Land Productivity Dynamics Indicator: LPDynR Package

Xavier Rotllan-Puig1,2,✉, Eva Ivits3, Pier Lorenzo Marasco1, Federico Gianoli1, and Michael Cherlet1

14/12/2020

1 Joint Research Centre – European Commission. Directorate D – Sustainable Resources. Unit D6 – Knowledge for Sustainable Development & Food Security Unit. Via Enrico Fermi 2749. I-21027 Ispra (VA), ITALY  
2 ASTER-Projects. Barri Reboll, 9, 1r. 08694 Guardiola de Berguedà (Barcelona), SPAIN  
3 Geospatial Information Services group, European Environment Agency, Copenhagen, DENMARK

✉ Correspondence: [Xavier Rotllan-Puig <[xavier.rotllan.puig@aster-projects.cat](mailto:xavier.rotllan.puig@aster-projects.cat)>](mailto:xavier.rotllan.puig@aster-projects.cat)

**Abstract**

As part of the UN Sustainable Development Goal 15 (Life on Land), the indicator 15.3.1 is adopted to measure the Land Degradation Neutrality (stable —or increasing— state regarding the amount and quality of land resources required to support ecosystem functions and services and enhance food security during a certain period of time). It is a binary indicator (i.e. degraded/not degraded), expressed as the proportion of land that is degraded over total land area, and is based on three sub-indicators: (1) Trends in Land Cover, (2) Land Productivity and (3) Carbon Stocks.

The Land Productivity sub-indicator (LP) refers to the total above-ground Net Primary Production and reflects changes in health and productive capacity of the land. Its declining trends can be usually understood as land degradation. LP is calculated using the Land Productivity Dynamics (LPD) approach, first developed by Ivits and Cherlet [“Land-Productivity Dynamics Towards Integrated Assessment of Land Degradation at Global Scales.” Technical Report EUR 26052. Joint Research Centre of the European Commission (2013)].

The LPD is the methodological basis of the R-based tool *LPDynR* presented in this document. It uses vegetation-related indices (phenology and productivity) derived from time series of remote sensed imagery to estimate ecosystem dynamics and change. The final result of the LPD indicator is a categorical map with 5 classes of land productivity dynamics, ranging from declining to increasing productivity. As an example of *LPDynR* functionalities, we present a case study for Europe.

# Introduction

The United Nations General Assembly designed in 2015 a collection of 17 global goals, so called Sustainable Development Goals (SDGs; UN 2015), with the general aim of “achieving a better and more sustainable future for all”, and which were intended to be accomplished by the year 2030. Each SDG is subdivided into a list of targets which, in turn, go together with indicators to be able to measure their success. Such indicators have to be credible, based on standardized methodologies and, often, spatially explicit (Dubovyk 2017).

The SDG-15, entitled Life on Land, has among its targets the 15.3, which expects “to combat desertification, restore degraded land and soil, including land affected by desertification, drought and floods, and strive to achieve a land degradation-neutral world”. In this context, Land Degradation Neutrality (LDN) is defined as the stable (or increasing) state regarding the amount and quality of land resources required to support ecosystem functions and services and enhance food security during a certain period of time (UNCCD 2015).

The indicator 15.3.1 is adopted to measure the LDN and is expressed as the proportion (%) of land that is degraded over total land area. It is a binary indicator (i.e. degraded/not degraded) based on three sub-indicators calculated separately: (1) Trends in Land Cover, (2) Land Productivity and (3) Carbon Stocks (Sims et al. 2017, 2020). While the first two can capture relatively fast changes, carbon stocks reflect slower changes which suggest a longer term trajectory (Orr et al. 2017). Following a “one-out-all-out” process, the indicator identifies an area as degraded if one of the sub-indicators shows degradation. The three sub-indicators must be comparable among territories and based on standardized sources and methods. The data can be collected through existing sources, such as maps, reports or databases, but also can be derived from Earth Observation (EO) imagery using remote sensing tools.

The Land Productivity sub-indicator (LP), addressed in this document, refers to the total above-ground net primary production (NPP), which can be defined as the total energy fixed by plants minus their respiration. Such energy is transformed into biomass which, in turn, allows ecosystems to develop their functions and deliver essential services. Therefore, LP reflects changes in health and productive capacity of the land and its declining trends can be usually understood as land degradation (Cherlet et al. 2018; Prince 2009; Yengoh et al. 2015). The LP sub-indicator is calculated using the Land Productivity Dynamics (LPD) approach, first developed by Ivits and Cherlet (2013), which is the methodological basis of the *LPDynR* tool presented in this document.

# Land Productivity Dynamics and *LPDynR*

The Land Productivity Dynamics (LPD) approach is based fundamentally on the use of time series of vegetation-related indices derived from remote sensed imagery, such as the normalized difference vegetation index (NDVI) or the plant phenology index (PPI). NDVI, for example, can be used as a proxy for land productivity, as many studies at global and local scales have identified a strong relationship between NDVI and NPP (Ivits and Cherlet 2013; Prince 2009; Yengoh et al. 2015, and references therein). The LPD approach often uses phenological and productivity-related variables derived from time series of NDVI, given that these can provide additional information on several aspects of vegetation/land cover functional composition in relation to ecosystem dynamics and change (Ivits, Cherlet, Mehl, et al. 2013). These dynamics of the ecosystems, which eventually might drive to land degradation, can be caused by human activities and/or biophysical processes, as well as other processes not tied to them, such as climate change (Yengoh et al. 2015). While the most commonly used phenological parameters are the beginning and the end date of the vegetation growing season, together with the season length in number of days, the ones related to land productivity are those which approximates the measures to NPP and growing season production.

The final result of the LPD indicator is a categorical map with 5 classes of land productivity dynamics, ranging from declining to increasing productivity. It is the result of a combined assessment of two sources of information, as seen in Figure 1. On the one hand, the first layer is the Long Term Change Map. In general terms, it shows the tendency of change of land productivity (positive or negative) and the effect that this tendency might have had on a particular original point after a certain period of time. On the other hand, the second layer is the Current Status Map, which provides information on the current levels of land productivity in relation to its potential. It compares the local productivity with the range of productivity across similar areas in terms of land cover or bioclimatic traits (Sims et al. 2017). Further explanations for both branches will be given in their own sections below.

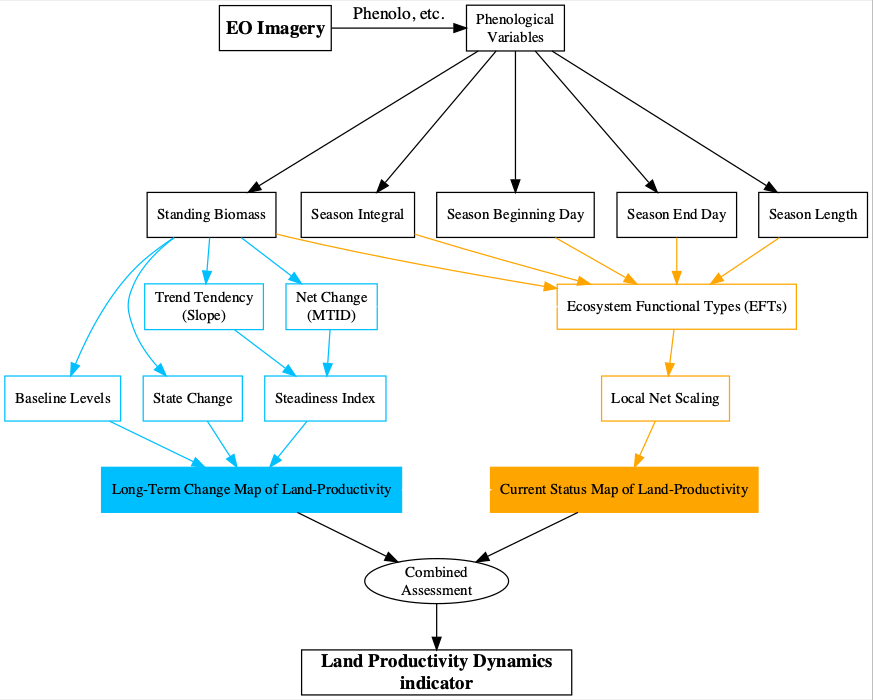


Figure 1: Flowchart of the process to calculate the Land Productivity Dynamics indicator and followed by *LPDynR*

Following the LPD approach, *LPDynR* is an R-based tool (i.e. an R package) which allows the user to produce the final Land Productivity Dynamics Map using as inputs a set of time series of phenological and/or productivity variables (multi-band GeoTIFF rasters). By means of the different functions included in the package, it produces intermediate layers (e.g. Steadiness Index, Ecosystem Functional Types, etc.) which are used to calculate both the Long Term Change Map and the Current Status Map. In addition, several parameters can be set along the process in order to fit them with the preferences of the user. The functions included in the package have no limitations regarding the number of years included in the time series, the variables to use or the spatial extent and resolution. The source code of the latest version of *LPDynR* can be found at <https://github.com/xavi-rp/LPDynR>.

# Data set preparation

A case study is presented in this document in order to illustrate the methodology implemented in the *LPDynR* package to calculate the LPD indicator. In this case, it is used a data set of 3 phenological and productivity-related variables, at European level and 0.5km of spatial resolution, produced and freely distributed by the European Environmental Agenct - Europeam Commission (<https://www.eea.europa.eu/data-and-maps/>). They are all derived from time series (2000-2016) of MODIS imagery and its derived product Plant Phenology Index (PPI; Jin and Eklundh 2014). PPI is linearly related to the canopy green leaf area index (LAI) and has a temporal pattern very similar to the one shown by the gross primary productivity (GPP) estimated by flux towers at ground reference stations. The three variables are produced using the software TIMESAT (Jönsson and Eklundh 2004). More information about them can be found in their own webpage:

* Above ground vegetation productivity (from now on, CF): <https://www.eea.europa.eu/data-and-maps/data/annual-above-ground-vegetation-productivity>
* Start of vegetation growing season (from now on, SBD): <https://www.eea.europa.eu/data-and-maps/data/annual-start-of-vegetation-growing>)
* Vegetation growing season length (from now on, SL): <https://www.eea.europa.eu/data-and-maps/data/annual-above-ground-vegetation-season>)

In the *LPDynR* v.1.0.0 the main function uses multi-band “GeoTIFF” rasters to start the process, one per phenological/productivity variable. Each band of each raster contains one of the years of the time series. Therefore, the initial data set needs to be prepared accordingly.

In addition, *LPDynR* comes with a small data set, which can be used to run tests, as well as some examples in the form of “vignettes” attached to the package. Once the package is installed, a basic example can be seen loading the following lines of R code:

# To install the latest version of LPDynR  
library(devtools)  
install\_github("xavi-rp/LPDynR")  
  
# Launching an example  
library(LPDynR)  
vignette(topic = "LPD\_simple\_example", package = "LPDynR")

# Long Term Change Map of Land Productivity

As seen in Figure 1 and explained above, the Land Productivity Dynamics indicator is produced based in two different main layers, being the first one the Long Term Change Map (also called “tendency map”). In turn, this tendency incorporates information both on the progression of the general process of land dynamics (positive or negative) and on the original level of productivity of the ecosystem, as well as whether it has changed its state or not in the period of study (Ivits and Cherlet 2013). The multi-source information used for the Long Term Change Map derivation is necessary because, for instance, even though an ecosystem presents a long term negative dynamics, it might have not been strong enough to decrease its level of productivity to change its original state. The way in which the three sources of information are calculated for the Long Term Change Map using a land productivity variable is described in the following subsections.

## Steadiness Index

The first of the three metrics which integrates the Long Term Change Map represents the long term tendency of degradation of the natural systems on study, either positive or negative. This metrics is the “Steadiness Index” which, in turn, is based on the combination of two other metrics calculated per pixel: (1) the slope derived from a linear regression of the different years of the time series and (2) the net change of the productivity level on the same period.

The use of a linear regression would imply to respect some strict statistical assumptions for confidence intervals and significance tests to be representative. This is why the Steadiness Index only keeps classes of tendency and no more tests are run for assessing its significance. See Ivits and Cherlet (2013) for further explanations on this point. Therefore, only the sign (positive or negative) of the slope of the trend is kept as the value of each pixel’s tendency of ecosystem dynamics. In addition, the net change of the productivity variable is calculated for the same time window and per pixel using the Multi-Temporal Image Differencing method (MTID; Guo et al. 2008). And, afterwards, it is transformed also into positive or negative net change. Finally, both metrics (slope of the linear function and net change category) are combined to get four “steadiness” categories as seen in Table 1.

Table 1: Description of the four Steadiness Index classes and how they are derived based on the combination of the signs of both the slope of the linear function and the net change

|  |  |  |  |
| --- | --- | --- | --- |
| Steadiness Class | Slope | Net Change | Description |
| Steadiness1 | - | - | Strong negative ecosyst. dynamics (possibility changing equil.) |
| Steadiness2 | - | + | Moderate negative ecosyst. dynamics (likely remain current equil.) |
| Steadiness3 | + | - | Moderate positive ecosyst. dynamics (likely remain current equil.) |
| Steadiness4 | + | + | Strong positive ecosyst. dynamics (possibility changing equil.) |

While Figure 2 represents a 4-class map of the Steadiness Index for the case study, the following lines of code show how to run the function *steadiness()* of the package to calculate the Steadiness Index.

?steadiness  
SteadInd <- steadiness(obj2process = cf, # 'cf' is the productivity variable with the time series   
 cores2use = 3, # for parallel processing  
 filename = "SteadInd.tif")

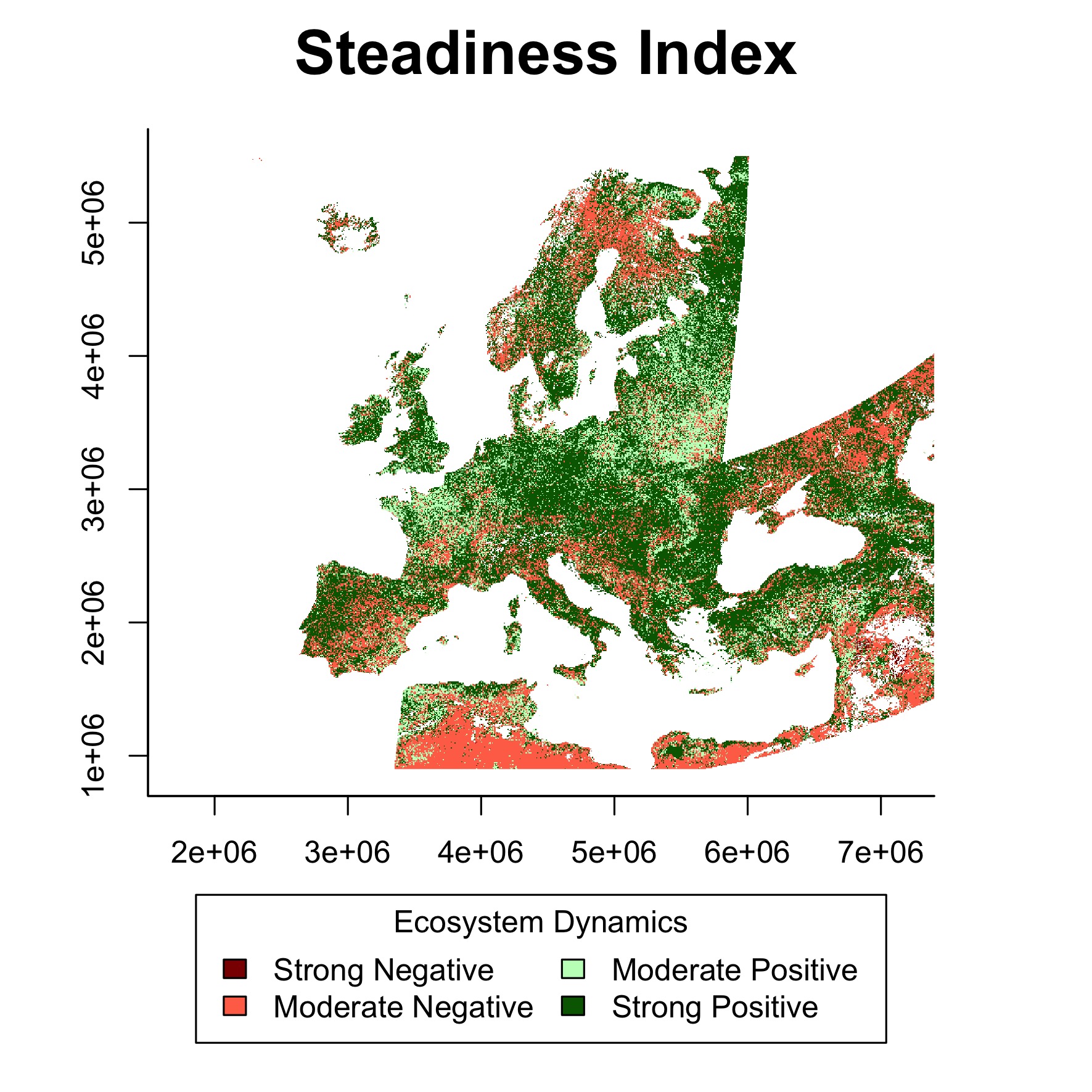


Figure 2: Representation of the Steadiness Index for the case study based on the ‘Above ground vegetation productivity’ variable

## Baseline levels of the productivity variable

The second source of information for the derivation of the Long Term Change Map is the baseline levels of the productivity variable in study, in the example case, the “Above ground vegetation productivity” variable.

For the calculation of the baseline levels of land productivity at the beginning of the time series on study, *LPDynR* categorizes the results into three classes: low, medium and high. To do that, the function *baseline\_lev()* averages the first *n* years of the time series in order to avoid extreme events such as abnormal droughts in wet areas, etc. This number of years can be set by passing the argument *yearsBaseline* to the function. Its default is 3, given that averaging more years would move the value closer to the mean of the time series, which is not desirable.

After the average is calculated, *baseline\_lev()* first classifies pixels into 10 classes instead of the final three (i.e. low, medium and high) using 10-quantiles. The reason for this intermediate step is that, if directly opted for three classes, the number of pixels per category would have been classified homogeneously (i.e. 33.3% of pixels/class), and this is in contrast with what is stated by the United Nations Development Programme (<https://www.undp.org>). UNPD declares that 40% of the World’s land resources are drylands (Middleton et al. 2011) and, therefore, 40% of pixels at global level must be classified as “low level” of productivity. Consequently, as default, *LPDynR* classifies the first four groups of pixels as “low”, whereas the five consecutive groups as “medium” and the rest 10% of pixels with the largest baseline levels, as “high”. Both the proportion of pixels classified as low level and high level of land productivity can be set by passing to *baseline\_lev()* the arguments *drylandProp* and *highprodProp*, respectively. The function classifies the rest of the pixels ((100 - (*drylandProp* + *highprodProp*)) as medium level. The assumption of classifying 40% of pixels as low productive is valid at global level, however, the proportion of drylands/low level of productivity should be modified for local studies. For example, at the European level, drylands cover 20% of total land (FAO 2019).

In the following lines of code it can be observed how to run *baseline\_lev()* to categorize the baseline levels of productivity, while adjusting the parameters to the European proportion of drylands. The result is a final 3-class map showing the estimation of the levels of productivity at the beginning of the time series (Figure 3).

?baseline\_lev  
  
Baseline\_Level <- baseline\_lev(obj2process = cf,   
 yearsBaseline = 3,   
 drylandProp = 0.2, # 20% as dryland  
 highprodProp = 0.1, # 10% highly productive land   
 cores2use = 3,   
 filename = "Baseline\_Level.tif")

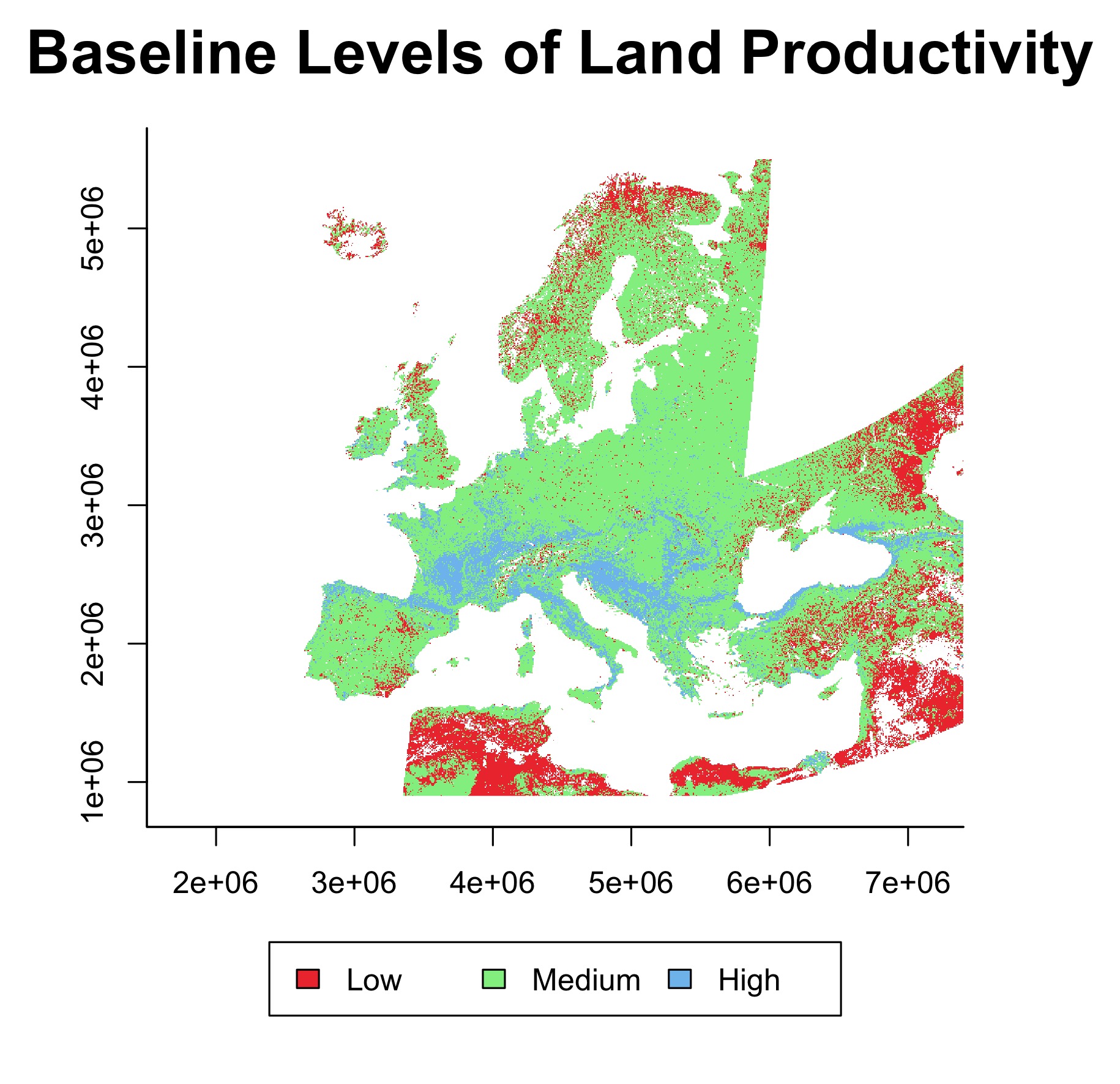


Figure 3: Representation of the baseline levels of land productivity for the case study

## State Change of the productivity variable

The third layer used for the land productivity Long Term Change Map is the change of the state of the productivity variable during the time window of the study. This point is necessary for land degradation assessments as it reports whether productivity thresholds have been passed or not. The state change can be a consequence of either the natural resilience thresholds have been surpassed or new land use/practices have been introduced by humans (Ivits and Cherlet 2013).

To calculate this state change per pixel, *LPDynR* uses both the state level at the beginning of the time series, as calculated in the previous subsection, and the state level at the end of the period. This final state is calculated in the same way than the base line level, i.e. (1) averaging the last 3 years and (2) classifying into 10 categories using 10-quantiles. The reason for not using the final 3-class classification is because it would be difficult to appreciate if the change of one state to another was due to a big or a small change. Instead, using the 10-class classification, one can see for instance if a pixel has moved from class 5 to 4 (small change) or from class 9 to 4 (big change).

Once the class change per pixel has been calculated, either with positive or negative results, the map is categorized into 3 final classes: (1) no change, (2) changed 1 class or (3) changed 2 or more classes. See Figure 4 for a map of the state change in the case study. The following code shows how to run the function *state\_change()* to perform the final classification.

?state\_change  
  
State\_Change <- state\_change(obj2process = cf,   
 yearsBaseline = 3,   
 cores2use = 3,  
 filename = "State\_Change.tif")

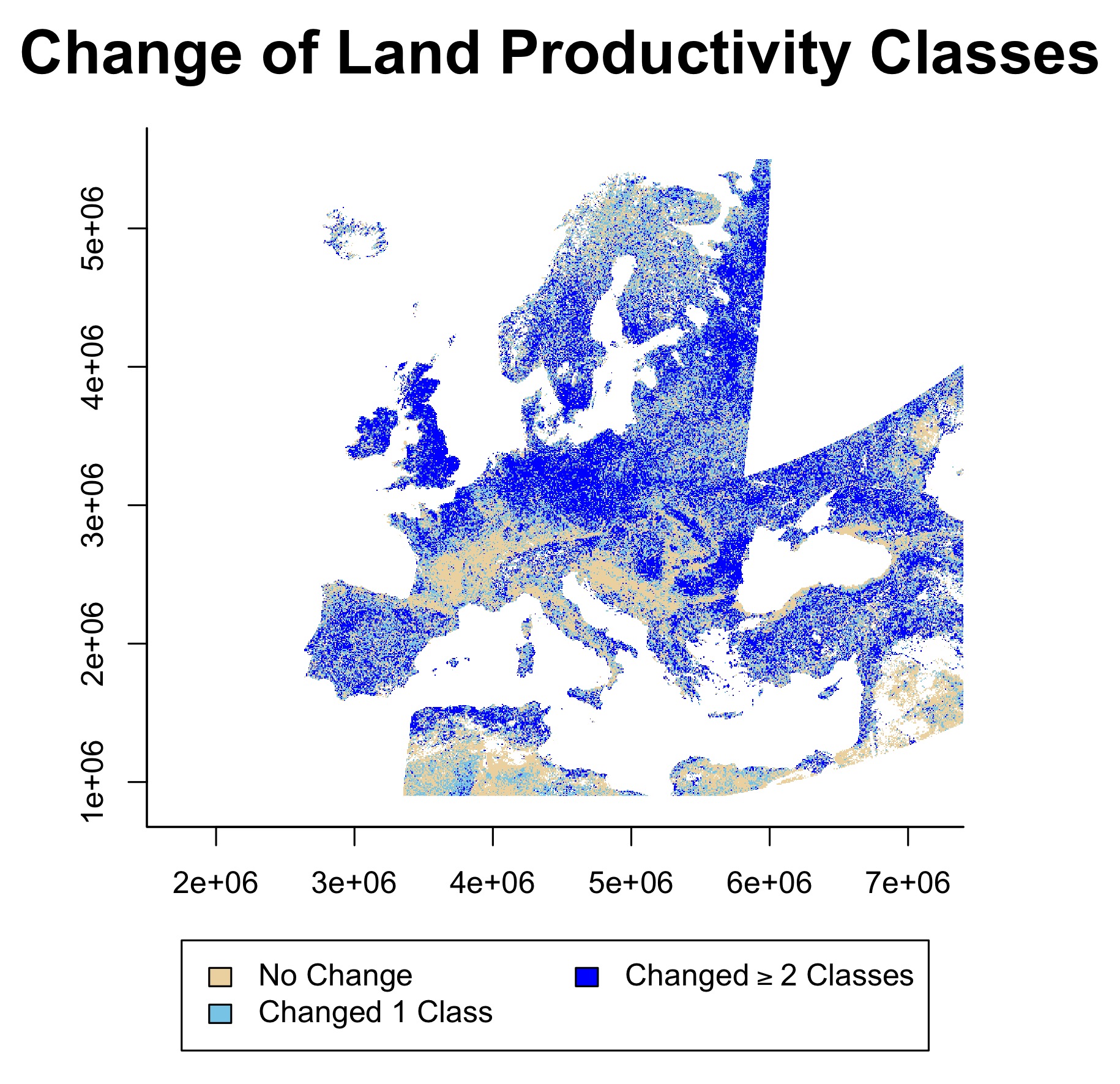


Figure 4: Representation of the state change map of the productivity variable (baseline levels minus final levels), after being reclassified into 3 categories (1: no change; 2: changed 1 class; 3: changed 2 or more classes) for the case study

## Land Productivity Long Term Change Map

The land productivity Long Term Change Map is one of the two pillars of the LPD indicator (Figure 1) calculated with *LPDynR*. This map is developed, in turn, by the combination of the Steadiness Index, the productivity levels at the beginning of the time series and the change of the state of productivity between the beginning and the end of the time series, all calculated with the productivity variable in use.

The function *LongTermChange()* (see an example of running it below) performs the combination of the three qualitative metrics mentioned before into the Long Term Change Map, resulting in 22 new categories as shown in Table 2. The resulting map for the case study is presented in Figure 5.

?LongTermChange  
  
Long\_Term\_Change\_Map <- LongTermChange(SteadinessIndex = SteadInd,   
 BaselineLevels = Baseline\_Level,  
 StateChange = State\_Change,   
 filename = "Long\_Term\_Change\_Map.tif")

Table 2: Lookup table for the land productivity Long Term Change Map (Steadiness Index + BaseLine Levels + State Change)

|  |  |  |  |
| --- | --- | --- | --- |
| **Class Change** | No Change | Changed 1 class | Changed ≥ 2 classes |
| **Stead.Ind./BaseL.** |  |  |  |
| **St1 low** | 1 | 2 | 3 |
| **St1 med.** | 4 | 5 | 6 |
| **St1 high** | 7 | 8 | 9 |
| **St2 low** | 10 | 10 | 10 |
| **St2 med.** | 11 | 11 | 11 |
| **St2 high** | 12 | 12 | 12 |
| **St3 low** | 13 | 13 | 13 |
| **St3 med.** | 14 | 14 | 14 |
| **St3 high** | 15 | 15 | 15 |
| **St4 low** | 16 | 17 | 18 |
| **St4 med.** | 19 | 20 | 21 |
| **St4 high** | 22 | 22 | 22 |

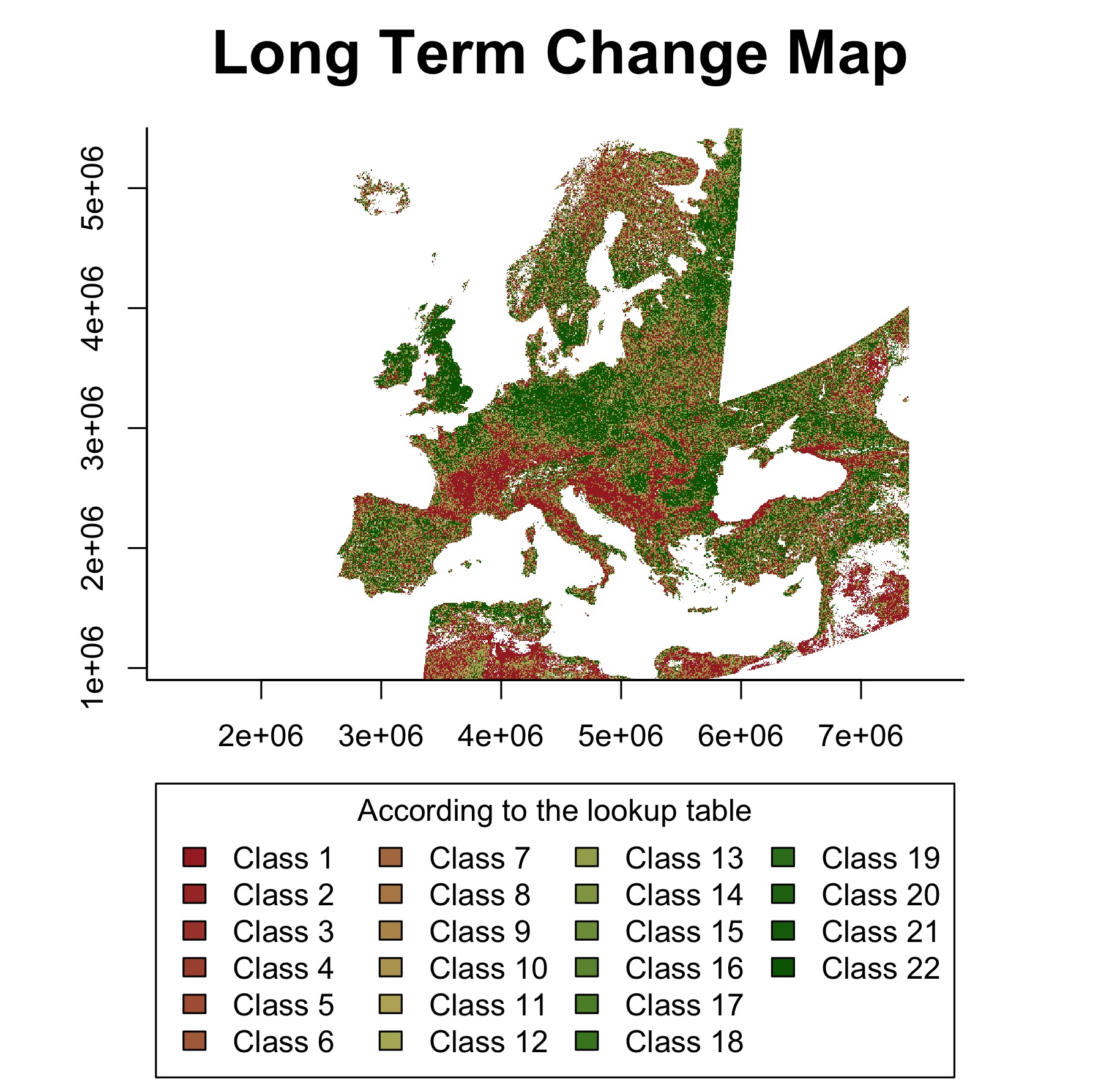


Figure 5: Land productivity Long Term Change Map for the case study based on the combination of the Steadiness Index, the baseline levels and change of states of land productivity. **I know that the colour palette is not ‘colour-blind friendly’, but it’s hard to find such a palette with so many colours… any suggestion very welcome!!**

At this point, the user might want to finalise the LPD calculation avoiding the second part of the methodology proposed by Ivits and Cherlet (2013), which is the Current Status Map of Land Productivity. In this case, the function *LPD\_CombAssess* (see further explanations in its own subsection below) can be called to reclassify the 22-class Long Term Change Map into the final 5 classes of LPD. In the following code it can be seen an example of how to run the function to perform the reclassification.

?LPD\_CombAssess  
  
LPD\_finalMap <- LPD\_CombAssess(LandProd\_change = "Long\_Term\_Change\_Map",   
 LandProd\_current = NULL,  
 filename = "LPD\_finalMap.tif")  
  
plot(LPD\_finalMap)

# Current Status Map of Land Productivity

The Land Productivity Dynamics indicator is composed by two base layers, as shown in Figure 1. After the long term productivity dynamics described previously (i.e. Long Term Change Map), the second source of information needed is the current level of land productivity. For this purpose, a Local Net Scaling approach is implemented (Prince 2009). Such approach estimates, by means of Earth Observation imagery and remote sensing tools, the level of land productivity of each pixel relative to its neighbours with similar characteristics. In other words, it calculates the potential level of productivity of each pixel within a homogeneous land unit at the time under study. The Current Status Map may help, for instance, to identify areas which, although having a positive trend of productivity over time, their levels of current productivity are low relative to the pixels in the same homogeneous land unit and, thus, they might be still suffering land degradation (Sims et al. 2017) **here would perfectly fit the examples that Eva mentioned in her emails, which are included in the GPGv2 (still unpublished) but not in GPGv1**. A first step for the calculation of the Current Status Map, therefore, must be the derivation of the homogeneous land units across the area of study.

## Ecosystem Functional Types (EFTs)

The methodology implemented in *LPDynR* to derive homogeneous land units, or Ecosystem Functional Types (EFTs), is adapted from Ivits, Cherlet, Horion et al. (2013). It is basically a clustering process which uses, in this case, phenological and productivity variables to create the groups. Among the different unsupervised clustering techniques available for data grouping, K-means has been chosen. K-means is widely used in data science mainly due to its relative simplicity.

Originally, the unsupervised classification was performed after a three-steps preprocessing of the variables: (1) removing highly correlated variables; (2) a first Principal Component Analysis to know the optimal number of PCs and their associated variables showing the highest loadings; and (3) a final PCA to clearly associate each PC with one variable. However, some recent tests have shown that the final LPD indicator does not differ significantly when it is derived using the raw phenological/productivity variables. Therefore, although the two-PCAs step is also implemented in *LPDynR*, only removing those highly correlated variables (e.g. |r| > 0.7) is recommended before to run the k-means clustering.

In order to check for multicollinearity among the variables, the function *rm\_multicol()* first calculates their averages among the years of the time series. Then, the process internally runs the function *removeCollinearity()* from the package *virtualspecies*. This function allows the user to set up the minimum Pearson’s correlation absolute value, which is established to be 0.7 as default. It is also allowed to use a subset of random points of the data set to calculate the correlation in case the rasters have a large number of pixels and the user wants to speed up the process. The default number of randomly selected points is 10% of total pixels in the raster. Finally, one of the variables of each group of correlation is randomly selected. A dendrogram to visualize the groups of intercorrelated variables can be plotted. Figure 6 shows a dendrogram for the present case study, which has been run with three variables and no intercorrelation has been found among them at the cutoff 0.7. In the following piece of code it can be seen an example of how to run *rm\_multicol()* with all the mentioned parameters.

?rm\_multicol  
  
variables\_noCor <- rm\_multicol(dir2process = variables\_dir,   
 multicol\_cutoff = 0.7,   
 cores2use = 3,  
 filename = "variables\_noCor.tif",  
 # The following arguments are passed to virtualspecies:removeCollinearity()  
 sample.points <- TRUE # using 'nb.points' (or 10% of pixels) to calculate multicollinearity  
 nb.points <- 1000000   
 plot <- FALSE # if TRUE it writes out a dendrogram  
 )

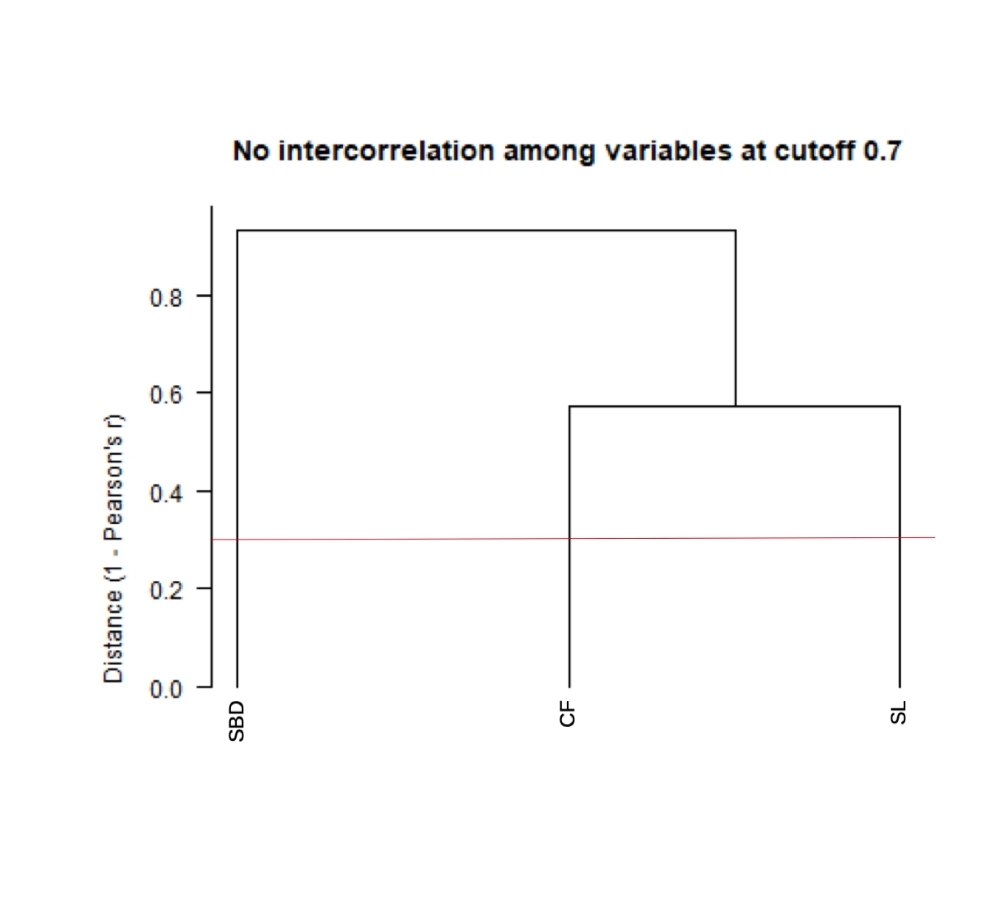


Figure 6: Dendrogram of groups of correlated variables (none in this case). Multicollinearity cutoff set to r > |0.7|. SBD: Start of vegetation growing season; SL: Vegetation growing season length; CF: Above ground vegetation productivity

In case the user would like to run the two-PCAs step, both the first “screening PCA”, which is done over the uncorrelated variables, and the “final PCA” are subsequently performed with the same function *PCAs4clust()*. In order to know the optimal number of variables to be used in the “final PCA”, a threshold of cumulative variance of the PCs is implemented. This threshold is established to be 0.9 as default. The following code shows an example of how to run this process.

?PCAs4clust  
  
pca\_final\_brick <- PCAs4clust(obj2process = variables\_noCor,   
 cumul\_var\_threshold = 0.9,  
 filename = "pca\_final\_brick.tif")

Finally, the clustering algorithm can be run over either the selected PCs or the uncorrelated raw variables using the function *EFT\_clust()*. This function uses *kmeans()* from the package *stats*. K-means is an iterative unsupervised method which has as one of the main limitations the fact that it is not able to optimize the number of clusters by itself. Instead, the optimal number of clusters needs to be calculated externally. In this case, it can be determined using the “scree-plot method”. Such method is implemented in *LPDynR* with the function *clust\_optim()* and it is based on running several K-means clustering with different number of clusters each, in order to assess how the quality of the models change with the number of clusters. Then, a plot is produced with the results and an “elbow” indicates where the quality of the model no longer improves substantially as the number of clusters (model complexity) increases. See Figure 7 for the results of the example presented in this document and the following code as an example of how to call the function. Notice that the clustering has been run with nine different number of clusters to give a good amount of points to plot the curve. And also that the maximum number of iterations is set up to 10. Although it is unlikely that the process gets any convergence with such a low number of iterations, the results are already valid for the purpose.

?clust\_optim  
  
jpeg("OptimalNumClusters.jpg")  
clust\_optim(obj2clust = pca\_final\_brick,   
 num\_clstrs = seq(5, 50, 5)  
 )  
dev.off()

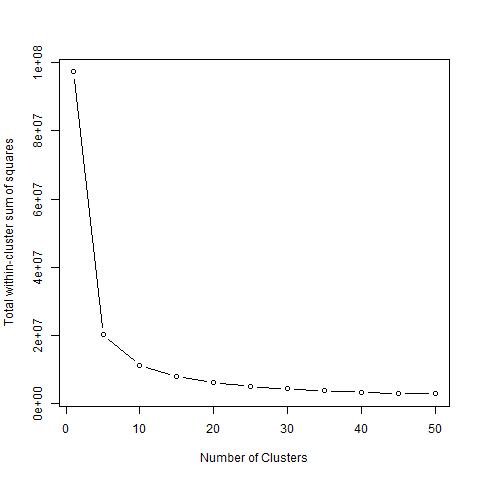


Figure 7: “Scree plot” method used to calculate the optimal number of clusters. The “elbow” indicates where the quality of the model no longer improves substantially as the number of clusters (model complexity) increases

The “scree plot” method undoubtedly has some level of subjectivity, as the user decides where the curve flattens enough. Alternatively, to remove such subjectivity, several numerical methods exist to calculate the optimal number of clusters, although they take also some statistical assumptions. These methods could be explored in the future if a higher level of accuracy is believed as necessary or if the whole process wants to be done without user’s intervention. In addition, other hierarchical clustering methods could be explored in order to avoid calculating the optimal number of clusters beforehand, although previous tests run with ISODATA have been shown to be highly resource demanding, especially in terms of computing time.

Once the optimal number of clusters is estimated, the final clustering is run with the function *EFT\_clust()*. It is important to notice that when setting *nstart*, the larger, the more accurate result. This is because the function uses different sets of starting random centroids and runs the clustering *nstart* times, and the best result is chosen. Therefore, a larger *nstart* increases the chances of having a better cluster classification. *EFT\_clust()* provides, together with a RasterLayer object with the clusters, a clustering performance evaluator. This value is calculated by the ratio of *betweenss* (i.e. between-cluster sum of squares) and *totss* (i.e. total sum of squares), in percentage, and it is expected to be as high as possible.

In addition, *k-means()* can use different algorithms to perform the clustering (e.g. “MacQueen”, “Hartigan-Wong”, etc.). As stated in the function documentation (*?k-means*), “Hartigan-Wong” usually gives better results, although it is recommended to try several starts (*nstart > 1*). However, when using “Hartigan-Wong” with a (too) large number of clusters, and a lot of values of the variables are very similar, *k-means()* is not able to converge in an acceptable amount of time (even increasing the number of iterations). In these cases, the user has to be careful because *k-means()* only gives a warning, so the final clustering is based on a non converged process. Diminishing the number of clusters or rounding variables’ values might be good strategies to help *k-means()* to converge.

Finally, as previous tests of K-means with up to 100 iterations were not converging, the maximum number of iterations is set to 500 as default. Within this limit and rounding variables, for almost all the tests performed, the process did achieve convergence.

For the running example, the EFTs resulted from the whole process can be seen in Figure 8. In addition, the following code shows an example of the function call using the *BrickLayer* object produced in the previous step with *rm\_multicol()*, and setting the optimal number of clusters estimated with *clust\_optim()*, together with a *nstart* value higher than 1 and the algorithm to be used.

?EFT\_clust  
  
EFTs <- EFT\_clust(obj2clust = variables\_noCor,   
 n\_clust = 20,   
 nstart = 5,   
 algorithm = "Hartigan-Wong",  
 filename = "EFTs.tif")  
  
clust\_eval <- EFTs[[2]] # Evaluation of clustering performance  
EFTs <- EFTs[[1]] # RasterLayer object with the clusters (i.e. EFTs)

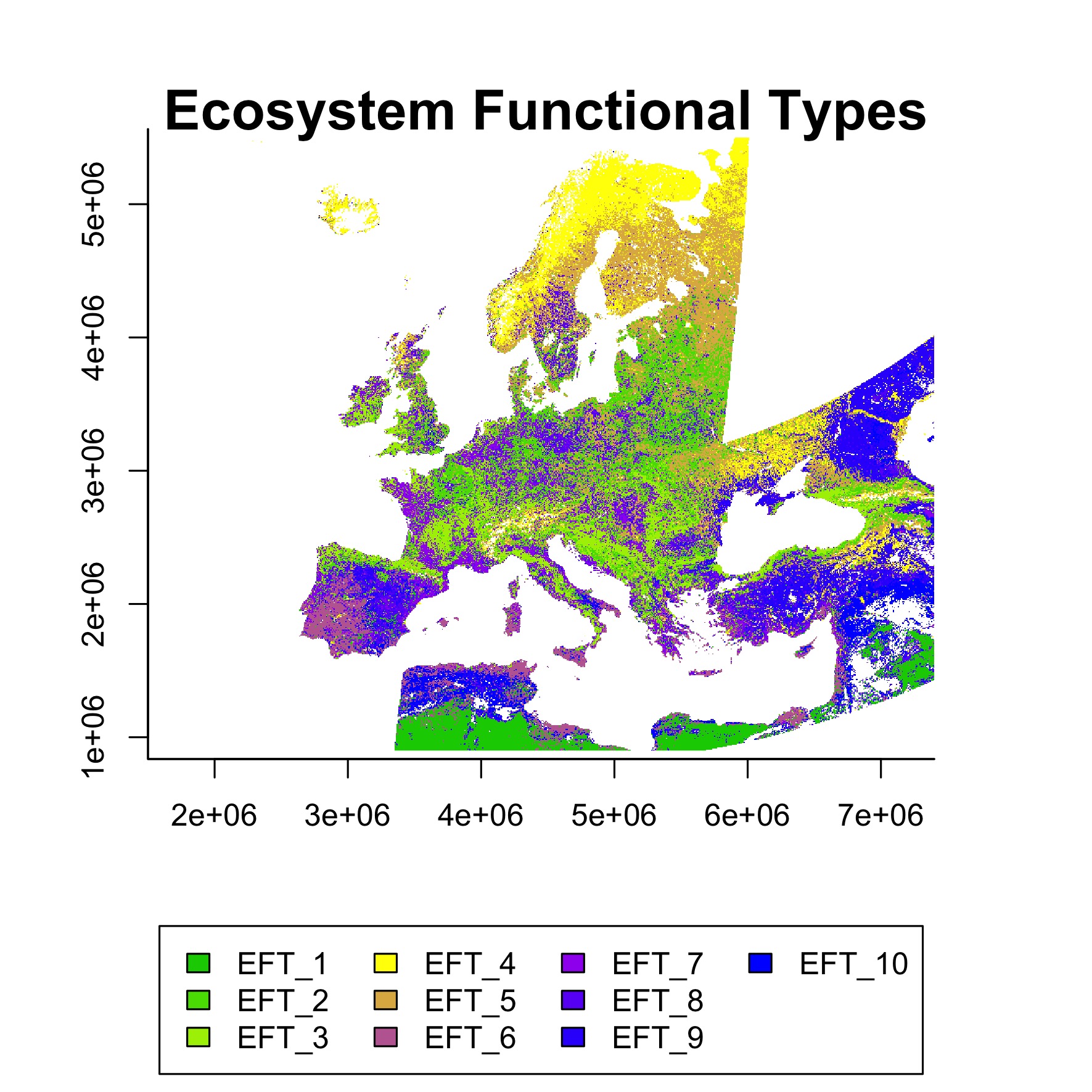


Figure 8: Ecosystem Functional Types (EFTs) derived from phenological and productivity variables using the K-means clustering method for the case study

## Local Net Production Scaling

The Local Net Primary Production Scaling (from now on, Local Net Scaling or LNS) method (Prince 2009) is based on the use of multi-temporal satellite data to calculate the difference between the potential and actual NPP for each pixel in homogeneous land areas. It will be considered potential productivity to that productivity which would have been without the influence of human factors, and it is estimated as the maximum value of productivity within each EFT (Prince 2009, and references therein). The current land production related to the local potential reflects the current level of productivity efficiency and, therefore, it is useful for the delineation of a productivity status map (Ivits and Cherlet 2013).

The cyclic fraction of vegetation productivity, or the summed NDVI over the growing season, is widely used as a proxy for the estimation of the current land productivity (Fensholt 2013), as it incorporates both natural and anthropogenic factors which define the inter-annual variability of land production. Therefore, it represents that part of the standing biomass which is potentially appropriated to be used by humans and the environment (Ivits and Cherlet 2013).

The function *LNScaling()* is implemented in *LPDynR* to calculate the LNS. In the case study, the above ground vegetation productivity varible (i.e. CF) for the period between 2011 and 2016 has been passed to the function, which runs the LNS using the average of the last 5 years. Together with the productivity variable, the EFTs calculated previously are also passed to the function, and the potential productivity within each EFT is calculated. However, instead of the maximum value within each cluster, its 90-percentile is established as the final potential, given that values higher than this threshold can be considered as outliers. Finally, the LNS per pixel is calculated as the proportion of its annual production (i.e. the average of 5 years cyclic fraction) over the potential production within its EFT (i.e. the 90-percentile). The result for the case study is represented in Figure 9, and the following lines of code show an example on how to call the function.

?LNScaling  
  
LNScal <- LNScaling(EFTs = EFTs, # RasterLayer with the EFTs  
 ProdVar = cf, # Productivity variable  
 cores2use = 3,  
 filename = "LNScal.tif")

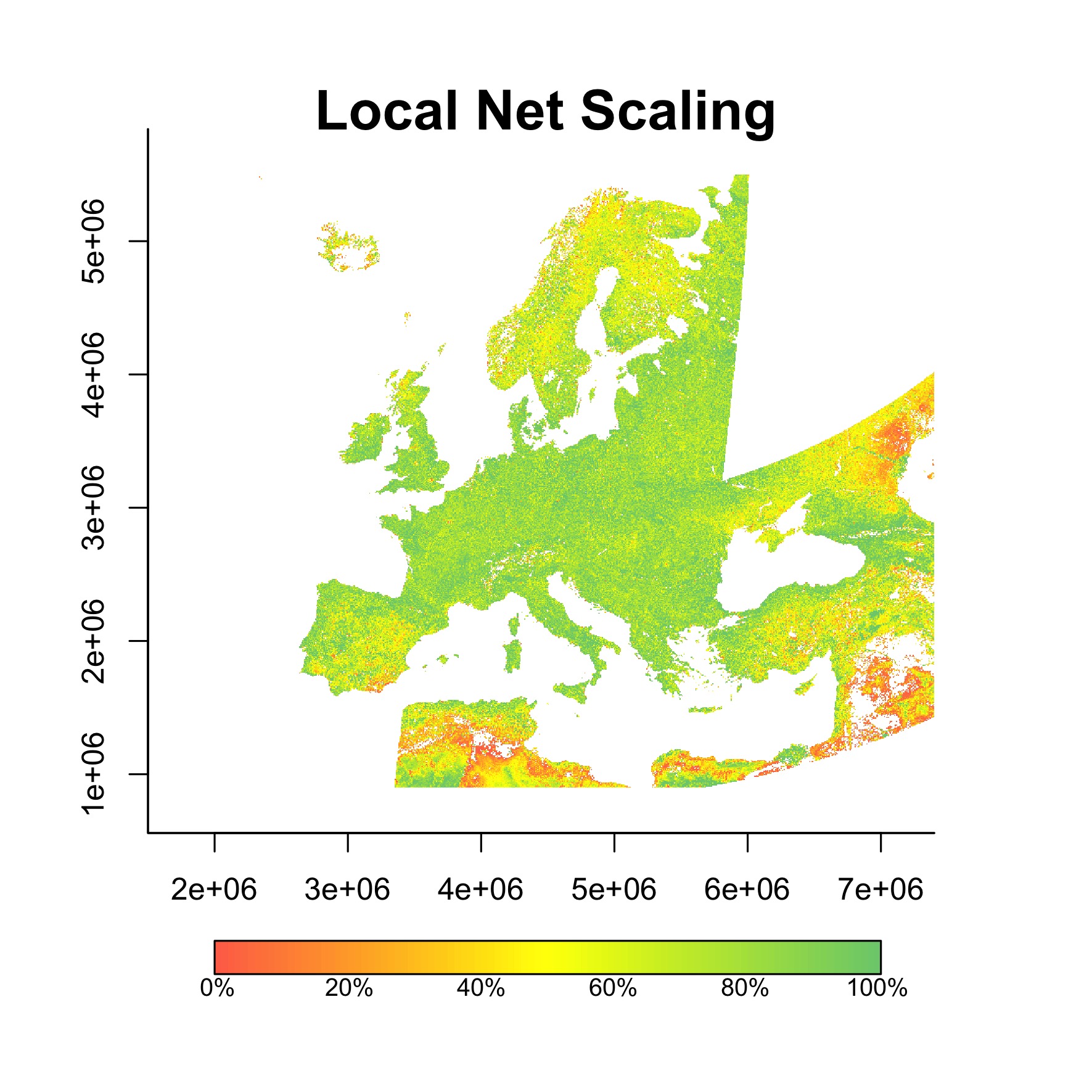


Figure 9: Local Net Primary Production Scaling (LNS): proportion of annual production (i.e. average of the last 5 years of cyclic fraction) over the local potential production (i.e. the 90-percentile within the Ecosystem Functional Type)

For the calculation of the final LPD indicator (i.e. Combined Assessment), these levels of local productivity will be aggregated into two categories: (1) pixels with less than 50% of the highest annual local production (within the EFT) and (2) pixels with more or equal to 50% of annual local production.

# Combined Assessment of Land Productivity

The Land Productivity Dynamics indicator, as shown in the flowchart of the process for its derivation in Figure 1, is based on the combination of two main sources of information: a map of the tendency, positive or negative, of the level of land productivity along the time series, and another map capturing the current level of productivity of each pixel relative to the maximum productivity in a homogeneous land area. As seen above in this document, both branches to calculate the indicator are qualitative methods. Therefore, the final LPD indicator is also a qualitative measure with 5 possible values or categories after the reclassification of each pixel as shown in Table 3. Such categories are (1) d - Declining, (2) ew - Early signs of decline, (3) nf - Stable but stressed, (4) pf - Stable and not stressed and (5) i - Increasing land productivity.

Table 3: Lookup table for the combination of the two branches assessment (i.e. Long Term Change Map and Current Status Map of land productivity) to derive the Land Productivity Dynamics categories (i.e. (1) d - Declining land-productivity, (2) ew - Early signs of decline of land productivity, (3) nf - Stable but stressed land productivity, (4) pf - Stable and not stressed land productivity and (5) i - Increasing land productivity)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Steadiness I. | Baseline L. | State Change |  | Local Sc. |
|  |  |  | < 50% | >= 50% |
| st1 | lo | 0 | **d** | **ew** |
| st1 | lo | 1 | **d** | **ew** |
| st1 | lo | 2 | **d** | **d** |
| st1 | me | 0 | **d** | **ew** |
| st1 | me | 1 | **d** | **ew** |
| st1 | me | 2 | **d** | **d** |
| st1 | hi | 0 | **ew** | **nf** |
| st1 | hi | 1 | **d** | **ew** |
| st1 | hi | 2 | **d** | **ew** |
| st2 | lo | 0 | **nf** | **nf** |
| st2 | me | 0 | **nf** | **nf** |
| st2 | hi | 0 | **nf** | **nf** |
| st3 | lo | 0 | **pf** | **pf** |
| st3 | me | 0 | **pf** | **pf** |
| st3 | hi | 0 | **pf** | **pf** |
| st4 | lo | 0 | **pf** | **i** |
| st4 | lo | 1 | **pf** | **i** |
| st4 | lo | 2 | **i** | **i** |
| st4 | me | 0 | **pf** | **i** |
| st4 | me | 1 | **i** | **i** |
| st4 | me | 2 | **i** | **i** |
| st4 | hi | 0 | **i** | **i** |

In the case study presented in this document, the Land Productivity Dynamics indicator final map (Figure 10) is the result of the combined assessment of the Long Term Change Map (Figure 5) and the Current Status Map of land productivity (Figure 9), both based on the “above ground vegetation productivity” variable.

The function to run the combined assessment to calculate the LPD indicator is *LPD\_CombAssess()* and the following code shows an example of how to call it.

?LPD\_CombAssess  
  
LPD\_finalMap <- LPD\_CombAssess(LandProd\_change = "Long\_Term\_Change\_Map",   
 LandProd\_current = "LNScal",  
 filename = "LPD\_finalMap.tif")  
plot(LPD\_finalMap)

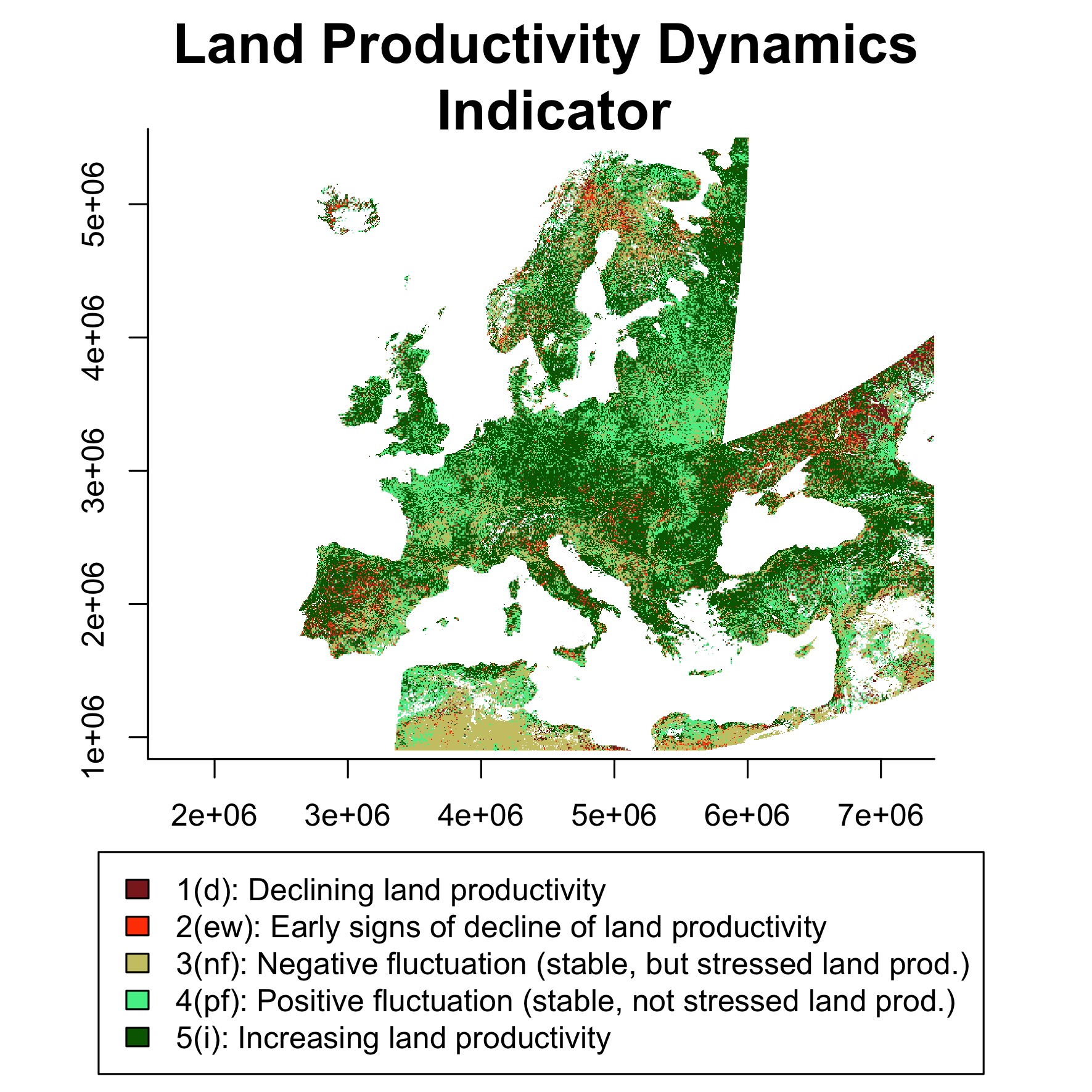


Figure 10: Land Productivity Dynamics indicator final map. Combined assessment of the Long Term Change Map and the Current Status Map of land productivity.(1) d - Declining land-productivity, (2) ew - Early signs of decline of land productivity, (3) nf - Stable but stressed land productivity, (4) pf - Stable and not stressed land productivity and (5) i - Increasing land productivity

## Alternative method for the LPD indicator

It has been shown previously that it is important to include in the LPD calculation the current level of land productivity relative to its potential, given that it may indicate degradation in areas with a positive tendency of productivity, but where the level of productivity still remains low relative to other similar areas nearby. Despite this, the user might want to derive the final product based only on the tendency map (i.e. Long Term Change Map), avoiding the inclusion of the Current Status Map derived with the Local Net Scaling approach. The function *LPD\_CombAssess()* has the potentiality to do it by passing the argument *LandProd\_current = NULL*. By doing so, the function reclassifies the Long Term Change Map into the same 5 categories of the LPD indicator described above, as shown in Table 4.

Table 4: Lookup table for the reclassification of the Long Term Change Map into the Land Productivity Dynamics categories (i.e. (1) d - Declining land-productivity, (2) ew - Early signs of decline of land productivity, (3) nf - Stable but stressed land productivity, (4) pf - Stable and not stressed land productivity and (5) i - Increasing land productivity)

|  |  |  |  |
| --- | --- | --- | --- |
| Steadiness I. | Baseline L. | State Change | LPD class |
| st1 | lo | 0 | **d** |
| st1 | lo | 1 | **d** |
| st1 | lo | 2 | **d** |
| st1 | me | 0 | **d** |
| st1 | me | 1 | **d** |
| st1 | me | 2 | **d** |
| st1 | hi | 0 | **ew** |
| st1 | hi | 1 | **d** |
| st1 | hi | 2 | **d** |
| st2 | lo | 0 | **nf** |
| st2 | me | 0 | **nf** |
| st2 | hi | 0 | **nf** |
| st3 | lo | 0 | **pf** |
| st3 | me | 0 | **pf** |
| st3 | hi | 0 | **pf** |
| st4 | lo | 0 | **pf** |
| st4 | lo | 1 | **pf** |
| st4 | lo | 2 | **i** |
| st4 | me | 0 | **pf** |
| st4 | me | 1 | **i** |
| st4 | me | 2 | **i** |
| st4 | hi | 0 | **i** |

A comparison of the final LPD indicator map produced using the Combined Assessment (i.e. Long Term Change Map + Current Status Map) with the one developed avoiding the Current Status Map for the case study can be seen in Figure 11 (Map 1 and Map 2, respectively). In addition, the “differences map” in the same figure represents those pixels showing a different class from one approach to the other. Notice that this difference is always equal to minus 1, which means that the pixels only change 1 class and that the Combined Assessment has always classes with larger values (i.e. better conditions in terms of land productivity). Table 5 shows the number of pixels which change from one class to another. From this table it can be seen how pixels never change from negative to positive dynamics (class 3 to 4) or from positive to negative (class 4 to 3).



Figure 11: Land Productivity Dynamics indicator final maps derived by the reclassification of the Long Term Change Map of land productivity (Map 1) and produced by the Combined Assessment (Map 2; Long Term Change Map + Current Status Map). Differences Map (Map 1 - Map2) represents in red those pixels showing different resulting classes from both approaches.

Table 5: Number of pixels showing different class in the Combined Assessment approach and in the non combined one (i.e. reclassification of the Long Term Change Map). Only these three combinations have been found in the case study

|  |  |  |  |
| --- | --- | --- | --- |
| No Combined Assessment - Class | Combined Assessment - Class | Number Pixels | Description |
| 1 | 2 | 1861147 | Declining to Early signs of decline |
| 2 | 3 | 334957 | Early signs of decline to Stable but stressed |
| 4 | 5 | 1772401 | Stable not stressed to Increasing |

Finally, Figure 12 shows the proportion of pixels per LPD class under each approach, both for the whole extent (i.e. Europe) and also splitting the map by biogeographical regions. The biogeographical regions have been defined with the official delineations used in the Habitats Directive (92/43/EEC) and for the EMERALD Network, which are freely distributed as a spatial data set by the European Environmental Agenct - Europeam Commission (<https://www.eea.europa.eu/data-and-maps/figures/biogeographical-and-marine-regions-in>).

The plots show that there are some differences in the proportion of pixels per class for each of the two approaches. For example, the Anatolian, the Arctic and the Steppic regions are the three showing more differences, which range from 10 to 12.5% for some LPD classes. This fact supports the importance of including the Current Status Map in the calculations to refine the LPD indicator final results.

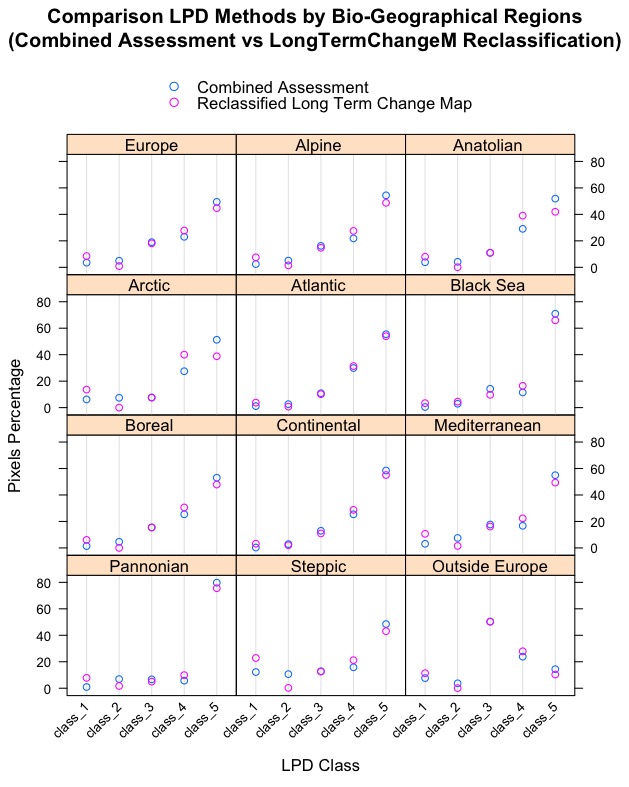


Figure 12: Proportion of pixels per LPD class for the Combined Assessment (light blue) and for the reclassified Long Term Change Map (purple), for Europe and by biogeographical regions

## Land Productivity Dynamics partial indicator

As seen in the previous subsections regarding the derivation of the tendency map (i.e. Long Term Change Map), its final result is quite related to the extremes of the time series. Then, if the time series is long, the LPD indicator shows a big picture of what has happened regarding the land productivity dynamics between the beginning and the end of the period under study. However, to understand the dynamics of the biomass within this period of time, as well as to assess the stability of the final product, it might be useful to produce several “partial LPD indicators” using different time windows of the time series.

This process is not yet implemented in *LPDynR* but we propose the following code to produce partial LPD maps of *n* years and with an overlap of *y* years between the end of the last period and the beginning of the next one. This example is implemented for the same case study shown along this document and the final partial LPD maps can be seen in Figure 13.

## Running LPDynR for partial time series ##   
   
ts\_length <- 5 # time series length to run 'partial LPD maps'  
ts\_years\_overlap <- 1 # number of years of overlapping  
partial\_dir <- "/LPD\_partial" # directory to save the 'partial LPD' results  
first\_year <- 1 # first year of the whole time series  
last\_year <- nlayers(cf) # last year of the whole time series  
last\_year\_run <- first\_year + ts\_length - 1 # last year of the 'partial LPD'  
  
while(last\_year\_run <= last\_year){  
 # subsetting the years (layers) to run  
 cf\_run <- cf[[first\_year:last\_year\_run]]  
  
 # a directory to save the data  
 dir2save0 <- paste0(getwd(), partial\_dir)  
 if(!dir.exists(dir2save0)) dir.create(dir2save0)  
 dir2save <- paste0(getwd(), partial\_dir, "/LPD\_", first\_year, "\_", last\_year\_run, "/")  
 if(!dir.exists(dir2save)) dir.create(dir2save)  
   
 ## ##  
 ## ##  
 ## Here all the steps to calculate the ##   
 ## final LPD map as in the examples ##  
 ## ##  
 ## ##   
  
  
 # Parameters for the loop  
 first\_year <- last\_year\_run - ts\_years\_overlap + 1  
 last\_year\_run <- first\_year + ts\_length - 1  
}

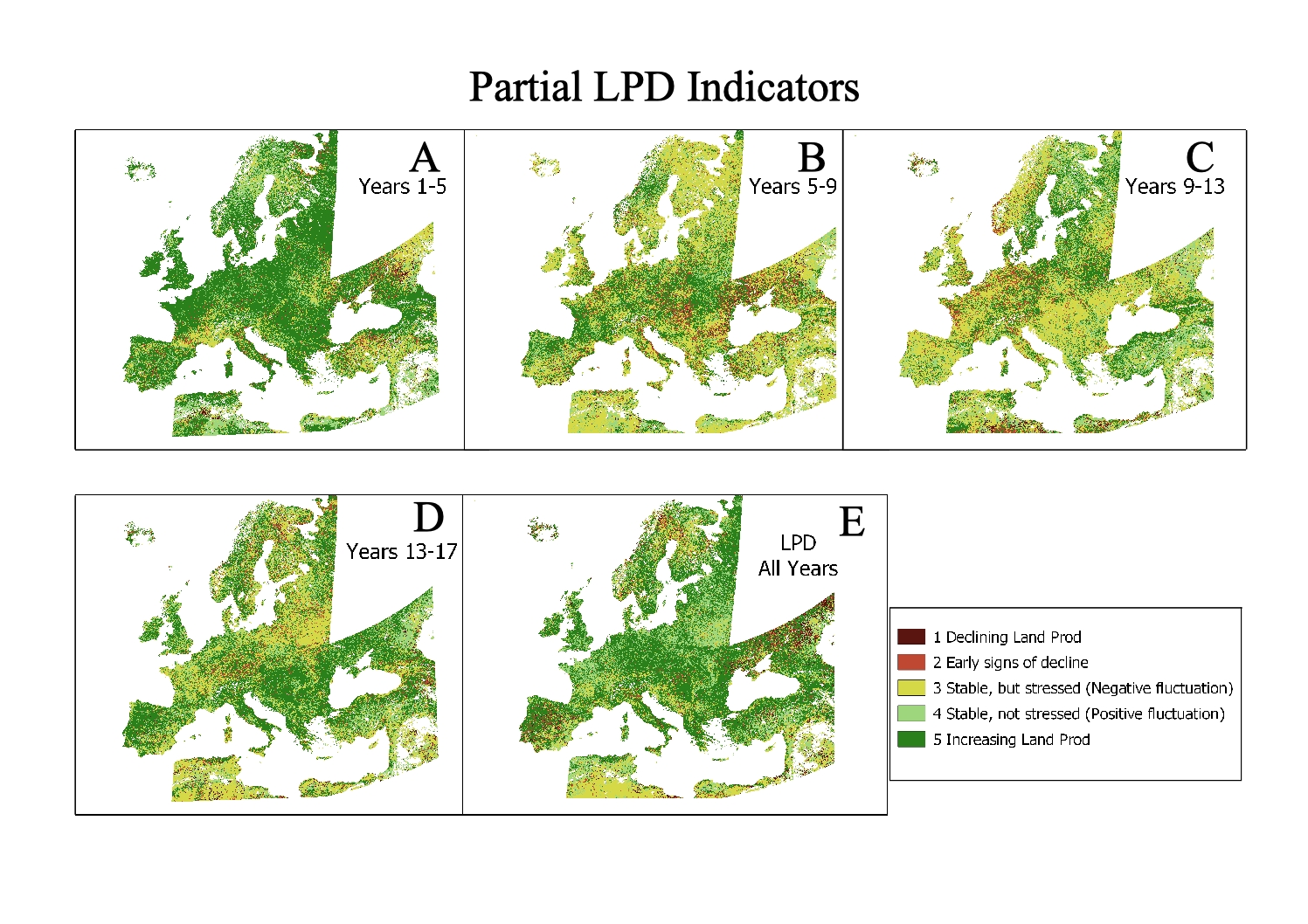


Figure 13: Partial LPD indicators (plots A to D) and LPD indicator for the whole time series (plot E). The partial LPD indicators have been produced for time windows of 5 years with an overlap of 1 year between the end of the last period and the beginning of the next one

The complete LPD indicator (i.e. for the whole time series; Figure 13E) shows, in general terms, a positive trend pattern across Europe (i.e. more pixels in greens). However, the intermediate plots, especially Figure 13B to D, show more negative trends (i.e. yellow and light red pixels). In addition, the first partial indicator shows much more increasing land productivity in the period between years 1 to 5. This, besides demonstrating the highly fluctuating character of vegetation, confirms the influence of the extremes of the time series on the final result. In this sense, in the time series of the example, the first year(s) seems to be quite low in terms of productivity for most of the pixels in Central/Northern Europe, and they suffer a big increase around years 5/6. Such increase causes a large number of areas belonging to the higher LPD class, and it still influences the dynamics of the following periods, resulting in areas with stressed vegetation.

The fact that the LPD indicator calculated with the approach included in *LPDynR* is quite influenced by the beginning and the end of the time series does not suppose a limitation, but it is convenient as the main goal of the LPD indicator is to know the current state of vegetation in realtion to a distant previous state, and not the fluctuations due, for example, to exeptional or unusual climatic events. **Maybe furhter explanations would be good**

# Conclusions

As stated by the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services (IPBES), land degradation leads to a loss of biodiversity and a reduction of ecosystem functions and delivered services all over the world. Therefore, combating land degradation and restoring degraded lands has become an urgent priority in order to protect all life on Earth as well as to ensure human well-being (IPBES 2018). In this sense, satellite observations provides valuable data which might help to monitor the Earth’s land cover to evaluate the state of land degradation.

The Land Productivity Dynamics indicator (LPD), as part of the SDG-15.3.1 indicator, aims at assessing the state of land degradation and desertification at global, regional and local scales. In turn, the *LPDynR* tool presented in this document has been developed to derive the LPD indicator using phenological and land productivity variables, which have been obtained from long-term time series of earth observation imagery.

*LPDynR* is a comprehensive set of programming code, written in the well-known R language and properly packaged, ready to be freely distributed in order to let the users calculate the LPD indicator with a minimum knowledge of the R language. The package, once installed, includes several examples and a small data set for testing the functionalities and the different parameters to tune them. The open source code of *LPDynR* version 1.0.0 can be found and installed from <https://github.com/xavi-rp/LPDynR>.

# References

Cherlet, M., C. Hutchinson, J. Reynolds, J. Hill, S. Sommer, and G. von Maltitz, eds. 2018. *World Atlas of Desertification*. Luxembourg: Publication Office of the European Union.

Dubovyk, Olena. 2017. “The Role of Remote Sensing in Land Degradation Assessments: Opportunities and Challenges.” *European Journal of Remote Sensing* 50 (1): 601–13. <https://doi.org/10.1080/22797254.2017.1378926>.

FAO. 2019. “Trees, Forests and Land Use in Drylands: The First Global Assessment.” FAO Forestry Paper 184. Rome: FAO.

Fensholt, Kjeld Kaspersen, Rasmus Rasmussen. 2013. “Assessing Land Degradation/Recovery in the African Sahel from Long-Term Earth Observation Based Primary Productivity and Precipitation Relationships.” *REMOTE SENSING* 5 (2): 664–86.

Guo, W. Q., T. B. Yang, J. G. Dai, L. Shi, and Z. Y. Lu. 2008. “Vegetation Cover Changes and Their Relationship to Climate Variation in the Source Region of the Yellow River, China, 1990–2000.” *International Journal of Remote Sensing* 29 (7): 2085–2103. <https://doi.org/10.1080/01431160701395229>.

IPBES. 2018. “The Ipbes Assessment Report on Land Degradation and Restoration. Montanarella, L., Scholes, R., and Brainich, a. (Eds.).” Bonn, Germany: Secretariat of the Intergovernmental Science-Policy Platform on Biodiversity; Ecosystem Services.

Ivits, E., and M. Cherlet. 2013. “Land-Productivity Dynamics Towards Integrated Assessment of Land Degradation at Global Scales.” Technical Report EUR 26052. Joint Research Centre of the European Commission.

Ivits, E., M. Cherlet, S. Horion, and R. Fensholt. 2013. “Global Biogeographical Pattern of Ecosystem Functional Types Derived from Earth Observation Data.” *Remote Sensing* 5 (7): 3305–30. <https://doi.org/10.3390/rs5073305>.

Ivits, E., M. Cherlet, W. Mehl, and S. Sommer. 2013. “Ecosystem Functional Units Characterized by Satellite Observed Phenology and Productivity Gradients: A Case Study for Europe.” *Ecological Indicators* 27: 17–28. <https://doi.org/10.1016/j.ecolind.2012.11.010>.

Jin, Hongxiao, and Lars Eklundh. 2014. “A Physically Based Vegetation Index for Improved Monitoring of Plant Phenology.” *Remote Sensing of Environment* 152: 512–25. <https://doi.org/https://doi.org/10.1016/j.rse.2014.07.010>.

Jönsson, P., and L Eklundh. 2004. “TIMESAT—a Program for Analyzing Time-Series of Satellite Sensor Data.” *Computers & Geosciences* 30: 833–45.

Middleton, N., L. Stringer, A. Goudie, and D. Thomas. 2011. “The Forgotten Billion. MDG Achievement in the Drylands.” New York, NY, 10017, USA: United Nations Development Programme.

Orr, B. J., A. L. Cowie, V. M. Castillo Sanchez, P. Chasek, N. D. Crossman, A. Erlewein, G. Louwagie, et al. 2017. “Scientific Conceptual Framework for Land Degradation Neutrality. In: A Report of the Science-Policy Interface.” Bonn, Germany: UNCCD.

Prince, I. Rishmawi, S. D. Becker-Reshef. 2009. “Detection and Mapping of Long-Term Land Degradation Using Local Net Production Scaling: Application to Zimbabwe.” *REMOTE SENSING OF ENVIRONMENT* 113 (5): 1046–57.

Sims, N. C., C. Green, G. J. Newnham, J. R. England, A. Held, M. A. Wulder, and et al. 2017. *Good Practice Guidance Sdg Indicator 15.3.1. Proportion of Land That Is Degraded over Total Land Area*. First. United Nations Convention to Combat Desertification.

Sims, Neil C., Nichole N. Barger, Graciela I. Metternicht, and Jacqueline R. England. 2020. “A Land Degradation Interpretation Matrix for Reporting on Un Sdg Indicator 15.3.1 and Land Degradation Neutrality.” *Environmental Science & Policy* 114: 1–6. <https://doi.org/https://doi.org/10.1016/j.envsci.2020.07.015>.

UN. 2015. “Transforming Our World: The 2030 Agenda for Sustainable Development.” A/RES/70/1. United Nations.

UNCCD. 2015. “Integration of the Sustainable Development Goals and Targets into the Implementation of the United Nations Convention to Combat Desertification and the Report of the Intergovernmental Working Group on Land Degradation Neutrality.” ICCD/COP(12)/4. Ankara, Turkey: UNCCD Conference of the Parties.

Yengoh, G. T., D. Dent, L. Olsson, A. E. Tengberg, and C. J. Tucker. 2015. *Use of the Normalized Difference Vegetation Index (Ndvi) to Assess Land Degradation at Multiple Scales*. Springer Briefs in Environmental Science. Springer.