Global biogeographical pattern of ecosystem functional types derived from earth observation data

Ivits, Eva; Cherlet, Michael; Horion, Stéphanie Marie Anne F; Fensholt, Rasmus

Published in:
Remote Sensing

DOI:
10.3390/rs5073305

Publication date:
2013

Document version
Publisher's PDF, also known as Version of record

Citation for published version (APA):
Global Biogeographical Pattern of Ecosystem Functional Types Derived From Earth Observation Data

Eva Ivits 1,*, Michael Cherlet 1, Stephanie Horion 2 and Rasmus Fensholt 2

1 Land Resource Management Unit, European Commission Joint Research Centre, I-21027 Ispra, Italy; E-Mail: michael.cherlet@jrc.ec.europa.eu
2 Department of Geosciences and Natural Resource Management, Faculty of Science, University of Copenhagen, Oster Voldgade 10, DK-1350 Copenhagen, Denmark;
   E-Mails: stephanie.horion@geo.ku.dk (S.H.); rf@geo.ku.dk (R.F.)

* Author to whom correspondence should be addressed;
   E-Mail: eva.ivits-wasser@ext.jrc.ec.europa.eu; Tel.: +39-332-78-5315.

Received: 22 May 2013; in revised form: 27 June 2013 / Accepted: 1 July 2013 / Published: 10 July 2013

Abstract: The present study classified global Ecosystem Functional Types (EFTs) derived from seasonal vegetation dynamics of the GIMMS3g NDVI time-series. Rotated Principal Component Analysis (PCA) was run on the derived phenological and productivity variables, which selected the Standing Biomass (approximation of Net Primary Productivity), the Cyclic Fraction (seasonal vegetation productivity), the Permanent Fraction (permanent surface vegetation), the Maximum Day (day of maximum vegetation development) and the Season Length (length of vegetation growing season) variables, describing 98% of the variation in global ecosystems. EFTs were created based on Isodata classification of the spatial patterns of the Principal Components and were interpreted via gradient analysis using the selected remote sensing variables and climatic constraints (radiation, temperature, and water) of vegetation growth. The association of the EFTs with existing climate and land cover classifications was demonstrated via Detrended Correspondence Analysis (DCA). The ordination indicated good description of the global environmental gradient by the EFTs, supporting the understanding of phenological and productivity dynamics of global ecosystems. Climatic constraints of vegetation growth explained 50% of variation in the phenological data along the EFTs showing that part of the variation in the global phenological gradient is not climate related but is unique to the Earth Observation derived variables. DCA demonstrated good correspondence of the EFTs to global climate and also to land use classification. The results show the great potential of Earth Observation derived
parameters for the quantification of ecosystem functional dynamics and for providing reference status information for future assessments of ecosystem changes.

**Keywords:** ecosystem function; gradient analysis; Earth Observation; phenology; productivity

1. Introduction

There is growing evidence that ecosystem degradation is largely due to intensive human resource use together with increasing awareness of the dependence of the global population on the globe’s ecosystem services [1–3]. Ecosystem degradation arises from long lasting loss of vegetation cover and biomass productivity over time and in space [4], as result of a combination of natural/environmental and human/socio-economic drivers. The need for improved information on the status of lands and on degradation processes, along with standardized methodologies and understandable mapping, has to be addressed urgently. To tackle these complex global challenges, a monitoring and assessment system offering up-to-date information on the status and trends of ecosystem degradation and their causes and effects need to be developed. A useful monitoring and assessment system will supply indicators that account for the climate dependence of ecosystem functioning that is responsive to land cover and land use change while supplying knowledge of the temporal and spatial patterns of ecosystem dynamics at larger spatial scales. At global scales, such indicators are best derived from information on vegetation/land cover dynamics as core components to assess the state of ecosystem degradation [5]. In particular, various aspects of vegetation productivity and phenology dynamics, reflecting land cover/use transitions that might lead to ecosystem degradation, need to be considered in a spatio-temporal context [6].

The Local Net Primary Productivity Scaling (LNS) approach relates actual ecosystem degradation to the potential degradation of ecosystems [7,8] where remote sensing estimated productivity of each pixel is expressed relative to the 90 percentile observed in all pixels falling within the same homogeneous ecosystem unit. Thus the LNS method requires the characterization of ecosystem units with similar biomass production potential, which might be defined by biophysical information on vegetation, soils, terrain and climate [7]. Besides this biophysical information ecosystems may be also characterized by the physiognomy and functional dynamics of the vegetation cover. Plant Functional Types (PFTs) are characterized by the physiognomy and functional dynamics of the vegetation and thus offer a possible framework for the identification of such homogenous ecosystem units providing a powerful tool in global change research [9–11]. This is because PFTs summarise the complexity of individual species and populations in recurrent patterns of plants that exhibit similar responses to biophysical environmental conditions. Hence, PFTs bridge the spatial and functional gap between plant physiology, land biophysical properties and ecosystem processes. Furthermore, identifying PFTs enables the assessment of ecosystem functions that might support predicting ecosystem response to human-induced changes [12], because the functional composition of PFTs changes in response to human disturbances [13,14] as well as in response to global change of ecosystems [15]. During the last decades the concept of Ecosystem Functional Types (EFTs) was developed for global scale assessment of ecosystems mostly based on the concept of PFTs, being defined as land surface areas with similar productivity (carbon gain/loss) dynamics.
Because of the large areal coverage and continuous temporal sampling, remotely sensed data provides a synoptic representation of vegetation dynamics in space and time and thus have a great potential for the assessment of ecosystems from regional to global scales [16]. Satellite derived vegetation indices, such as the NDVI, have proven their value in the field of large scale monitoring of vegetation cover and dynamics for land degradation and desertification studies (for an overview see [4,7,17]). Especially the de-convolution of the original time series into phenological and productivity metrics yield additional information on various aspects of vegetation dynamics and ecosystem functioning [18] that can be related to land use [19]. Monitoring vegetation phenology using satellite remote sensing [20–25] offers an optimal framework for ecosystem studies because \textit{in situ} phenological data are comparatively scarce in many parts of the world [26] and because long time and large spatial scales can be addressed simultaneously. Monitoring vegetation phenology has also improved the understanding of the interactions between the biosphere, the climate and biogeochemical cycles [16,27–29]. Lloyd [30] has proposed the use of remote sensing derived phenology for the description of ecosystem functioning whereas Soriano and Paruelo [31] proposed the classification of biozones for the assessment of EFTs based on the seasonal dynamics of vegetation indices derived from satellite observations. Using remote sensing for the evaluation of the functional components of the ecosystems and the characterisation of EFTs has been gaining importance in recent decades (see [32] for a thorough review).

There is a need for quantitative methods to map the status and/or change of ecosystem services and thus supplying a standardized framework for ecosystem degradation studies. Therefore, in the present study we aim at identifying global EFTs [12,31–34] in order to supply ecosystem degradation studies relating actual to potential degradation with homogenous biophysical units such as the LNS method [8]. In our study, EFTs are defined as spatial units with similar patterns of seasonal phenology and productivity dynamics, also showing similar responses to changing natural and human induced environmental conditions following the ideas of Paruelo \textit{et al.} [12] and Stow \textit{et al.} [34]. Seasonal phenological and productivity dynamics are calculated from the GIMMS3g NDVI time-series (1982–2010) by decomposing the yearly NDVI signal into various metrics. Since we describe the functioning of global ecosystems within large climatic and biophysical gradients we derived a larger set of parameters indicating vegetation seasonal dynamics when compared to Paruelo \textit{et al.} [12] and Alcaraz-Seguera \textit{et al.} [35–37]. As human activities such as abrupt land cover/land use change also play a significant role in land degradation processes [7], EFTs should also relate to land use. Therefore, when characterising ecosystems, a compound set of functional attributes that describe vegetation dynamics should also be included. We illustrate that the global EFTs reflect climate and land use and therefore offer a meaningful, transparent and objective stratification for global ecosystem monitoring studies.

2. Data and Methods

2.1. Data

The Normalised Difference Vegetation Index (NDVI) dataset was acquired from the Global Inventory Modeling and Mapping Studies (GIMMS) project. The dataset covers the period July 1981
to December 2011 and is produced as bi-monthly maximum-value composite (MVC) NDVI images referred to as NDVI3g (third generation GIMMS NDVI). The channel 1 and 2 data used for the GIMMS data are calibrated as suggested by Vermote and Kaufman [38], and the derived NDVI is further adjusted using the technique of Los [39]. The NDVI3g applies an improved cloud masking as compared to older versions of the GIMMS dataset and the cloud detection algorithm is based on reflectance and brightness temperature values [39,40]. No atmospheric correction is applied to the GIMMS data except for volcanic stratospheric aerosol periods (1982–1984 and 1991–1994) [41]. A satellite orbital drift correction is performed using an empirical mode decomposition/reconstruction (EMD) method of Pinzon et al. [42] minimizing effects of orbital drift by removing common trends between time series of solar zenith angle (SZA) and NDVI. The GIMMS3g NDVI data is provided in 1/12-degree resolution. Only vegetated areas were included in the present study by masking out pixels with a standard deviation of NDVI values below 0.02 [43].

In different parts of the world temperature, radiation, and water interact to impose complex and varying limitations on vegetation activity [44]. In order to assess these limitations over the derived EFTs, global climatic plant growth constraints [29] were included in the present study. These climatic plant growth constraint data were created from long-term climate statistics and includes gridded information of the relative importance of water availability (wat), air temperature (temp) and incoming shortwave radiation (rad) for vegetation across the globe. In order to demonstrate the correspondence of the EFTs in relation to climatic zones the updated Köppen-Geiger classification was acquired for the study [45]. Information on land use and management is also indispensable if the goal is to understand global environmental change and the loss of ecosystem services. Only a few global databases are available that allow the characterization of the land management concentrating on irrigation, the presence of livestock and of protected areas. In order to address land use and management on the global scale the present study utilized the Land Use Systems (LUS) maps of the world produced by the United Nations Food and Agricultural Organization [46]. The LUS classification includes several layers as the Global Land Cover dataset (GLC 2000), agro maps indicating cropping patterns (e.g., dominant crop types), livestock data, major ecosystem types, biophysical resources base layers (e.g., temperature regime, length of growing period, dominant soil units and terrain information) and socio-economic attributes.

2.2. Derivation of Phenological and Productivity Parameters

Phenological and productivity parameters were derived using the Phenolo software developed in the EC Joint Research Centre [19]. In order to generate results comparable between data sources with different time aggregation windows, daily values were calculated by an iterative linear interpolation of the NDVI time-series decades. Subsequently, a Savitzky-Golay filter with fourth polynomial degrees and a length of 50 days was applied to the time series in order to identify and remove short peaks and drop-offs caused by noise. This pre-processing resulted in the reference time series on which a set of phenological and productivity parameters were computed. The methodology is based on computing a delayed and a forward shifted time-series using a moving average filter and extracting the intersection points of these with the reference time-series. While Reed et al. [47] determined a general lag based on a priori knowledge about the average phenology of the study area, Phenolo is strictly data driven using
for each individual pixel its time series dynamics to determine the lag in order to render it applicable to a wide geographical extent (see [19] for a complete description of the method). Once the intersection points are defined a series of phenological and productivity metrics were computed for each year in the time-series (see Figure 1).

**Figure 1.** Schematic representation of the phenological and productivity variables calculated by Phenolo. Phenolo calculates all the areas under the time-series curve and all time and value of the depicted points on the yearly time-series curve. In the present study only the most important parameters were employed, which are listed on the right hand side.

### Productivity parameters:

- **SB** = a+b+c+d+e+f+g (Standing Biomass)
- **MPI** = a+b+c (Minimum Permanent Integral)
- **SBV** = Season Begin Value of the vegetation index
- **MXV** = Maximum Value of the vegetation index
- **MXD** = Maximum Day
- **SBD** = Season Begin Day
- **SER** = d+f (Season Exceeding Residual Integral)
- **CF** = g (Cyclic fraction)
- **PF** = d+e+f (Permanent Fraction)

**Approximation of the Total Surface Biomass.**

**Area characterising the perennial vegetation component and depending on the vegetation index used may also contain components of soil substrate.**

**The amount of senescent vegetation outside of the growing season. May be more indicative to before/after harvest of agricultural areas.**

**This parameter may be directly related to purely seasonal growth.**

**The amount of vegetation that does not have a characteristic seasonal cycle within the growing season.**

### Phenological parameters:

- **SBD** = Season Begin Day.
- **SL** = Season Length.
- **MXV** = Maximum Value of the vegetation index.
- **MXD** = Maximum Day.
- **SBV** = Season Begin Value of the vegetation index.

### 2.3. Statistical Screening of the Phenological and Productivity Parameters

The yearly phenological and productivity variables were summarized in their temporal mean (1982–2010) and consecutively screened for multicollinearity based on the correlation matrix. Variables with very high correlation (>0.7) were removed from the analysis. Principal Component Analysis (PCA) was run on the Phenolo variables using the correlation rather than the covariance matrix in order to standardize the remote sensing variables with different measurement scales. This first, “screening PCA” served for (1) the selection of that set of Principal Components (PC) that explained the highest amount of total variance in the Phenolo variables and (2) for the selection of those Phenolo variables that showed the highest loadings on the selected components. The loadings of the individual variables were normalized by multiplication with the square root of the eigenvalues in order to present the values as correlation with the PCA axes. Once the n number of Principal Components was determined and the remote sensing variables with the highest loadings were selected, the “final PCA” was run on the selected n number of variables extracting n number of axes. The PCs were rotated with the “varimax” technique [48] in order to clearly associate each PC axes with one Phenolo variable only. The spatial patterns of the final rotated components were calculated by multiplying the loadings with the selected phenological variables.
2.4. Classification of Global Ecosystem Functional Types Based on the Phenological and Productivity Parameters

An ISODATA cluster analysis was used to identify major Ecosystem Functional Types (EFTs) at global scale. The analysis was run on the spatial patterns of the final rotated components because the PCA with correlation matrices normalized the eigenvectors with zero mean and 1 SD. This is a statistical property desirable for the calculation of class means evenly distributed in the data space and for the iterative clustering of the pixels using minimum distance techniques. Several test runs showed that if the range of the number of classes requested was allowed to be large (min = 10 and max = 500) the number of iterations had the highest influence on the resulting number of clusters. The algorithm was run 15 times where for each run the number of iterations increased with 1 (from 1 to 15) and another 5 times with 20, 25, 30, 50 and 100 iterations. We used discriminant analyses on the resulting EFTs of the 20 ISODATA classification runs to assess the number of EFTs best separated by the rotated eigenvectors. For each discriminant analysis we calculated the Wilks’ lambda and the canonical correlation. The former is a multivariate test of significance, where values close to 0 indicate that the group means are different whereas values close to 1 indicate similar group means. The canonical correlations of the discriminant functions indicate the proportion of total variability explained by differences between the groups. We have selected the ISODATA iteration run for which the Wilks’ lambda measure was lowest and the canonical correlation was highest. This run created EFTs significantly different in their phenology and productivity. To facilitate summarizing and reporting results the EFTs were further classified into a smaller number of clusters. For this, the rotated eigenvectors were spatially averaged within the EFTs and submitted to a dendrogram analysis joining EFTs into clusters such that the variance within the clusters was minimized. The sum of squared deviations was used as a measure for an EFT to enter the cluster, i.e., an EFT entered a cluster if its inclusion in the cluster produced the least increase in the error as measured by the sum of squared deviations.

2.5. Relation of the EFTs to the Phenological and Productivity Parameters and to Climatic Vegetation Growth Constraints

In order to characterize the EFTs in terms of their vegetation productivity and phenology dynamics a gradient analysis was run using another PCA (gradient PCA). The input for this analysis was a data table with the EFTs as rows and the Phenolo variables (selected through the screening PCA and final PCA analysis) averaged within the EFTs as columns. The data was centered and standardized to bring the variables to the same measurement scale. Redundancy Analysis (RDA) was used to assess the association of the EFTs and their phenological and productivity values with the climatic vegetation growth constraint variables (water, temperature and radiation). These variables were also averaged within the EFTs and added to the data table as columns. The significance and F-value of the RDA run was determined with the Monte-Carlo permutation test with 999 permutations. The association of the EFTs, Phenolo variables and climatic constraints were presented in a PCA tri-plot instead of an RDA tri-plot. This indicates correlation of the variables, rather than the variance the climatic variables explain, and thus facilitates focusing interpretation of results on the EFTs. The climatic and Phenolo variables were presented by arrows. Arrows pointing in the same direction in the tri-plot indicate
positive correlation of the variables whereas arrows closing an approx. 90 degree indicate no
correlation between the variables. The EFTs were represented by different icons depending on the
clusters they were classified to with the dendrogram analysis. Projecting the EFTs perpendicularly
onto one of the arrows of the Phenolo or climatic variables shows the value of the variable in that EFT.
This tri-plot interprets the EFTs with vegetation phenological and productivity dynamics and
additionally demonstrates the correlation of the Phenolo variables and the EFTs with the climatic
constraints. Finally, to visualize the relationship of the EFTs with the remote sensing derived vegetation
phenology and productivity variables the PCA values of the EFTs along the first two PCA axes were
mapped. The relationship of the EFTs with radiation, temperature and water constraints of vegetation
growth was visualized by mapping the RDA values of the EFTs along the first two RDA axes.

2.6. Relation of the EFTs to Global Climate and Land Use Classifications

In order to correlate the EFTs of the ecosystems with existing global classifications Detrended
Correspondence Analysis (DCA) was carried out with the Köppen climate zones and with the FAO
Land Use System classes (Table 1), respectively. The LUS classes were intersected with the major
climate zones Tropical (Tr), Arid (Ar), Temperate (Tm), and Cold (Cd) in order to better describe the
climate dependent seasonal dynamics of LUS classes. The input for the analysis was a data table with
the EFTs as rows and the number of pixels belonging to the Köppen zones respectively to the LUS
classes within the given EFTs as columns. The DCA axes were detrended by second order polynomials
in order to avoid the arch effect caused by the strong environmental gradients [49] and for better
interpretability of the ordination tri-plot. The tri-plots of the first two DCA dimensions were presented
plotting the EFTs, the Köppen zones respectively the LUS classes and the selected Phenolo variables.
Close Euclidean distances between the spatial units of the EFTs and the Köppen classes respectively
the FAO land use classes indicate good correspondence with the EFTs. The spatial patterns of the
DCA ordinations were presented in maps plotting the distance of each EFT from the center of the
biplot. These maps further facilitated the interpretation of the DCA tri-plots. In order to find EFTs with
low correspondence to the Köppen zones and LUS classes respectively, the percentage of fit of each
EFTs in the ordination space was mapped by calculating the chi-square distance between the EFTs and
the centroid of all samples.

Table 1. Köppen climatic zones and FAO Land Use Systems classes used in the present study.

<table>
<thead>
<tr>
<th>LUS Classes</th>
<th>Code</th>
<th>Köppen Climatic Zones</th>
<th>Code</th>
<th>Köppen Climatic Zones</th>
<th>Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>Forests, unmanaged</td>
<td>FRu</td>
<td>Tropical rainforest</td>
<td>Af</td>
<td>Temperate, dry winter, cold summer</td>
<td>Cwc</td>
</tr>
<tr>
<td>Forests, managed</td>
<td>FRm</td>
<td>Tropical Monsoon</td>
<td>Am</td>
<td>Cold, dry and hot summer</td>
<td>Dsa</td>
</tr>
<tr>
<td>Grassland, unmanaged</td>
<td>GRU</td>
<td>Tropical Savannah</td>
<td>Aw</td>
<td>Cold, dry and warm summer</td>
<td>Dsb</td>
</tr>
<tr>
<td>Grassland, managed</td>
<td>GRm</td>
<td>Arid, desert, hot</td>
<td>Bwh</td>
<td>Cold, dry and cold summer</td>
<td>Dsc</td>
</tr>
<tr>
<td>Shrubland, unmanaged</td>
<td>SHu</td>
<td>Arid, desert, cold</td>
<td>Bwk</td>
<td>Cold, dry summer, very cold winter</td>
<td>Dsd</td>
</tr>
<tr>
<td>Shrubland, managed</td>
<td>SHm</td>
<td>Arid, steppe, hot</td>
<td>Bsh</td>
<td>Cold, dry winter, hot summer</td>
<td>Dwa</td>
</tr>
<tr>
<td>Rainfed agriculture</td>
<td>AGr</td>
<td>Arid, steppe, cold</td>
<td>Bsk</td>
<td>Cold, dry winter, warm summer</td>
<td>Dwb</td>
</tr>
<tr>
<td>Irrigated agriculture</td>
<td>AGi</td>
<td>Temperate, Dry and hot summer</td>
<td>Csa</td>
<td>Cold, dry winter, cold summer</td>
<td>Dwc</td>
</tr>
</tbody>
</table>
Table 1. Cont.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Wetlands</td>
<td>WTL</td>
<td>Temperate, Dry and warm summer</td>
<td>Csb</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Temperate, no dry season, hot summer</td>
<td>Cfa</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Temperate, no dry season, warm summer</td>
<td>Cfb</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Temperate, no dry season, cold summer</td>
<td>Cfc</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Temperate, dry winter, hot summer</td>
<td>Cwa</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Temperate, dry winter, warm summer</td>
<td>Cwb</td>
</tr>
</tbody>
</table>

3. Results

3.1. Statistical Screening of the Phenological and Productivity Parameters

The first five rotated axes of the screening PCA (10 input variables as in Figure 1) explained 98.6% of the variation in the global ecosystems (Table 2). The first axis alone explained 41.1% of the variation while the second, third, fourth and fifth axes explained 19.4%, 17.4%, 12.5% and 8.2%, respectively. The final PCA was run with the five variables loading highest on the first five rotated axes: Standing Biomass (SB), Cyclic Fraction (CF), the Maximum Day (MXD), the Season Length (SL) and the Permanent Fraction (PF) (Figure 1). Standing Biomass and Maximum Day have been used in other studies and are similar to EVI mean, NDVI sum and DMAX [10,32,33], while the other variables are new, extending the available set of phenological and productivity metrics in view of characterizing global EFTs. Some differences between the global distribution of the Standing Biomass as a productivity measure calculated in this study and spatial results from other productivity models (for a review see [48]) are partly due to the different nature of the methods. Standing Biomass is a yearly integration of all observed above ground biomass whereas NPP models are based on the fixation rate of atmospheric CO₂. The Cyclic Fraction, representing a rate of annual active biomass growth, is highest over high latitude areas with frequent snow cover. This reflects the very strong intra-annual dynamics of seasonal vegetation cover that starts from very low values due to snow cover. Low Cyclic Fraction values are observed over areas either without seasonal snow cover or with very low Standing Biomass (arid areas) or over evergreen vegetation cover lacking yearly greening up dynamics. The Season Length presented in this study and Length of Growing Period (LGP) presented in other studies [50,51] are of different nature. The Season Length variable reflects a full growth cycle of all observed standing vegetation whereas LGP data models refer to the number of days when temperature and moisture conditions are considered adequate for crop growth. On the other hand, short EO-derived Season Length values over the African tropical forest possibly relate to persistent cloud cover affecting satellite data quality in the visible bands. Furthermore, the Maximum Day variable did not consider areas with double or multiple growing seasons and future studies should consider the proper characterization of such territories.

In the final PCA run, the first rotated axis explained 21.2% of the total variance (Table 2) and was mostly associated with the Standing Biomass of the global ecosystems (strongest positive loading, Table 3). The second PCA axis also explained 20.4% of the variation and was associated with the Season Length. The third rotated PCA axis was mostly related to the timing of the NDVI maximum
Remote Sens. 2013, 5

(MXD) of the ecosystems and it explained 20.3% of the variance. The fourth rotated PCA axis was mostly dominated by the Cyclic Fraction explaining 19.9% of the variance whereas the fifth rotated axis represented an environmental gradient dominated by the Permanent Fraction explaining 18.2% of the variation in the data. The Isodata cluster analyses of the spatial pattern of the rotated Principal Components reached highest significance (Wilk’s lambda of 0.121 and canonical correlation of 0.937) at the 10th iteration. This iteration resulted in the creation of 232 global EFTs with significantly different group means and with a high percentage of variance of the rotated Principal Components explained. The 232 EFTs were subsequently regrouped into 15 major categories based on minimizing the variance of the rotated eigenvectors within the EFTs and maximizing the variance between them (Figure 2).

**Table 2.** Principal Component Analysis (PCA) of the phenological and productivity variables. Statistics for the first five varimax rotated components are shown.

<table>
<thead>
<tr>
<th>Component</th>
<th>Eigenvalues</th>
<th>% of Explained Variance</th>
<th>Cumulative % of Explained Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Screening PCA with 10 variables (as in Figure 1)</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>4.1</td>
<td>41.1</td>
<td>41.1</td>
</tr>
<tr>
<td>2</td>
<td>1.9</td>
<td>19.4</td>
<td>60.5</td>
</tr>
<tr>
<td>3</td>
<td>1.7</td>
<td>17.4</td>
<td>77.9</td>
</tr>
<tr>
<td>4</td>
<td>1.3</td>
<td>12.5</td>
<td>90.4</td>
</tr>
<tr>
<td>5</td>
<td>0.8</td>
<td>8.2</td>
<td>98.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Final PCA with 5 selected variables</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1.06</td>
<td>21.2</td>
<td>21.2</td>
</tr>
<tr>
<td>2</td>
<td>1.02</td>
<td>20.4</td>
<td>41.6</td>
</tr>
<tr>
<td>3</td>
<td>1.01</td>
<td>20.3</td>
<td>61.9</td>
</tr>
<tr>
<td>4</td>
<td>0.99</td>
<td>19.9</td>
<td>81.8</td>
</tr>
<tr>
<td>5</td>
<td>0.91</td>
<td>18.2</td>
<td>100.0</td>
</tr>
</tbody>
</table>

**Table 3.** Normalized loadings of the selected phenological variables on the first five rotated axes of the final PCA.

<table>
<thead>
<tr>
<th>Variables</th>
<th>PC1</th>
<th>PC2</th>
<th>PC3</th>
<th>PC4</th>
<th>PC5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Season Length</td>
<td>0.184</td>
<td>0.928</td>
<td>0.086</td>
<td>0.199</td>
<td>0.239</td>
</tr>
<tr>
<td>Permanent Fraction</td>
<td>0.392</td>
<td>0.289</td>
<td>-0.016</td>
<td>0.148</td>
<td>0.861</td>
</tr>
<tr>
<td>Cyclic Fraction</td>
<td>0.032</td>
<td>0.193</td>
<td>0.265</td>
<td>0.936</td>
<td>0.122</td>
</tr>
<tr>
<td>Standing Biomass</td>
<td>0.928</td>
<td>0.178</td>
<td>-0.104</td>
<td>0.024</td>
<td>0.308</td>
</tr>
<tr>
<td>Maximum Day</td>
<td>-0.093</td>
<td>0.074</td>
<td>0.963</td>
<td>0.242</td>
<td>-0.016</td>
</tr>
</tbody>
</table>
3.2. Relation of the EFTs to the Phenological and Productivity Parameters

The first two axes of the gradient PCA ordination explained 79.1% of the variance in the phenological and productivity variables and widely spanned the EFTs indicating good description of the global environmental gradients (Figure 3). The arrows of Standing Biomass and Permanent Fraction point to the same part of the triplot at the positive end of the first gradient PCA axis (clusters 6, 8, 13 and 14) indicating that EFTs with high Standing Biomass tend to have higher Permanent Fractions than the EFTs in other clusters. The Cyclic Fraction and the Maximum Day is almost perpendicular to Standing Biomass indicating that their values appear to be independent at global scale.
Figure 3. Ordination triplot of the gradient PCA analysis. The icons represent the 232 EFTs (numbers) classified into 15 clusters (icons). Full arrows represent the Phenolo variables (see Figure 1) whereas dashed arrows show climatic constraints of vegetation growth (wat = water, temp = temperature, rad = radiation). The horizontal axis shows the first PCA axis (55.9% variation explained) whereas the vertical axis shows the second axes (23.2% variation explained).

Thus, EFTs with high positive scores on the first PCA axis exhibit high Standing Biomass, high Permanent Fraction, longer seasons whereas the timing of their yearly maximum and their Cyclic Fraction can take a large range of values. These EFTs are mapped with high positive values in Figure 4 (top) showing the spatial patterns of the first axis of the gradient PCA ordination of the Phenolo variables averaged within the EFT clusters. High positive values appear covering the Great Plains and the Appalachian Mountains in North America, the most part of Canada, Europe (without the Mediterranean regions) and parts of the Sahel region. EFTs with high negative scores on the first PCA axis exhibit low Standing Biomass, low Permanent Fraction, shorter seasons and low Cyclic Fractions whereas their Maximum Day are large/late, average or low/early, respectively dependent on their tri-plot position. These EFTs are mapped with high negative scores in Figure 4 (top) mostly covering
the Northern Boreal regions, Patagonia, Northern part of the Sahel region and the Horn of Africa and the most part of Australia. The Maximum Day is latest in the EFTs at the negative end of the second PCA axis (Figure 3) with high Cyclic Fraction but low Standing Biomass and Permanent Fraction values. These EFTs are mapped with high negative scores in Figure 4 (bottom, spatial pattern of the second axis of the gradient PCA) covering the Northern/Boreal regions, the Sierra Madre region in Mexico, the Great Indian Desert and the Tibet region in China. High positive values in Figure 4 (bottom) indicate early Maximum Days with large Standing Biomass, moderate Permanent Fraction but low Cyclic Fraction values covering the Coastal Plain in North America, the Tropics, Central East Latin America, Indonesia and the temperate regions of Southern Australia.

**Figure 4.** Spatial patterns of the first (top) and second (bottom) axis of the gradient PCA ordination of the selected Phenolo variables averaged within the EFT clusters. White areas represent pixels without vegetation cover.
3.3. Relation of the EFTs to Climatic Constraints of Vegetation Growth

Redundancy Analysis (RDA) indicated that the climatic constraints radiation, temperature and water explained 50% of the variation in the phenological data (Table 4, sum of canonical eigenvalues) along the gradient represented by the EFTs. This confirms the association of climate and the global ecosystems gradient inherit in the EFTs but it also shows that much of the variation in the global phenological and productivity gradient, derived from Earth Observation, cannot be explained by climate only. This is also shown by the relatively high importance of the fourth, not climate related axis that still explained 22.6% variation in the global ecosystems gradient indicating, that this variability should be attributed to other factors in explaining the spatial distribution and functional types of global ecosystems.

Table 4. Redundancy Analysis (RDA) of the phenological and productivity variables averaged within the EFTs as dependent variables and the climatic constraints averaged within the EFTs as explanatory variables.

<table>
<thead>
<tr>
<th></th>
<th>Axis1</th>
<th>Axis2</th>
<th>Axis3</th>
<th>Axis4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eigenvalues</td>
<td>0.343</td>
<td>0.139</td>
<td>0.018</td>
<td>0.226</td>
</tr>
<tr>
<td>Cumulative % of explained variance</td>
<td>34.3</td>
<td>48.3</td>
<td>50.0</td>
<td>72.7</td>
</tr>
<tr>
<td>Sum of canonical eigenvalues</td>
<td></td>
<td></td>
<td></td>
<td>0.500</td>
</tr>
<tr>
<td>Monte Carlo test: F-ratio/significance</td>
<td></td>
<td></td>
<td></td>
<td>76.135/p &lt; 0.001</td>
</tr>
</tbody>
</table>

The PCA triplot (Figure 3) shows the correlation of the vegetation growth constraint variables with the ordination axes and with the EFTs whereas these relationships are mapped in Figure 5. The first axis is mostly related to radiation (rad, positive correlation) and water (wat, negative correlation) constraints of vegetation growth. Water is a limiting factor in the EFTs with low cluster values on the negative end of the first axis (PC1) covering mostly arid and semi-arid areas. These ecosystems are plotted with negative values in Figure 5 (top) showing the correlation of the climatic constraint variables with the first PC axis.

The EFTs are located in the Sierra Madre region in Mexico, Patagonia, the Sahel region and Horn of Africa and the arid regions of Western South Africa and Australia. In EFTs at the positive end of the first PC axis radiation, instead of water, is the limiting factor for vegetation growth (Figure 3). These EFTs are mapped with high positive PC values on Figure 5 (top) mostly covering the Northern and Boreal regions of Canada and Western US, Atlantic, Continental and Northern Europe and the most part of Russia. The second PCA axis shows a strong negative relation to temperature (temp) constraints of vegetation growth (Figure 3, negative end of the axis). EFTs where temperature is a limiting factor for vegetation growth are mapped with high negative values in Figure 5 (bottom), covering mostly Northern and cold ecosystems. Ecosystems with high positive values in Figure 5 (bottom) exhibit the opposite pattern, i.e., where temperature does not constrain vegetation growth. These ecosystems are mapped with high positive values in Figure 5 (bottom) mostly covering the Tropical and Mediterranean ecosystems.
3.4. Relation of the EFTs to Global Climate Classifications

The first two DCA axes explained 47% of the variance between the EFTs and the Köppen climate zones (Table 5) and as shown by the tri-plot most of the EFT clusters could be associated to one of the zones (Figure 6). For instance, some of the Cold Köppen zones (Dsd, Dwd, Dfd, Dwc, Dsc and Dfc, see Table 1) were associated with the EFTs on the positive end of the first DCA axis with high Cyclic
Fraction, close to (global) average Standing Biomass and moderate Permanent Fraction. These ecosystems are mapped with high positive DCA sample score values on Figure 7 (top).

**Table 5.** Detrended Correspondence Analysis (DCA) of the EFTs with the Köppen climate zones.

<table>
<thead>
<tr>
<th></th>
<th>Axis1</th>
<th>Axis2</th>
<th>Axis3</th>
<th>Axis4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eigenvalues</td>
<td>0.641</td>
<td>0.354</td>
<td>0.157</td>
<td>0.065</td>
</tr>
<tr>
<td>Cumulative % of explained variance</td>
<td>30.1</td>
<td>46.7</td>
<td>54.1</td>
<td>57.1</td>
</tr>
</tbody>
</table>

**Figure 6.** Detrended Correspondence Analysis (DCA) triplot of the EFTs classified in 15 clusters (icons with colors) and the Köppen climate zones (crosses). Arrows represent the phenological and productivity variables averaged within the EFTs. Their positions indicate the correlation with the DCA axes.
Figure 7. Spatial patterns of the first (top) and second (bottom) axis derived from DCA of the EFTs with the Köppen zones. White areas represent pixels without vegetation cover.

On the negative end of the first axis gradient, the Temperate zones were mostly associated with the EFT clusters exhibiting moderate Standing Biomass and Permanent Fractions and low Cyclic Fractions. These ecosystems are mapped with high negative DCA sample score values on Figure 7 (top). On the negative end of the second DCA axis, the Tropical zones (Af, Tropical/Rainforest and Am, Tropical/Monsoon) were strongly associated with the EFT clusters exhibiting high Standing Biomass, moderate Permanent Fraction and low Cyclic Fraction. These clusters are mapped with high negative values on Figure 7 (bottom) whereas EFTs associated with arid ecosystems (Bwh, Bwk, Bsh and Bsk) in the positive end of the second axis exhibit the opposite phenological and productivity pattern and
are mapped with high positive values. The phenological values Season Length and Maximum Day reacted only moderately to the ordination space (shorter arrows) indicating small rate of change and thus, are less important in distinguishing the climate zones. The EFTs discriminated the Cold climate zones especially well at the positive end of the first DCA axis (Figure 6) where the Köppen zones are plotted with larger distances between the individual zones. Köppen zones are also well discriminated in the Arid EFTs at the positive end of the second DCA axis whereas the in the Tropical and Temperate EFTs the Köppen zones display a more aggregated thus less interpretable pattern. Figure 8 presents for each EFTs the % of fit of the ordination with the Köppen zones. Most EFTs presented a good (>50%) or very good (>70%) fit showing good correspondence with the Köppen climate zones. Notable disagreements were observed over the Great Plains of the United States, over the central part of the Sahel, along the North China Plain and over the Temperate regions of Southern Australia.

Figure 8. Percentage fit for the DCA of the EFTs with the Köppen zones. High values indicate good fit of the two datasets. White areas represent pixels without vegetation cover.

3.5. Relation of the EFTs to the Global Land Use Systems Classification

The first two DCA axes explained 53.4% of the variance between the EFTs and the LUS classes within the four major climatic zones Arid, Cold, Temperate and Tropical (Table 6). The positive end of the first DCA axis grouped LUS classes and EFTs in the Cold climatic zone with high Cyclic and moderate Permanent Fractions and average Standing Biomass values (Figure 9). These EFTs are mapped with high to moderate positive values in Figure 10 (top).

Table 6. DCA of the EFTs with the Land Use System classes.

<table>
<thead>
<tr>
<th></th>
<th>Axis1</th>
<th>Axis2</th>
<th>Axis3</th>
<th>Axis4</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Eigenvalues</strong></td>
<td>0.586</td>
<td>0.306</td>
<td>0.077</td>
<td>0.043</td>
</tr>
<tr>
<td><strong>Cumulative % of explained variance</strong></td>
<td>35.1</td>
<td>53.4</td>
<td>58.0</td>
<td>60.6</td>
</tr>
</tbody>
</table>
Figure 9. DCA triplot of the EFTs classified in 15 clusters (icons with colors) and the Land Use System classes (crosses). Arrows represent the phenological and productivity variables averaged within the EFTs. Their position indicates the correlation with the DCA axes.

At the negative end of the first DCA axis, Land Use Systems in the Tropical EFTs are grouped whereas land use categories in the Temperate climatic zone are plotted towards the center of the biplot. These EFTs have higher Standing Biomass but low Permanent and Cyclic Fractions and are mapped with high to moderate negative values in Figure 10 (top). The positive end of the second DCA axis clearly separates land use categories in the Arid EFTs like managed and unmanaged sparse vegetation, grasslands, shrublands and wetlands. These ecosystems have very low Standing Biomass and low Permanent and Cyclic Fraction values and are mapped with high positive DCA values in Figure 10 (bottom). The EFTs discriminated managed and unmanaged land use categories especially well in the Cold climatic zones at the positive end of the first axis where the EFTs and LUS classes are plotted with larger distances between the categories. LUS classes are also well discriminated in the Arid EFTs at the positive end of the second DCA axis whereas the Tropical and Temperate EFTs
display a more aggregated thus, less interpretable pattern. The phenological values Season Length and Maximum Day reacted only moderately to the ordination space (shorter arrows) indicating small rate of change and thus are less important in distinguishing the LUS classes. Most EFTs presented a good (>50%) or very good (>70%) fit showing good correspondence with the LUS classes (Figure 11). Notable disagreements were observed over central Mexico (Sierra Madre del Sur), the arid agricultural areas of the Sahel, agricultural areas in Turkey, West of Iran, the Steppes in Kazakhstan, the agricultural areas in Northern India and along the eastern agricultural areas of China.

**Figure 10.** Spatial patterns of the first (top) and second (bottom) axis derived from DCA of the EFTs with the Land Use System classes. White areas represent pixels without vegetation cover.
Figure 11. Percentage fit for the DCA of the EFTs with the land Use System classes. High values indicate good fit of the two datasets. White areas represent pixels without vegetation cover.

4. Discussion

Recent decades have seen an increasing need for information on ecosystem functioning to evaluate the human impact on the environment [32]. A baseline characterization of ecosystem functioning is to be based on an up-to-date, effective and repeatable indicator system. Furthermore, regular updates are required as ecosystems rapidly change due to short time events such as droughts, floods, fires and human land use/intensification and longer term impacts due to climatic variability. Responsive and meaningful indicators are needed to assess the state and change of these ecosystems, including establishing causes and effects of degradation. With the aim to support monitoring of ecosystem functioning studies, we derived global EFTs from satellite measurements of seasonal vegetation dynamics. The EFTs showed a distinct spatial pattern in the PCA ordination but with overlap between the single categories conforming fuzzy boundaries of different ecosystems. Some overlap might also arise from the limits of remote sensing when observing seasonal vegetation dynamics over areas with low seasonality as e.g., over tropical areas or over areas with low vegetation cover. In fact, the tropical EFTs displayed strongly aggregated, thus less interpretable pattern when analyzing their correspondence with existing climate or with land use classifications. Nevertheless, the PCA ordination confirmed the potential of the EFTs in assessing the seasonal functioning of global ecosystems and the triplot enhanced the understanding of their phenological and productivity dynamics. Furthermore, it must be noted that droughts, floods, fires, land use intensification and longer term impacts due to climatic variability might change the functioning of the ecosystems. This is not reflected in the present study as we assessed EFTs based on the 30-year average functioning of the ecosystems. In order to assess changes in ecosystem dynamics, future studies should focus on yearly assessments and inter-annual variability of EFTs.
Climate is a main regional driver of ecosystem structure and functioning [52]. Therefore, to be valuable indicators of ecosystem functioning, the EFTs need to agree with climate gradients. The RDA indicated that the gradient of the EFTs and the phenology and productivity parameters along this gradient follow the global pattern of climatic vegetation growth constraints. The high percentage of variation explained by the first two axes in the gradient analysis showed that the EFTs captured two strong global environmental gradients. The first and stronger gradient positively related to water constraints and negatively related to radiation. The second, weaker gradient of the EFTs showed strongest correlation with the temperature growth constraint and moderate correlation to radiation constraints. For example, EFTs in high latitudes were shown to be controlled by temperature whereas EFTs in the arid and semi-arid areas were associated with water constraints of vegetation growth. We found a good correspondence of the global spatial pattern of these constraints within the EFTs and results of other authors on patterns of ecosystem-climate relationships (e.g., [29,32,44]). Nemani et al. [29] showed that water availability strongly limits vegetation growth over 40% of Earth’s vegetated surface, whereas temperature limits growth over 33% and radiation over 27%. Also Churkina and Running [44] showed that especially water but also temperature controls global NPP accumulation whereas radiation has considerably lower importance. Although we used a different variable for productivity than Churkina and Running [44], we obtained similar results showing that lowest productivity values were found in EFTs where water availability is the major climate constraint. We could further map a combined dependency of temperature and radiation using our extended set of variables. For instance, the yearly productivity expressed by the Cycling Fraction showed a strong association to temperature and radiation constraints for vegetation growth especially in the high latitudes. Furthermore, RDA suggested that the Maximum Day of the yearly vegetation signal is not related to water but rather to the constraining factor of temperature.

However, 50% of the variance in the global gradient of the EFTs could not be explained by the data indicating climatic constraints and variability and this should be attributed to other factors. On one hand, the climate and phenological data both have different time and spatial scales and associated noise/error that might play a role in the limited amount of explained variance. On the other hand, these results are consistent with Ivits et al. [19] showing a limited temperature and precipitation dependency of the remote sensing derived phenological and productivity parameters for Europe. That it is not only climate controlling the distribution of EFTs is further supported by the correspondence analysis between the Köppen climate zones and the EFTs, which was 47%. Especially the cold and arid zones scored high, showing that over these regions vegetation phenology and productivity agrees with bio-climatic classification systems, which are mostly dependent on climatic constraints [19]. In other biomes, however, the correspondence was very low (low DCA fit) indicating that the Earth Observation derived variables comprise ecosystem functional information that is not inherent in bio-climatic classifications. Indeed, variations in vegetation signals are not only driven by climatic cycles but among others, by anthropogenic management practices [7]. This suggests that the functional types of global ecosystems might be substantially influenced by anthropogenic activities as well and that the EFTs, to be valuable indicators of ecosystem functioning, should also reflect differences in land use.

Differences in land use vary locally depending on the nature of the human pressure [32]. For example, harvest will characteristically cause earlier end of vegetation growing season and thus shorter season length as compared to natural areas. We therefore analyzed the relation between the EFTs and
major global Land Use Systems testing how phenological and productivity parameters could enhance the understanding and characterization of these classifications. The influence of anthropogenic impacts on the EFTs was demonstrated by the 53.4% correspondence with the global Land Use Systems classes. Land use classes in the Cold and Arid zones were especially well discriminated by the EFTs whereas land use classes in the Tropical and Temperate zones, with more heterogeneous land use, were less interpretable. In a similar study, Ivits et al. [19] showed 85% correspondence of EFTs and European land covers using the 1km resolution SPOT VEGETATION data. In this global study based on the GIMMSs3g dataset, the results suggest that over more complex or intensively used areas the coarse spatial resolution of the remote sensing derived variables reach their limit in describing the spatial variability of vegetation cover. Hence, the response of EFTs to land use is dependent on the detail of spatial heterogeneity of the pixels potentially comprising all types of natural, semi-natural or intensively managed areas. Nevertheless, the present study demonstrated that EFTs do not only account for the climate dependence of ecosystem functioning but that they are also responsive to land use and thus supply indicators of the spatial patterns of ecosystem dynamics at larger spatial scales and offer an objective and repeatable method to characterise the functioning of ecosystems.

5. Conclusions

Using a time series of 30 years of Earth Observation derived phenological and productivity metrics, we stratified global ecosystems into Ecosystem Functional Types exhibiting similar patterns of seasonal phenology and productivity dynamics and similar responses to climate and land-use induced environmental conditions. The main findings of our analysis can be summarized as follows:

(1) Gradient analysis confirmed the potential of the Ecosystem Functional Types in assessing the phenological and productivity dynamics of global ecosystems.

(2) Redundancy Analysis showed that the gradient of the Ecosystem Functional Types indicates the global pattern of climatic vegetation growth constraints.

(3) Correspondence Analysis confirmed that Ecosystem Functional Types relate to classification of global climatic zones.

(4) Correspondence Analysis also indicated that it is not only climate controlling the distribution of Ecosystem Functional Types but that they also correspond to classifications of Land Use Systems, showing that the functional types of global ecosystems might be substantially influenced by anthropogenic activities as well.

(5) By incorporating the spatial information of Earth Observation derived metrics into a new ecosystem functional map, we demonstrated that Ecosystem Functional Types comprise functional information that is not inherent in bio-climatic classifications and have potential in the monitoring of human influence on ecosystem functioning and in supporting ecosystem degradation studies.

Acknowledgement

The authors would like to thank the NASA Global Inventory Modelling and Mapping Studies (GIMMS) group for producing and sharing the GIMMS3g NDVI dataset. Thanks also to Ramakrishna
Nemani for providing the potential climatic plant growth constraints data used in the analyses. This research was partly funded by the Danish Council for Independent Research (DFF) Sapere Aude programme under the project entitled “Earth Observation based Vegetation productivity and Land Degradation Trends in Global Drylands”.

Conflict of Interest

The authors declare no conflict of interest.

References


© 2013 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license (http://creativecommons.org/licenses/by/3.0/).