation of the ECOCLIMAP-II database

The procedure is summarized in Fig. 3. Before stratifying the domain of interest into land cover types, we used the C14 map (Fig. 1) to mask out pixels identified as water (either inland or oceanic) or urban. Then, the classification was performed according to 3 steps: (i) use of the K-means algorithm to disentangle all data sets, (ii) reduction of the number of clusters, and (iii) integration of the information from the existing land cover maps. After the classification process, the last step of our procedure was to define, for each cover, the percentage covered by the 4 main surface types (land, sea, inland water and towns) according to the "tile" approach, and to determine the 12 fractions of PFTs inside the nature tile, and the LAI profiles, the root and soil depths, and the heights of trees for each of the 12 functional types of natural land areas represented in the cover.

4.1 Disentangling datasets

Samples were formed using the K-means clustering algorithm of Hartigan and Wong (1979). This is an unsupervised learning algorithm that is suitable for clustering multidimensional data sets. The K-means method seeks to partition all points into k clusters such that, in a multivariate attribute space, the total sum of squares (or squared deviations) from a set of individual points – represented here by the pixels – is minimized with respect to an optimum number of cluster centroids. The algorithm can be parsed as follows: (i) a number, k, of points is randomly placed in the space represented by the objects to be clustered; (ii) each object of the group is assigned to the closest centroid; (iii) once all objects have been assigned, the positions of the k centroids are recalculated. The process is repeated until the position of the centroids is stable. The final step is then to minimize the metric of the objects with respect to the k clusters.

The K-means algorithm is sensitive to the initial configuration of cluster seeds and does not necessarily find the optimal configuration corresponding to the global objective function minimum (Kanungo et al., 2002). In practice, the K-means algorithm can be run several times to reduce this effect. We start with randomly chosen centres of clusters. Then, after a few iterations, rapid convergence is obtained for the centres of clusters based on the criterion of stability to decide that an optimal distribution of clusters has been reached. The K-means algorithm is well-known for its efficiency and robustness in handling the bulk of datasets. However, some constraints and limitations seem to exist for this approach. Firstly, the number of clusters ascertains the precision for obtaining a sharp description of the whole data set. Secondly, Jain et al. (1999) calculated the sensitivity of the algorithm to the initial positions of cluster centres because, in some situations, important populations may be misrepresented. Thirdly, the use of the Euclidian distance as a unique criterion may be too restrictive to describe the dynamics of NDVI time profiles. Nevertheless, in our approach, we did try to circumvent such difficulties by using first a large number of clusters, and then refining the selection using other criteria such as optimizing the combination between the correlation and the distance. Therefore, due to the relatively large number of clusters, no specific patterns were buried in clusters that were too big.

4.2 Reducing the number of clusters

In order to reduce the number of classes to the target number (set between 200 and 300), several criteria were tested on the centres of the clusters obtained. Finally a resemblance criterion (referred as RC in the rest of the paper) was selected:

$$RC = \frac{d}{r^2} \tag{1}$$

where d and r refer to the Euclidean distance and the Pearson's correlation, respectively, between the mean 10-day NDVI profiles. This criterion is a trade-off that serves to account for both the dynamics and intensity of the NDVI signal. The use of a squared correlation gave stronger weight to the correlation at this stage of the process. Families of clusters were formed by grouping NDVI mean profiles of clusters if the RC criterion

GMDD
5, 3573–3620, 2012

computed over their cluster centres relative to all the other clusters in the family was below a fixed threshold. Several tests were performed on the threshold value. This provided a supervised step and the new classes were thoroughly examined at each step, in particular their spatial organization, NDVI mean profiles and the total number of clusters. Visual inspection was ultimately the main criterion at this stage of the classification. The value of the threshold was modulated according to the size of the clusters, a higher threshold being assigned to smaller clusters in order not to multiply poorly representative covers. Also, the maximum value of NDVI was examined: the threshold was increased if the NDVI value was considered doubtful in order not to segregate clusters having a low vegetation rate.

The above operations revealed that clusters built with only the NDVI information were geographically consistent. Inspection of the distribution of clusters within the study domain generally revealed bundles of pixels for which a justification could be found by looking either at the orography, or at a coarse climatic zoning (latitude, proximity of sea and even of country boundaries for arable land). Such outcomes validated the choice of the NDVI as the main classifier. Even if NDVI did not succeed in describing all surface types characteristics, it was at least able to capture most of the variability of the land cover at the continental scale, thereby resulting in 270 final clusters for the classification product based on NDVI alone.

4.3 Integration of information derived from existing land cover maps

At the third step, to strengthen the coherency between the 270 NDVI-based classification map (described in Sect. 4.2) and the C14 map (described in Sect. 3 and complemented by CLC2000 and GLC2000 when necessary), for each NDVI-based cluster, the mean NDVI time profile per main land cover type was calculated. Pixels that did not belong to the two dominant land cover types were moved to another cluster where their type would be more appropriate. This option was activated based on RC again, with an adapted threshold, but only if the Pearson's correlation was above 0.9. This operation was supervised and minimized the discrepancy between the NDVI-based clusters and

GMDD
5, 3573–3620, 2012
ECOCLIMAP-II/Europe
S. Faroux et al.

the land cover maps. It is worth noting here that this criterion would disregard the geographical distance. Actually, it was not even necessary to introduce it as geographical coherency of the grouped pixels was obtained by construction.

Climate maps were finally used to avoid the segregation of pixels belonging to different climate units but this actually concerned a very few situations. For instance, three classes were split into two to better fit the land cover maps, which means that there were three more classes after this step. A dedicated treatment was nonetheless provided for urban covers. Suburban areas based on both GLC2000 and CLC2000 information were classified using the above method based solely on NDVI, while the other urban covers were directly inherited from the land cover maps. The new, final ECOCLIMAP-II classification includes 273 covers.

4.4 Defining the surface parameters

A correspondence was established between the 273 covers, the 4 surface types and the 12 functional types of natural land areas. This step only focused on suburban areas and natural land areas, while no further parameters were needed for the other surfaces (other urban classes, inland water bodies and sea). For each cover, a thorough interpretation was made of the CLC2000/GLC2000 classes appearing in a given cover. Using the several classes for a given cover was found to be of great benefit in removing any ambiguity in the classification of the classes in terms of the functional types. The indications of density in the classes were used to fix the percentages of functional and surface types.

At the beginning of this step, only LAI profiles (coming from MODIS satellite data) and functional type fractions (see their assessment in Sect. 3.2.1) of the covers were known. An iterative technique based only on these two sources of information was implemented to determine the LAI profiles of functional types inside classes. For each non-zero fraction of a functional type within a cover, we determined the LAI profile. Then, a disentanglement process was implemented to obtain the LAI profiles for functional types from the LAI profiles of the covers. Figure 4 describes this process in a diagram.

First, an approximation identified the LAI profile of the main functional type of a cover with the mean LAI profile of this cover. Then, the principle of disentanglement was to search other covers in the vicinity, in which each minority functional type of the initial cover yielded the major functional type. The LAI profiles of the minority functional types of the first cover were then associated with the LAI of the selected covers. The selective criteria for the procedure were, in order of decreasing importance: preponderance of the type in the cover, geographic proximity (with reference to the climate map), maximum correlation between the two associated classes with respect to the NDVI temporal profile. At this stage of the process, the LAI temporal profile of secondary PFTs was known for a given cover. Then, the LAI temporal profile of the major PFT was re-calculated by subtracting the LAI temporal profiles of secondary PFTs, weighted according to their fraction, from the initial LAI. In a very small number of cases where negative LAI values were identified, we selected other covers of reference for the minor PFTs. In this case, a new iteration was performed using the LAI from the first step as a guess. This algorithm showed rapid convergence after the second step was reached.

The determination of root depth, total soil depth and tree heights for the functional types was inherited from ECOCLIMAP-I as it was judged that no reliable additional source was available to improve the values of these parameters in the framework of ECOCLIMAP-II development.

5 Results of validation

5.1 Analysis of NDVI per land cover class

The method for building the ECOCLIMAP-II land cover map is essentially based on a trimmed analysis of the spatial distribution of the NDVI time profiles over the domain of interest. Actually, this highlights how suitably such variations of NDVI seasonal patterns

GMDD
5, 3573–3620, 2012

are represented, as can be confirmed by a dependable comparison, with correlations up to 0.9, between SPOT/VGT NDVI and MODIS LAI calculated for each cover (except for urban areas, wetlands, snow, bare soil and water bodies) (see Table 2). The following subsections are devoted to a description of land cover characteristics in terms of intensity and evolution with respect to their geographic location. This emphasis on NDVI patterns for the extended European domain will complete the information provided by GLC2000 and CLC2000 to characterize the new land covers of ECOCLIMAP-II. Firstly, major cover types are reviewed along with illustrations of their NDVI time profiles (Fig. 5).

5.1.1 Forests

The presence of forests evolves from north-east to south-west. In northern Russia, NDVI values reach their peak during summertime and fall to spurious values in wintertime due to snow contamination for most scenarios. Hence, the narrow seasonal peak mirrors the short warm season and activity (Fig. 5a). Approaching central Europe, the NDVI annual cycles of forests take on square shape, i.e. high NDVI values last over a longer time, with a reduced amplitude, signifying less variability in climatic conditions, although with various degrees of severity (Fig. 5b and c). Near the Mediterranean Sea, the annual amplitude of NDVI time series decreases further, and observed profiles sometimes even become flat (Fig. 5d). Clearly, permanently cool temperatures coupled with an increasing number of sunny days and proximity to the sea seem to support quiescent periods of thriving vegetation.

5.1.2 Herbaceous plants and shrubs

Over northern, central and western Europe, NDVI time profiles for herbaceous plants and shrubs resemble those of forest with a strong annual amplitude and sharp peak in the north and east (Fig. 5e). Moving towards the south-west, the annual NDVI variations become broader and more square-shaped again. In particular Atlantic moors can be distinguished by a regular, smooth, rounded NDVI time profile (Fig. 5f). Other noteworthy profiles are those of grassland located in the Massif Central (France) (Fig. 5g) and the mosaic of grassland and crops in the Vendée region (France) (Fig. 5h). Their inter-annual variability and annual cycles have no equivalent within the study area. Around the Mediterranean basin, about 5 types of herbaceous plants and shrubs can be catalogued:

- The first has a "triangular" profile, with a first peak triggered in spring and a second, weaker one in autumn. Classes inside this type are notably distinguished by the position of the time-shifted moderate second peak relative to the summer peak. This kind of herbaceous vegetation and shrubs can be found in central Asia and Turkey (Fig. 5i and j).

- A second noticeable type is characterized by a single NDVI peak starting during springtime. It is rather similar to the previous type except that the second peak is flattened. This type is located exclusively in north African and north Arabian regions with smooth NDVI variations (Fig. 5k).

- In some cases, the peak of the second type is shifted towards the wintertime, which yields a third type that is also present in North Africa and northern Arabia (Fig. 5l).

- A fourth type is formed by a secondary peak occurring in wintertime, which, in contrast to the first type, is associated with a triangular profile. This type principally occurs in North Africa, Spain and Portugal (Fig. 5m and n).

- A final singular type also has a square-shaped NDVI present in wintertime. A somewhat similar NDVI pattern is noticeable in spring and summer for forested areas of west-central Europe and is also significant in areas surrounding the Mediterranean Sea (Fig. 5o).

GMDD
5, 3573–3620, 2012
ECOCLIMAP-II/Europe
S. Faroux et al.

So far, these five types show a rather sparse distribution around the Mediterranean basin, in poorly delimited locations, rather than forming a complete set of Mediterranean ecosystems. It is likely that different botanical properties in connection with different environments (climatic and/or geological) could explain such apparent differences.

5.1.3 Crops

Crop areas are well managed and thus generally well delimited in space, having characteristics that form an integral part of a climatic region. Nonetheless, an exception has to be made for the Mediterranean basin, where the organization of covers is complex with a high level of mixed plots. For this region, NDVI profiles for crops resemble those of herbaceous plants and shrubs. The fifth type of profile corresponds to Spanish Estremadura agro-forestry areas (according to CLC2000). Crops and herbaceous plants are probably quite mixed there.

In the eastern part of Europe (Russia, Kazakhstan), NDVI for crops show stretched profiles, also triangular-shaped with a peak in the middle of summer (Fig. 5p). The most smooth, rounded NDVI profiles are found in southern Europe, notably along the Po plain (Italy) and in southern France. A tiny difference between winter and summer crops is noticeable within this type as the peaks are about 3 months apart (Fig. 5q and r).

In western Europe, for example in the Paris Basin, a typical NDVI profile consists of a first high peak during spring followed by a secondary, very small peak in early winter. This is characteristic of winter crops being sowed in early autumn immediately after the harvest and showing some growth before the dormant period that precedes further growth in the following spring (Fig. 5s).

Some regions are very well delimited and are marked by compact areas of crops: Bulgaria (Fig. 5t), Hungary (Fig. 5u), Turkey at the Bosphorus, Poland, Germany, south-west England, French Brittany, French Vendée, the Po plain, Spanish Castile. Nile delta crops have specific NDVI profiles with 2 peaks of equal amplitude (Fig. 5v). These

GMDD
5, 3573–3620, 2012

profiles are structured in very small clusters and extend, although very locally, from Turkey to Syria. In contrast, several classes of crops are quite scattered and their NDVI profiles resemble those of forests and herbaceous plants. In such situations, it is believed that crops grow alongside other types of vegetation.

5.1.4 Areas of sparse vegetation and bare land

NDVI profiles for sparse vegetation look like those of herbaceous plants and shrubs but have a lower intensity. This points out the effect of density of vegetation on the intensity of the NDVI signal. Bare land areas show NDVI profiles that can be directly related to the height of the vegetation.

5.1.5 Miscellaneous

A clear difference is noticeable in soil occupancy between the land surface surrounding the Mediterranean basin and the rest of Europe. Generally speaking, classes that are located outside the Mediterranean region are geographically well-outlined and rather compact. For these classes, the changes in NDVI time profiles can be ranked according to latitude but also depend on the presence of a sea or ocean nearby. In this case, land cover types can be referred to as pure. On the other hand, ecosystems bordering the Mediterranean region are made up of mosaics of vegetation kinds spread over broad geographic areas and are often referred to as mixed land cover types. In this respect, and unlike the rest of the domain, regions bordering the Mediterranean Sea do not permit a straightforward analysis of the spatial distribution of the land covers because of possible overlapping between vegetation units. In this case, the partitioning of land cover types into PFTs (or patches) is rather challenging.

Figure 6 proposes a simplified visualization of the map obtained. The 273 classes kept for modelling are grouped by proximity of content into 103 named classes.

GMDD
5, 3573–3620, 2012

Title Page
Abstract | Introduction
Conclusions | References
Tables | Figures
◂◂ | ▸▸
◂ | ▸
Back | Close
Full Screen / Esc
Printer-friendly Version
Interactive Discussion

5.2 Elements of validation for ECOCLIMAP-II

In this section, the goal is to carry out a validation exercise concerning the fractions of PFTs to be assessed for all pixels of a given land cover of the ECOCLIMAP-II database. For this, we use independent sources of information to build maps of new spatial resolution, which then serve as references.

5.2.1 Comparison with AGRESTE for France

Statistics from the French Ministry of Agriculture and Forests are brought together in the AGRESTE database (http://agreste.agriculture.gouv.fr/). AGRESTE is used here to compute the fractions of PFTs in ECOCLIMAP-II land covers for each of the 95 administrative divisions (departments) of metropolitan France. The fractions of PFTs are first weighted by the representative fractions of the covers in each department and then summed. These fractions are also reduced to 6 common distinctive PFTs: forests, grassland, C3 crops, C4 crops, permanent crops, plus all the other PFTs grouped into one.

Further, the information contained in AGRESTE hectares and ECOCLIMAP-II kilometre pixels is converted into percentages of land use in the administrative divisions, and compared at department level for these 6 types. The findings of the comparison are displayed in Fig. 7. An error estimate is given per department, and consists of the root mean square error (rmse) between the representative fractions of the 6 PFTs for ECOCLIMAP-II and AGRESTE:

$$\mathrm{Err(dept)} = \left(\sum_{\mathrm{type1}}^{\mathrm{type6}} (\mathrm{Frac}_{\mathrm{ecov2}} - \mathrm{Frac}_{\mathrm{agreste}})^2 \right)^{\frac{1}{2}} \quad (2)$$

It can be seen that this error falls below 15 % for the great majority of departments, which validates the approach. It should be stressed that it represents a cumulated error for all 6 PFTs and that, for a single PFT, it would come down to only 2.5 % on

GMDD
5, 3573–3620, 2012
ECOCLIMAP-II/Europe
S. Faroux et al.

average. Nevertheless, it can be observed that a few departments, well spread over France (Loire-Atlantique, Cantal, Hérault and Ile-de-France) have errors higher than the average estimates, for instance up to 35 % for Ile-de-France. But this department is strongly urbanized, which brings out a somewhat thorny problem of representation.

5.2.2 Comparison with ISLSCP2 C4 map

The International Satellite Land Surface Climatology Project, Initiative II (ISLSCP2) notably addresses land-atmosphere interactions focusing on land cover, hydrometeorology, radiation, and soils. In this respect, ISLSCP2 proposes a stratification of the landscape at 1° resolution. The purpose here is to compare the C4 fraction from ECOCLIMAP-II (crops + herbaceous) with ISLSCP2 data at European scale. The reprojection of ECOCLIMAP-II C4 at the 1° resolution of ISLSCP2 is achieved by simply performing a linear aggregation of the 120 × 120 ECOCLIMAP pixels. Figure 8 shows the fractions of C4 vegetation for ISLSCP2 and for ECOCLIMAP-II. It can be seen that higher values are generally obtained for C4 fractions with ECOCLIMAP-II, except for the Paris Basin, in Romania and in part of Ukraine (as low as −10 % in places). Larger differences are particularly noteworthy in northern Italy (Po plain, +30 % to +50 %), in Hungary (+20/25 %), south-west France (around +10 %), northern Egypt (Nile delta, around +25 %), around the Aegean Sea and the Black Sea, and south of the Caspian Sea (+5 % to +10 %). A low C4 fraction (2–3 %) is also prominent with ECOCLIMAP-II almost everywhere, with the exception of southern deserts and northern Russia.

The Food and Agricultural Organization (FAO) (http://www.fao.org/es/ess/top/country.html) holds freely available yearly yield statistics per country for several categories of agricultural products. Over our domain of study, the first producers of maize in 2005 were France, Italy, Romania, Hungary, Ukraine and Egypt. This is in full agreement with the consistent results between the ECOCLIMAP-II and ISLSCP2 maps. For instance, higher fractions of C4 for ECOCLIMAP-II are observed in Italy, Hungary and Egypt. The interpretation of the discrepancies between the two maps can be oriented in two directions. First, owing to a better spatial resolution, the ECOCLIMAP-II products seem superior for the inventory of C4 crops. Moreover, the landscape stratification performed in the framework of ECOCLIMAP-II suggests a land cover classification based purely on a multi-temporal analysis of the surrogate variable NDVI. Therefore, caution is advisable as to the homogeneity of the land covers concerning C4 fraction representation because a low fraction of C4 NDVI would be drowned in the pixel integration of the NDVI signal. Unfortunately, the attempt to adjust C4 fractions in France with the aid of AGRESTE information could not be duplicated elsewhere. Clearly, the extrapolation of these C4 fractions outside France through the land covers must contain a part of uncertainty.

Generally, the spatial attributions of C4 crops in ECOCLIMAP-II are in agreement with ISLSCP2 and FAO statistics (FAO, 2008) whereas the method for implementing ECOCLIMAP-II leads to some imprecision on the depiction of C4 fractions. Certainly, an equivalent of AGRESTE at European scale would bring new insights but information extrapolated from AGRESTE already helps in the setting up of a continental-scale map of C4 crops with a level of reliability that allows it to be fully exploited further.

5.2.3 Comparison with FORMOSAT-2 products

FORMOSAT-2 is an NSPO (Taiwan National Space Program Office) Earth imaging satellite with the objective of collecting high-resolution panchromatic (2 m) and multi-spectral (8 m) imagery for a wide variety of applications, such as land use, agriculture and forestry. FORMOSAT-2 is able to revisit the same point on the globe every day in the same viewing conditions. FORMOSAT observations of a very small area (60 × 60 km south-west of Toulouse, France) at very high resolution (about 2 m) have been acquired to establish a classification product at a resolution of 20 m with 21 land covers. The study area is composed as follows: 23 % of wheat crop, 21 % of grassland, 17 % of sunflower, 10 % fallow, 9 % of man-made material, 6 % of deciduous forest and 6 % of maize, the rest being a mosaic of different landscape units (sorghum, soybean, barley, rapeseed, conifers, river). The strategy for verifying ECOCLIMAP-II is different here than previously. For each land cover of ECOCLIMAP-II present in the area of interest,

GMDD
5, 3573–3620, 2012
ECOCLIMAP-II/Europe
S. Faroux et al.

the sub-grid information from FORMOSAT is aggregated by first estimating the percentage of the fractions of PFT inside each land cover. As a result, given the resolution of ECOCLIMAP-II, the FORMOSAT classification is given the same nomenclature as ECOCLIMAP-II for the sake of a fair comparison. Figure 9 shows the percentage of presence for the 8 most representative PFTs (broadleaf forest, needleleaf forest, C3 crops, C4 crops, grassland, urban areas, bare/waste land and water bodies) as a function of the 12 most representative ECOCLIMAP-II and FORMOSAT land covers. Note that FORMOSAT classes were grouped together in order to match the ECOCLIMAP-II nomenclature as far as possible. In particular, meaningful agreement was obtained between the two 1-km classifications for broadleaf trees, C3 crops and urban areas. For grassland, the distribution was somewhat more erratic but the main lines were still present. However, the C4 fractions were underestimated in ECOCLIMAP-II compared to FORMOSAT. This is believed to be due to the fragmented nature of the landscape with C4 crops being present on small plots of land. For such cases, it is clear that having a moderate, 1-km resolution, ECOCLIMAP-II will always fail to capture the true presence of C4 crops. The coniferous curves also diverge, but conifers are only present at the edges of the FORMOSAT domain, which means that the comparison is less relevant in this case than for broadleaf pixels for instance. Water body curves do not concur as partly expected because water represents a fraction of less than <1 %. Also, the overestimation noticed for bare land in the case of ECOCLIMAP-II seems to be related to the fact that, at 1-km scale, land covers are diluted into larger bands of bare land areas.

Considering the comparisons with finer (FORMOSAT) and coarser (ISLSCP2) land cover classifications, along with the satisfactory level of consistency with both of them, ECOCLIMAP-II can, at this stage, be deemed successful in properly aggregating small scale information and harmonizing broad scale information. This is a necessary condition for the efficient use of ECOCLIMAP-II in seamless climate models.

GMDD
5, 3573–3620, 2012
ECOCLIMAP-II/Europe
S. Faroux et al.

5.3 Comparison with ECOCLIMAP-I map

With the elaboration of the new ECOCLIMAP-II database, a key objective is to update the former, rather obsolete ECOCLIMAP product. In this regard, it is necessary to quantify the consequences of the modifications and input data sets used by making a comparison between the LAI values of the two versions and also with respect to the fractions of PFTs. This verification exercise is performed at 1-km resolution on a pixel-by-pixel basis independently of land covers. The outcomes are composed of raster maps of the differences, which provide a critical tool for the application upgrade and the maintenance strategy of ECOCLIMAP versions.

5.3.1 Comparison of LAI temporal profiles (Fig. 10)

The evolution of LAI with time is analysed by grid point, which already embraces the landscape aggregation into PFTs. The following three quantities are considered for further analysis as explained below:

- The correlation between LAI temporal profiles of the two ECOCLIMAP versions, in order to assess the potential changes in the dynamics of LAI. This correlation is given only when the two ECOCLIMAP PFTs other than just bare land, rock and snow are present (Fig. 10a).
- The relative difference, diffrel, in maximum and minimum values of LAI (Fig. 10b and c), defined as:

$$\text{diffrel} = \frac{V_{\text{ec2}} - V_{\text{ec1}}}{V_{\text{ec2}} + V_{\text{ec1}}} \times 200 \tag{3}$$

where V is the maximum or the minimum value of LAI on a given grid point. These two quantities give information on the detection of changes in the LAI.

Figure 10 displays maps for these three quantities, i.e. the correlation and the two cases of maximum and minimum of LAI applied to the formula in Eq. (3).

GMDD
5, 3573–3620, 2012

The LAI correlations (Fig. 10a) are dependable (values >90 %) over the whole north-eastern part of the domain. The correlation values decrease towards the Mediterranean sea, even falling below 50 % along the Mediterranean coastline. Negative correlation occurs over semi-deserts but this is not drastic because LAI remains low in these regions. The representation of LAI in the Mediterranean region is expected to be improved with ECOCLIMAP-II because the discrimination between land covers relies on a method that offers refinement and focuses, above all, on homogeneity for a given land cover with respect to NDVI and LAI.

The value of LAI_{max} (Fig. 10b) in ECOCLIMAP-II is higher for the main forests and herbaceous areas, and lower for main crops and semi-desert regions. It has already been mentioned that MODIS LAI maximum values, on which the ECOCLIMAP-II LAI is built, are higher for forested areas than, for instance, the LAI values from CYCLOPES, which made use of SPOT/Vegetation observations (Baret et al., 2007). Otherwise, and perhaps incidentally, maximum heights were found to be generally equivalent between MODIS and CYCLOPES LAI values (Weiss et al., 2007), and the reason why MODIS was selected was the availability of longer time series of data sets. The results of this comparison in terms of correlation are summarized in Table 2. Median correlation is always higher than 0.9, showing the good correlation between the three datasets.

5.3.2 Comparison of the fractions of PFTs (Fig. 11)

This comparison is complementary to the LAI temporal profiles because the quality of the update of ECOCLIMAP should also be judged through the new distribution of PFTs within the land covers. Incidentally, note that LAI has repercussions on some other parameters like the root zone, soil depth and aerodynamic roughness (for low vegetation) but this is not important enough to deserve a dedicated analysis. The parameter investigated in this section is the simple difference $F_{ec2} - F_{ec1}$, where F is the fraction (expressed in %) of the PFT considered in a given grid point. Figure 11 shows these differences on a map for the 12 PFTs.

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

GMDD

5, 3573–3620, 2012

ECOCLIMAP-II/Europe

S. Faroux et al.

Bare soil, bare rock, snow (Fig. 11 a, b and c)

The percentage of bare land (Fig. 11a) increased almost everywhere and particularly on the marked topography near the Mediterranean Sea (close to +50 %) and in pre-desert zones north of the Caspian sea. Different elements of an explanation can be found, like an increased number of burn scars for the Mediterranean region but also a better pixel resolution in the case of ECOCLIMAP-II. However, the greater area of bare soil in place of rock zones, notably at high altitudes in Norway and in the Alps, seems to be an artefact, rather than being inherent in the differences in method and data quality between the two versions of ECOCLIMAP. The fraction of bare land is lower along the coasts of the Mediterranean (−10 %), the explanation for which may be found in management policy. In the southern deserts of the domain, some bare soil has become rock and conversely, which is thought to be due to the method. Concerning snow targets (Fig. 11 c), some snow plots in ECOCLIMAP-I are replaced by land cover with 10 % of snow and 85 % of bare land in ECOCLIMAP-II.

These changes are noticeable because of the will to include a large range of nuances between the pure land cover types in ECOCLIMAP-II, in order to access a continuity between land covers that did not exist in ECOCLIMAP-I. On the other hand, the choice to focus on the NDVI homogeneity rather than on the pureness of the land covers could be accompanied by some imprecision. The mean distribution of PFTs inside certain land covers may not be totally exact, even if this point has been verified, e.g. by checking that CLC2000 and GLC2000 classes are well-blended in ECOCLIMAP-II covers where they appear.

Deciduous broadleaf trees, needleleaf trees, evergreen broadleaf trees (Fig. 11d, e and f)

Broadleaf trees (Fig. 11d) are now more present in central Russia (+40 %, +70 %) while they have tended to disappear from the Mediterranean region, especially near the coastline (−25 %). It is worth emphasizing that needleleaf trees (Fig. 11b) are notably

GMDD
5, 3573–3620, 2012

less represented in northern and western European mountains. They have been replaced by broadleaf trees in northern Europe and grassland in western Europe. The few evergreen trees initially found in Mesopotamia in ECOCLIMAP-I have vanished in ECOCLIMAP-II.

For the forest scenario, as for bare land areas, an effort was made to adequately reproduce the complexity of mosaics inside the land covers, hence discrepancies are naturally observed. Here again, this is at the expense of a clear discrimination of dominant land cover types.

C3 crops, C4 crops, irrigated crops (Fig. 11g, h and i)

In ECOCLIMAP-II, there are obviously less irrigated crops (Fig. 11i) in the north of the Mediterranean region, except on the Turkish west coast and in the south west of France. These differences are related to the mixing of vegetation types in ECOCLIMAP-II. Nile valley crops now appear irrigated in ECOCLIMAP-II. The fractions of C3 crops are often lower in ECOCLIMAP-II because there are more mosaics of grassland with crops in the new classification. The category of C4 crops (Fig. 11h) is evenly more present in the Po plain, French Alsace and Hungary, thanks to the dating of maximum of NDVI profiles for the corresponding land covers. The fractions are slightly lower (−10 %) for the majority of the domain.

Temperate grassland, tropical grassland, wetlands (Fig. 11j, k, and l)

The area of temperate grasslands (Fig. 11j) has shrunk in Russia, north-western Europe and in arid south-eastern areas (−20 % to −50 %). They are more present in a great part of central and western Europe, especially continentally (around +30 %), which counterbalances the tree loss as seen previously. Tropical grasslands (Fig. 11k) present in the Maghreb area and also in Mesopotamia in the case of ECOCLIMAP-II are no longer conserved in ECOCLIMAP-II. It is probable that the way fractions of vegetation types are decided tends to disregard this distinction, bearing in mind that

GMDD
5, 3573–3620, 2012
ECOCLIMAP-II/Europe
S. Faroux et al.

Europe is not really an appropriate place to observe tropical grasslands. The distribution of wetlands (Fig. 11l) is completely modified in ECOCLIMAP-II. What seems to have happened here is that wetlands in ECOCLIMAP-II often show a high NDVI signal now, leading trees to be added in these areas. Conversely, the merging of CLC2000 and GLC2000 classes into ECOCLIMAP-II results in new areas of wetlands for some land covers.

6 Summary and conclusion

The new classification ECOCLIMAP-II/Europe with 273 distinct ecosystems has been exposed for Europe, together with a verification exercise. This upgraded information about the physiography is intended to improve the representation of the continental surface in the SURFEX model (Masson et al., 2012) and also to foster advanced investigations related to the carbon and water cycles. The spatial distribution and association of the land surface properties as previously defined within ECOCLIMAP-I has been revisited using enhanced, consistent long-term series of moderate resolution maps of NDVI and LAI originating from the new generation of onboard remote sensing instruments. The natural evolution of the landscape in connection with hazards (floods, fires) and human impacts (high concentration of habitat and population) mean that the European domain needs to be regularly redrawn. The popular method of K-means has once again proved its ability to help respect the main features of the landscapes assembled into a rather limited number of clusters (classes). Interestingly, the conversion of these clusters into PFTs, as is required for many applications and also for validation purposes, was possible without being detrimental to the fine quality of information. The LAI and NDVI tools of discrimination have proved that information initially compiled from land cover maps could be still more trustworthy. Incidentally, it is worth underlining the commendable coherence between the SPOT/VGT NDVI product used for the classification and MODIS LAI used for aggregation, without which the study would not have

GMDD
5, 3573–3620, 2012
ECOCLIMAP-II/Europe
S. Faroux et al.

been possible. Combining two sources of information has no doubt strengthened the reliability of the ECOCLIMAP-II classification product as it is less sensor dependent.

It emerges that, at moderate resolution, typically of one kilometre, the majority of pixels are mixed at the scale of measurement, particularly near the Mediterranean. Since the genuine asset of the ECOCLIMAP-II database is to include coherent sets of biophysical variables primarily used in meteorology, the consistency of variables like LAI and fraction of vegetation was verified after broad-scale aggregation. The uncertainty given on ECOCLIMAP-II surface parameters according to their intra-class and inter-annual variability favours their use in data assimilation systems of carbon and water budget models. The ECOCLIMAP-II product is freely available (http://www.cnrm.meteo.fr/surfex/spip.php?article19).

Acknowledgements. The authors are indebted to Centre d'Etudes Spatiales de la Biosphere (CESBIO, Toulouse) for providing access to the high-resolution land cover map product generated with FORMOSAT observations.

The publication of this article is financed by CNRS-INSU.

References

Avissar, R. and Pielke, R. A.: A parameterization of heterogeneous land surfaces for atmospheric numerical models and its impact on regional meteorology, Mon. Weather Rev., 117, 2113–2136, 1989.

Baret, F., Hagolle, O., Geiger, B., Bicheron, P., Miras, B., Huc, M., Berthelot, B., Nino, F., Weiss, M., Samain, O., Roujean, J., and Leroy, M.: LAI, FAPAR and FCover CYCLOPES global

GMDD

5, 3573–3620, 2012

ECOCLIMAP-II/Europe

S. Faroux et al.

products derived from Vegetation. Part 1: principles of the algorithm, Remote Sens. Environ., 110, 305–316, 2007.

Bartholomé, E. and Belward, A.: GLC2000: A new approach to global land cover mapping from Earth observation data, Int. J. Remote Sens., 26, 1959–1977, 2005.

Bicheron, P., Leroy, M., Brockmann, C., Krämer, U., Miras, B., Huc, M., Nino, F., Defourny, P., Vancutsem, C., Arino, O., Ranéra, F., Petit, D., Amberg, V., Berthelot, B., and Gross, D.: GLOBCOVER: a 300 m global land cover product for 2005 using ENVISAT/MERIS time series, Proceedings of the Recent Advances in Quantitative Remote Sensing Symposium, Valencia, 2006.

Bonan, G., Levis, S., Sitch, S., Vertenstein, M., and Oleson, K.: A dynamic global vegetation model for use in climate models: concepts and description of simulated vegetation dynamics, Glob. Change Biol., 9, 1543–1566, 2003.

Boone, A., de Rosnay, P., Balsamo, G., Beljaars, A., Chopin, F., Decharme, B., Delire, C., Ducharne, A., Gascoin, S., Grippa, M., Guichard, F., Gusev, Y., Harris, P., Jarlan, L., Kergoat, L., Mougin, E., Nasonova, O., Norgaard, A., Orgeval, T., Ottle, C., Poccard-Leclercq, I., Polcher, J., Sandholt, I., Saux-Picart, S., Taylor, C., and Xue, Y. K.: The Amma Land Surface Model Intercomparison Project (ALMIP), B. Am. Meteorol. Soc., 90, 1865–1880, 2009.

Dawson, T., North, P., Plummer, S., and Curran, P.: Forest ecosystem chlorophyll content: implications for remotely sensed estimates of net primary productivity, Int. J. Remote Sens., 24, 611–617, 2003.

De Colstoun, B., Hall, F., Meeson, B., Los, S., and Landis, D.: The International Satellite Land Surface Climatology Project (ISLSCP) Initiative II data collection, Proceedings IEEE International Geoscience And Remote Sensing Symposium, 4, 2326–2327, 2002.

Defries, R., Hansen, M., and Townshend, J.: Global Discrimination of land cover types from metrics derived from AVHRR Pathfinder data, Remote Sens. Environ., 54, 209–222, 1995.

Di Gregorio, A. and Jansen, L.: Food & the United Nations, Land cover classifications system (LCCS), classification concepts and user manual, Rome, Italy, 2000.

EC: Regionalization and stratification of European forest ecosystems, European Commission, 1995.

EEA: Sustainable use and management of natural resources, European Environment Agency, 2005.

FAO: Climate change and food security: a framework document, Food and Agriculture Organisation of the United Nations, Italy, 2008.

GMDD

5, 3573–3620, 2012

ECOCLIMAP-II/Europe

S. Faroux et al.

Friedl, M., Mciver, D., Hodges, J., Zhang, X., Muchoney, D., Strahler, A., Woodcock, C., Gopal, S., Schneider, A., Cooper, A., Baccini, A., Gao, F., and Schaaf, C.: Global land cover mapping from MODIS: algorithms and early results, Remote Sens. Environ., 83, 287–302, 2002.

Fritz, S. and See, L.: Quantifying uncertainty and spatial disagreement in the comparison of global land cover for different applications, Glob Change Biol., 14, 1057–1075, 2008.

GEOSS: The global Earth Observation System of Systems GEOSS 10-Year Implementation Plan, 2005.

Gitay, H. and Noble, I.: Plant functional types: their relevance to ecosystem properties and global change. Part I: What are functional types and how should we seek them?, edited by: Smith, T., Shugart, H., and Woodward, F., Cambridge University Press, 122–152, 1997.

Gong, D.-Y. and Shi, P.-J.: Northern hemispheric NDVI variations associated with large-scale climate indices in spring, Int. J. Remote Sens., 24, 2559–2566, 2003.

Hagolle, O., Lobo, A., Maisongrande, P., Cabot, F., Duchemin, B., and De Pereyra, A.: Quality assessment and improvements of temporally composited products of remotely sensed imagery by combination of VEGETATION 1 and 2 images, Remote Sens. Environ., 94, 172–186, 2004.

Han, K., Champeaux, J.-L., and Roujean, J.-L.: Land cover mapping over France in using SPOT-4/VEGETATION data, Remote Sens. Environ., 92, 52–66, 2004.

Hansen, M., Defriers, R., Townshend, J., and Sohlberg, R.: Global land cover classification at 1 km spatial resolution using a classification tree approach, Int. J. Remote Sens., 21, 1331–1364, 2000.

Hartigan, J. and Wong, M.: Algorithm AS 136: A K-means clustering algorithm, Appl. Stat., 28, 100–108, 1979.

Herold, M., Mayaux, P., Woodcock, C., Baccini, A., and Schmullius, C.: Some challenges in global land cover mapping: an assessment of agreement and accuracy between existing 1 km datasets, Remote Sens. Environ., 112, 2538–2556, 2008.

Hill, M. and Donald, G.: Estimating spatio-temporal pattern of agricultural productivity in fragmented landscapes using AVHRR NDVI time series, Remote Sens. Environ., 84, 367–384, 2003.

Holben, B.: Characteristics of maximum-value composite images from temporal AVHRR data, Int. J. Remote Sens., 7, 1417–1434, 1986.

Jain, A., Murty, M., and Flynn, P.: Data clustering: a review, ACM Computing Surveys, 31, 264–323, 2003.

GMDD

5, 3573–3620, 2012

ECOCLIMAP-II/Europe

S. Faroux et al.

Kanungo, T., Mount, D., Netanyahu, N., Piatko, C., Angela, R., and Wu, Y.: An efficient k-means clustering algorithm: Analysis and implementation, IEEE Trans. Pattern Anal. Mach. Intell., 24, 881–892, 2002.

Kaptué Tchuenté, A. T., Roujean, J.-L., and Faroux, S.: ECOCLIMAP-II: An ecosystem classification and land surface parameters database of Western Africa at 1 km resolution for the African Monsoon Multidisciplinary Analysis (AMMA) project, Remote Sens. Environ., 114, 961–976, 2010.

Katpué, T., De Jong, S., Roujean, J.-L., Favier, C., and Mering, C.: Ecosystems mapping at the African continent scale using a hybrid clustering approach based on 1 km resolution multiannual data from SPOT/VEGETATION, Remote Sens. Environ., 115, 452–464, 2011a.

Kaptué Tchuenté, A. T., Roujean, J., Bégué, A., Los, S., Boone, A., Mahfouf, J.-F., Carrer, D., and Badiane, D.: A new characterization of the land surface heterogeneity for use in land surface models, J. Hydrometeorol., 12, 1321–1336, 2011b.

Kattge, J. and Co-authors: A global database of plant traits, Glob. Change Biol., 17, 2905–3935, 2011.

Koeppe, C. E. and De Long, G. C.: Weather and climate, McGraw-Hill, 341 pp., 1958.

Krinner, G., Viovy, N., de Noblet-Ducoudré, N., Ogée, J., Polcher, J., Friedlingstein, P., Ciais, P., Sitch, S., and Prentice, I. C.: A dynamic global vegetation model for studies of the coupled atmosphere-biosphere system, Global Biogeochem. Cy., 19, GB1015, doi:10.1029/2003GB002199, 2005.

Le Moigne, P., Boone, A., Calvet, J.-C., Decharme, B., Faroux, S., Gibelin, A.-L., Lebeaupin Brossier, C., Mahfouf, J. F., Martin, E., Masson, V., Mironov, D., Noilhan, J., Tulet, P., and Van Den Hurk, B.: SURFEX scientific documentation, Météo France, 2009.

Loveland, T., Reed, B., Brown, J. F., Ohlen, D.O., Zhu, Z., Yang, L. and Merchant, J. W.: Development of a global land cover characteristics database and IGBP DISCover from 1 km AVHRR data., Int. J. Remote Sens., 21, 1303–1330, 2000.

Maisongrande, P., Duchemin, B., and Dedieu, G.: VEGETATION/SPOT: an operational mission for the Earth monitoring; presentation of new standard products, Int. J. Remote Sens., 25, 9–14, 2004.

Masson, V., Champeaux, J.-L., Chauvin, F., Meriguet, C., and Lacaze, R.: A Global Database of Land Surface Parameters at 1-km Resolution in Meteorological and Climate Models, J. Climate, 16, 1261–1282, 2003.

GMDD
5, 3573–3620, 2012
ECOCLIMAP-II/Europe
S. Faroux et al.

Masson, V. and Coauthors: The SURFEXv7.2 externalized platform for the simulation of Earth surface variables and fluxes, Geosci. Model Dev. Discuss., submitted, 2012.
Matthews, E.: Global vegetation and land use: new high resolution databases for limited studies, J. Climatol. Appl. Meteorol., 22, 474–487, 1983.
Mayaux, P., Bartholomé, E., Fritz, S., and Belward, A.: A new land cover map of Africa for the year 2000, J. Biogeogr., 31, 861–877, 2004.
Molod, A. and Salmun, H.: A global assessment of the mosaic approach to modeling land surface heterogeneity, J. Geophys. Res., 107, 4217, doi:10.1029/2001JD000588, 2002.
Muller, H., Meixner, F., Kramm, G., Fowler, D., Dollard, G. J., and Possanzini, M.: Determination of HNO_3 deposition by modified Bowen ratio and aerodynamic profile techniques, Tellus, 45B, 346–367, 1993.
Noilhan, J. and Mahfouf, J. F.: The ISBA land surface parameterisation scheme, Global Planet. Chang., 13, 145–159, 1996.
Olson, J., Watts, J., and Allison, L.: Carbon in Live Vegetation of Major World Ecosystems, Oak Ridge National Laboratory, Oak Ridge Tennessee, 1983.
Rodell, M., Houser, P., Jambor, U., Gottshalck, J., Mitchell, K., Meng, C.-J., Arsenault, K., Cosgrove, B., Radakovich, J., Bosilovich, M., Entin, J., Walker, J., Lohmann, D., and Toll, D.: The Global Land Data Assimilation System, B. Am. Meteorol. Soc., 85, 381–394, 2004.
Sarrat, C., Noilhan, J., Lacarrere, P., Donier, S., Lac, C., Calvet, J. C., Dolman, A. J., Gerbig, C., Neininger, B., Ciais, P., Paris, J. D., Boumard, F., Ramonet, M., and Butet, A.: Atmospheric CO_2 modeling at the regional scale: Application to the CarboEurope Regional Experiment, J. Geophys. Res.-Atmos., 112, D12105, doi:10.5194/bg-4-1115-2007, 2007.
Still, C., Berry, J., Collatz, G., and DeFries, R.: Global distribution of C3 and C4 vegetation: Carbon cycle implications, Global Biogechem. Cy., 17, 1006–1019, 2003.
Sutherland, W. and Co-authors: One hundred questions of importance to the conservation of global biological diversity, Conservation Biol., 23, 557–567, 2009.
Telesca, L. and Lasaponara, R.: Fire-induced variability in satellite SPOT-VGT NDVI vegetational data, Int. J. Remote Sens., 37, 3087–3095, 2006.
Tucker, C.: Red and photographic infrared linear combinations for monitoring vegetation, Remote Sens. Environ., 8, 127–150, 1979.
Vitousek, P., Mooney, H., Lubchenco, J., and Melillo, J.: Human domination of earth ecosystems, Science, 277, 494–499, 1997.

GMDD
5, 3573–3620, 2012
ECOCLIMAP-II/Europe
S. Faroux et al.

Weiss, M., Baret, F., Garrigues, S., and Lacaze, R.: LAI and fAPAR CYCLOPES global products derived from VEGETATION, Part 2: validation and comparison with MODIS collection 4 products, Remote Sens. Environ., 110, 317–331, 2007.

Wilson, M. and Hernderson-Sellers, A.: A global archive of land cover and soils data for use in general circulation models, J. Climatol., 5, 119–143, 1985.

Yang, W., Tan, B., Huang, D., Rautiainen, M., Shabanov, N., Wang, Y., and Jeffrey, L. P.: MODIS leaf area index products: from validation to algorithm improvement, IEEE Trans. Geosci. Remote Sens., 44, 1885–1898, 2006.

GMDD

5, 3573–3620, 2012

ECOCLIMAP-II/Europe

S. Faroux et al.

Table 1. Summary of the main characteristics of the data used in this study.

Product	10 day NDVI SPOT VGT	8 day LAI MODIS	10 day LAI CYCLOPES	CLC2000	CLC2006	GLC2000
Project or reference	Maisongrande et al. (2005)	Yang et al. (2006)	Baret et al. (2007)	CORINE	CORINE	Bartholme and Belward (2005)
Geographic extent	[25° N, 75° N] × [11° W, 62° E]	[30° N, 70° N] × [10° W, 60° E]	[30° N, 7° 0N] × [10° W, 60° E]	European Union	European Union	Global and regional
Spatial resolution	1/112°	1/120°	1/112°	100 m	100 m	1/112°
Projection	Lat-lon	ISG	Lat-lon	Lambert	Lambert	Lat-lon
Period considered	Jan 1999–Dec 2005	Feb 2000–Dec 2005	Jan 1999–Dec 2003	2000	2006	2000
Input (I) or Validation (V)	I	I	I	I	V	I

Product	GLOBCOVER	Koeppe classification	FIRS	Agricultural statistics	C4 plants	FORMOSAT
Project or reference	Bicheron et al. (2006)	Koppe and De Long (1958)	EC (1995)	AGRESTE	ISLSCP-II	CESBIO
Geographic extent	Global	Global	European Union	France	Global	60 km × 60 km, Upper Left corner: [0°58′6.72″ E, 43°36′5.61″ N]
Spatial resolution	1/360°	1°	1/112°	French department (hectares)	1°	20 m
Projection	Lat-lon	Lat-lon	Lambert		1°	Lambert II extended
Period considered	2005	1958	1995	Average 2002–2005	2003	2002–2005
Input (I) or Validation (V)	V	I	I	V	V	V

GMDD
5, 3573–3620, 2012

ECOCLIMAP-II/Europe

S. Faroux et al.

Table 2. Statistics of comparison between SPOT/VGT NDVI and MODIS LAI for the temporal profiles of ECOCLIMAP-II classes. 25 % of correlations are higher than the value for Q1; 50 % are higher than the median value; 75 % are higher than the Q3 value.

Correlation	Min	Q1	Median	Q3	max
LAI CYCLOPES/NDVI	−0.30	0.85	0.93	0.95	0.99
LAI MODIS/NDVI	−0.29	0.84	0.92	0.95	0.99
LAI CYCLOPES/MODIS	−0.36	0.84	0.94	0.97	0.99

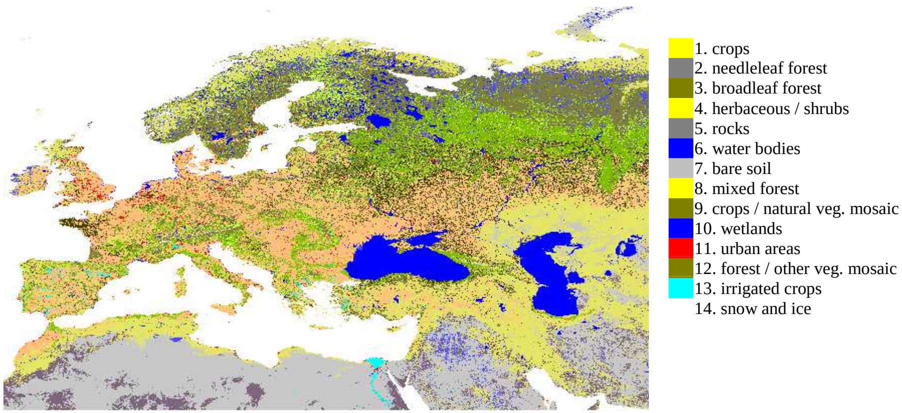

Fig. 1. Classification map with 14 main land covers (C14) resulting from the combination of CLC2000 and GLC2000.

GMDD
5, 3573–3620, 2012

ECOCLIMAP-II/Europe

S. Faroux et al.

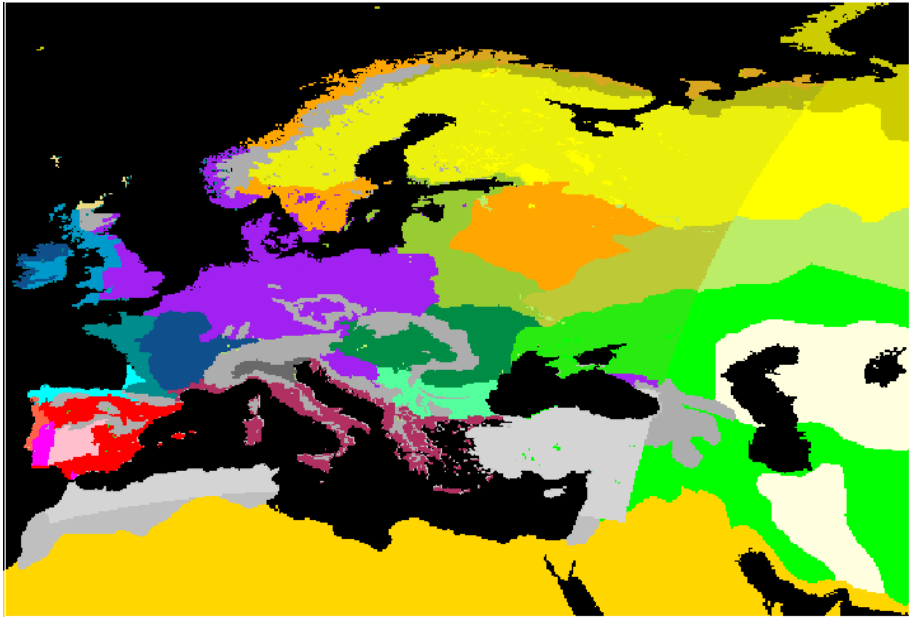

Fig. 2. Climate map built using the FIRS climate map completed by the climate map of Koeppe and De Long (1958) for the eastern part of the domain.

GMDD
5, 3573–3620, 2012

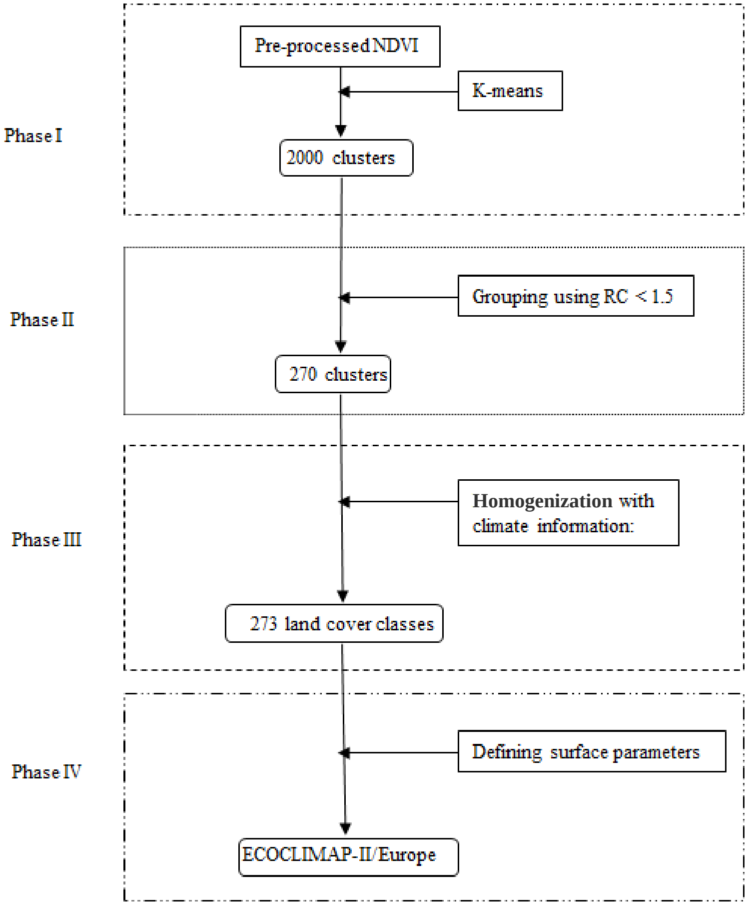

Fig. 3. Schematics of the different steps for the building of the map.

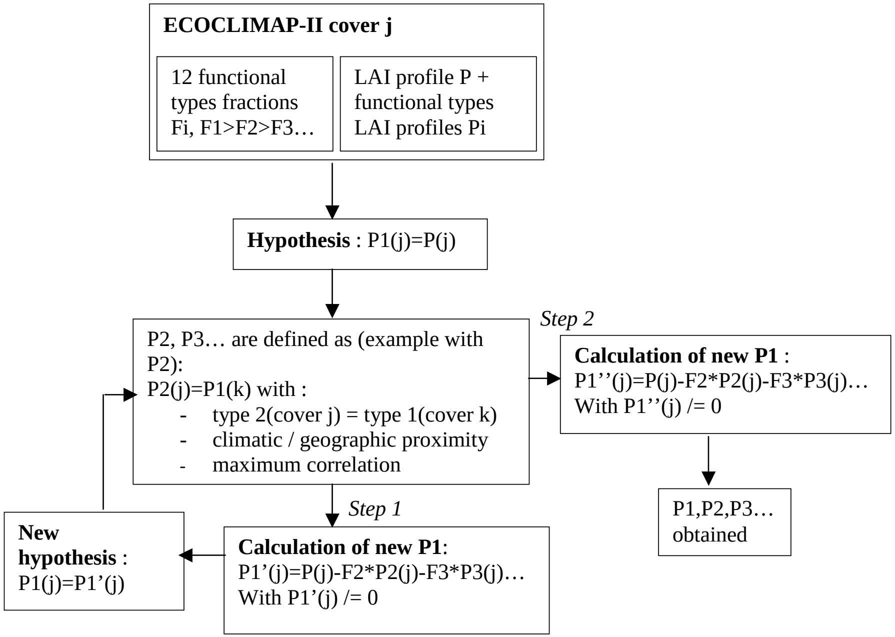

Fig. 4. Disentanglement of LAI profiles from covers to functional types.

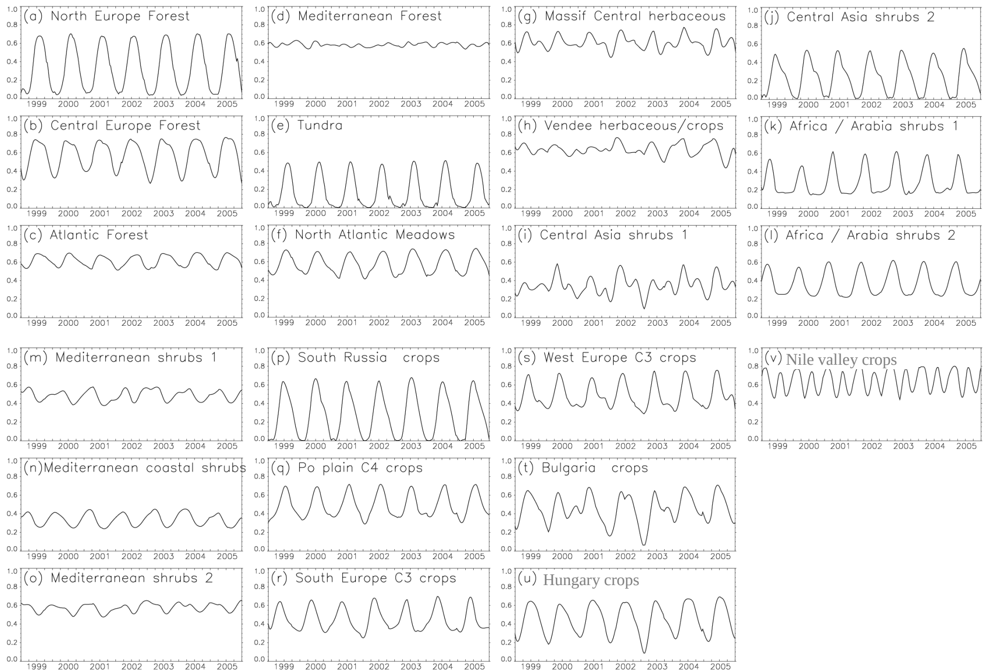

Fig. 5. Examples of NDVI profiles for several covers of ECOCLIMAP-II/Europe.

GMDD

5, 3573–3620, 2012

ECOCLIMAP-II/Europe

S. Faroux et al.

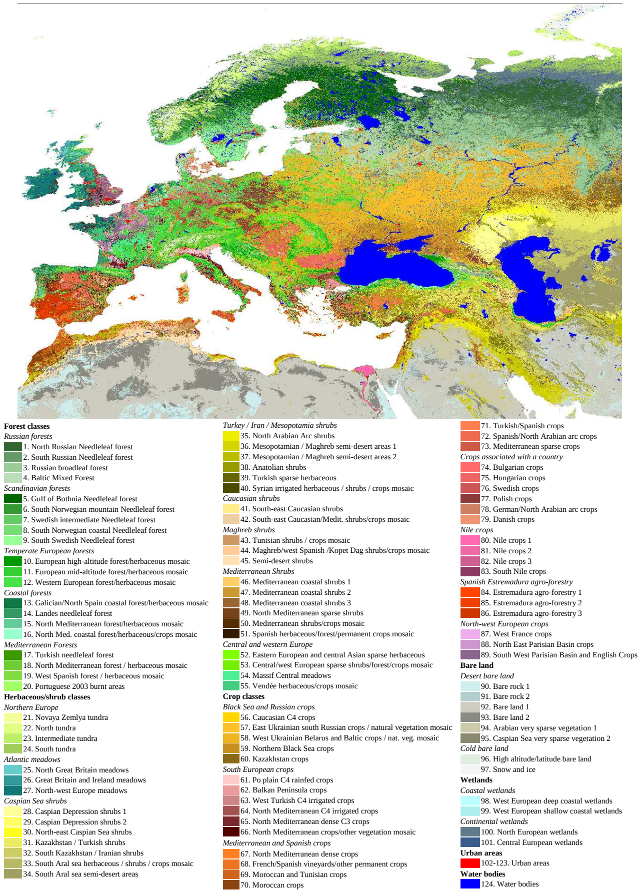

Fig. 6. Simplified map ECOCLIMAP-II/Europe with 103 classes enhancing the dominant patterns.

GMDD
5, 3573–3620, 2012
ECOCLIMAP-II/Europe
S. Faroux et al.

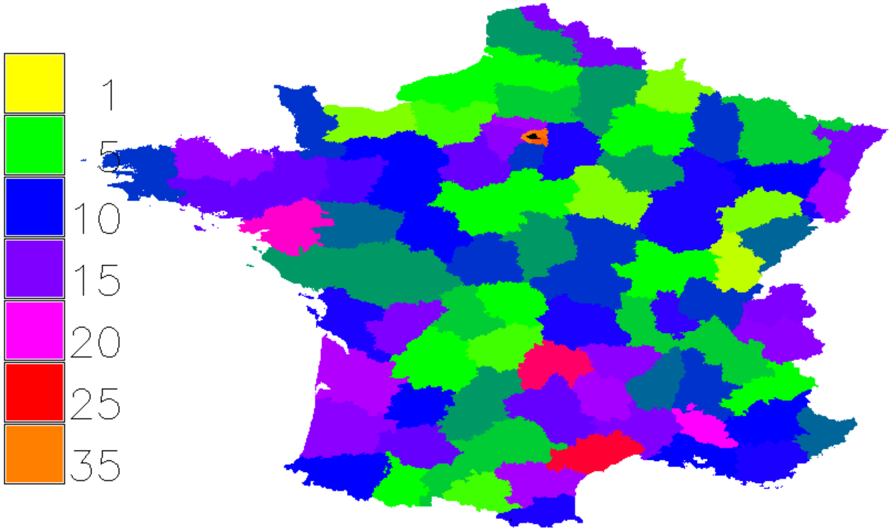

Fig. 7. Results of the comparison of ECOCLIMAP-II and AGRESTE for each administrative department of France. Quantities plotted are cumulated errors (in percentage) for the six land use types examined according to Eq. (2).

GMDD
5, 3573–3620, 2012

Title Page
Abstract | Introduction
Conclusions | References
Tables | Figures
◂◂ | ▸▸
◂ | ▸
Back | Close
Full Screen / Esc
Printer-friendly Version
Interactive Discussion

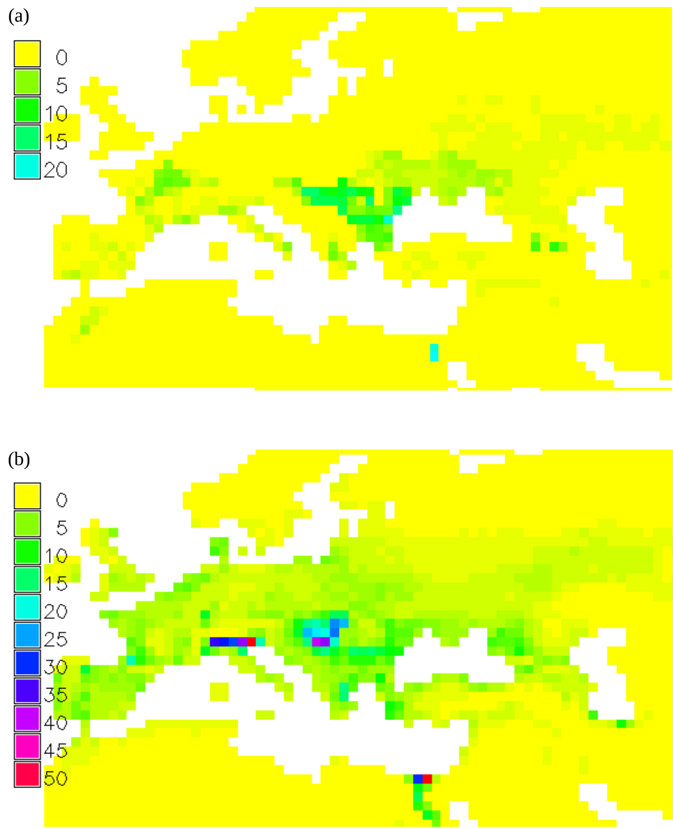

Fig. 8. Fraction (%) of C4 plants at 1° resolution: **(a)** ISLSCP-II; **(b)** ECOCLIMAP-II.

GMDD
5, 3573–3620, 2012

ECOCLIMAP-II/Europe

S. Faroux et al.

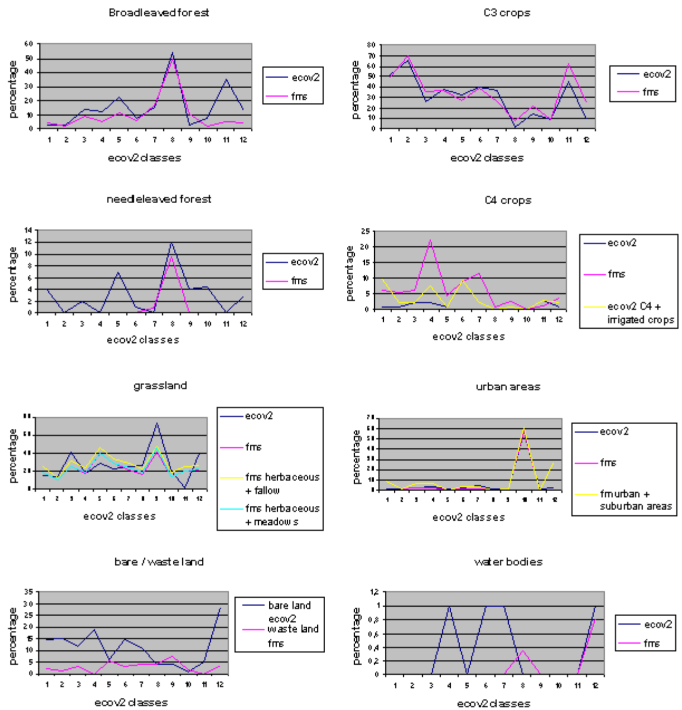

Fig. 9. Comparison of ECOCLIMAP-II and FORMOSAT for 12 ECOCLIMAP-II covers (abscissas) and 8 land cover types (plots). The curves show percentages of the land cover types considered inside ECOCLIMAP-II covers, according to ECOCLIMAP-II (blue) and FORMOSAT (magenta).

GMDD
5, 3573–3620, 2012
ECOCLIMAP-II/Europe
S. Faroux et al.

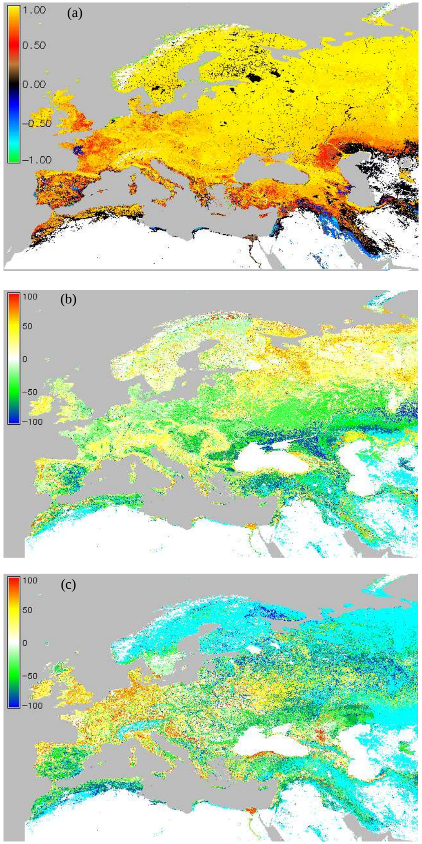

Fig. 10. Comparison of ECOCLIMAP-II and ECOCLIMAP-I LAI: Pearson's correlation coefficient **(a)**, normalized difference (%) of maximum LAI **(b)**, and minimum LAI **(c)** where the normalized difference is defined as the difference between ECOCLIMAP-II and ECOCLIMAP-I divided by the sum of the two.

GMDD
5, 3573–3620, 2012
ECOCLIMAP-II/Europe
S. Faroux et al.

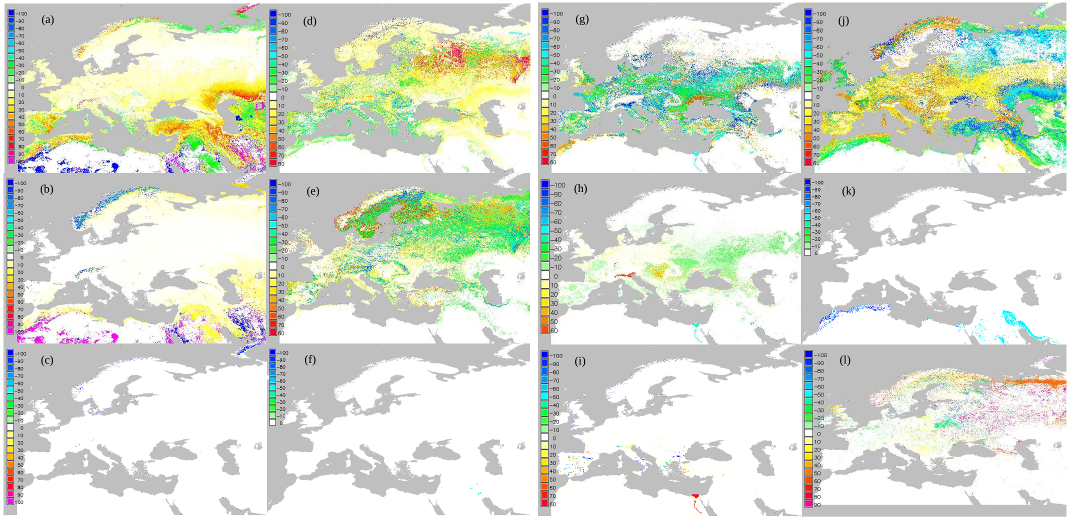

Fig. 11. Comparison of ECOCLIMAP-II and ECOCLIMAP-I fractions of vegetation types: **(a)** bare land, **(b)** bare rock, **(c)** snow, **(d)** deciduous broadleaved forest, **(e)** needle-leaved forest, **(f)** evergreen broadleaved forest, **(g)** C3 crops, **(j)** temperate grasslands, **(h)** C4 crops, **(k)** tropical grasslands, **(i)** irrigated crops, and **(l)** wetlands.

GMDD
5, 3573–3620, 2012

ECOCLIMAP-II/Europe

S. Faroux et al.