

# OpenLand: Software for Quantitative Analysis and Visualization of Land Use and Cover Change

by Reginal Exavier and Peter Zeilhofer

**Abstract** There is an increasing availability of spatially explicit, freely available land use and cover (LUC) time series worldwide. Because of the enormous amount of data this represents, the continuous updates and improvements in spatial and temporal resolution and category differentiation, as well as increasingly dynamic and complex changes made, manual data extraction and analysis is highly time consuming, and making software tools available to automatize LUC data assessment is becoming imperative. This paper presents a software developed in R, which combines LUC raster time series data and their transitions, calculates state-of-the-art LUC change indicators, and creates spatio-temporal visualizations, all in a coherent workflow. The functionality of the application developed is demonstrated using an LUC dataset of the Pantanal floodplain contribution area in Central Brazil.

## Introduction

Land use and land cover (LUC) monitoring provides key information on the ecological state and biophysical properties of the land surface and is widely used in climatic, hydrological, and ecological modelling (Brovkin et al., 2013; Verburg et al., 2015). Global population growth over the last decades has led to increased rates of LUC change (LUCC), which has affected all major ecosystem services, including biodiversity, climate, and water supply, and has altered carbon cycling (Ballantyne et al., 2015; Nelson et al., 2010; Song et al., 2018).

The importance of these applications is supported by the continuous expansion of remote sensing data acquisition programs, making freely available an increasing amount of spatially explicit LUC time series data worldwide (Prestele et al., 2016). The huge volume of these datasets, their timely actualization, ever-increasing spatial resolution and category differentiation, and a variety of data formats, render their manual extraction and analysis increasingly time consuming. This is especially true when trying to compare or harmonize LUC datasets of the same study area generated from different sources, with different methodological approaches across different spatial and temporal scales, in order to output the factual LUC in a region of interest (Yang et al., 2017).

Human interference on LUC is complex, and the intensity and frequency of LUC transitions are increasing. Therefore, the development of sound LUC models is highly dependent on a deep understanding of past and ongoing LUCC processes (Müller and Munroe, 2014). This means that LUC patterns have to be constantly reviewed to underpin our understanding of these processes (Lambin, 1997; Lambin and Geist, 2006), develop a baseline analysis for projections of future LUC (Hurtt et al., 2011), and to construct, calibrate and validate LUCC simulations (Prestele et al., 2016). Such efforts require tools for data extraction, pre-processing, visualization, and calculation of LUC metrics to relieve the burden from labour-intensive and time-consuming manual data processing and analysis (Yu et al., 2019). If the procedures implemented to do so follow standards in land cover characterization, and use formalized methodological approaches in LUC analysis, the resulting analytics become more transparent, robust, and auditable (Herold et al., 2006; Müller and Munroe, 2014; Yang et al., 2017).

Aldwaik and Pontius (2012) developed an approach called Intensity Analysis (IA), which examines changes in LUC categories by comparing the intensity of change between categories during a given time interval with a hypothesized uniform change intensity. Since then, several case studies have emphasized the potential of Intensity Analysis to synthesize complex LUCC under different spatial and temporal scales, such as urban environments (Akinyemi et al., 2017; Subasinghe et al., 2016), regional studies (Melo et al., 2018; Mwangi et al., 2017; de Souza et al., 2017), and country-wide comparisons (Chaudhuri and Mishra, 2016). Furthermore, Huang et al. (2018) concluded that intensity analysis metrics outperform other indicators of land use dynamics in the comparison between candidate regions, while Varga et al. (2019) showed that IA metrics are a helpful tool for assessing the quality of LUC modelling outputs.

In IA, the uniform change intensity of a time period is compared with the observed intensities of transitions between LUC categories. The calculations are carried out using transition matrices between LUC categories, assessing specific: (i) time intervals, (ii) categories or classes and (iii) types of transitions. The first assessment level examines in which time intervals global annual change rates are faster, slower or comparable to an average rate of change. The second level determines which category transitions are relatively dormant or more active within a given time interval, based on an

analysis of gross change. The gross gains and losses for each category in a given time interval are compared with an averaged, uniform annual change. The third level of assessment seeks to identify which transitions are particularly intensive during the time interval considered. In order to do so each transition is compared with a uniform transition intensity. The mathematical notations used for the IA indicators measuring size and intensities of temporal changes among LUC categories (Equations 1-8) can be found in [Aldwaik and Pontius \(2012\)](#) and [Aldwaik and Pontius \(2013\)](#).

Analytical tools for the analysis of LUC time series are commonly available in compiled languages and distributed as software packages or extensions to proprietary geographic information systems such as ArcGIS or Idrisi ([Moulds et al., 2015](#)). Consequently, the source code for such tools, used for land use change analysis and modelling, is often unavailable ([Rosa et al., 2014](#)). This makes adopting the applications of new approaches and reproducing scientific results difficult ([Morin et al., 2012](#); [Peng, 2011](#)). GIS software has made widely available spatially explicit visualization capabilities for multi-temporal LUC, which is crucial for documentation and analysis of LUC distribution. However, maps can become blurred and non-interpretable when extensive time series or complex landscapes with different category transitions need to be analysed. This is particularly true when multi-category changes have to be analysed at multiple time steps to improve the understanding of landscape processes, in which case non-spatial forms of LUCC will have to be explored.

By contrast, the availability of packages to perform analyses of LUC categories in time series data is limited. **intensity.analysis**, developed by [Pontius Jr. and Khallaghi \(2019\)](#), is an application of IA, with functionality limited to the calculation of the three IA assessment levels described above and their plotting; this package does not allow any personalization. Little pre-processing capability is available. If needed, input rasters cannot be checked for spatial and thematic consistency, and they have to be manually imported to R and stored. No other commonly used LUC metrics can be calculated or plotted. Tools to visualize LUC changes are limited to standard plots of the three IA analysis levels. **lulcc** ([Moulds et al., 2015](#)) is another package developed for LUC change analysis and focuses on the application of tools for LUC change modelling.

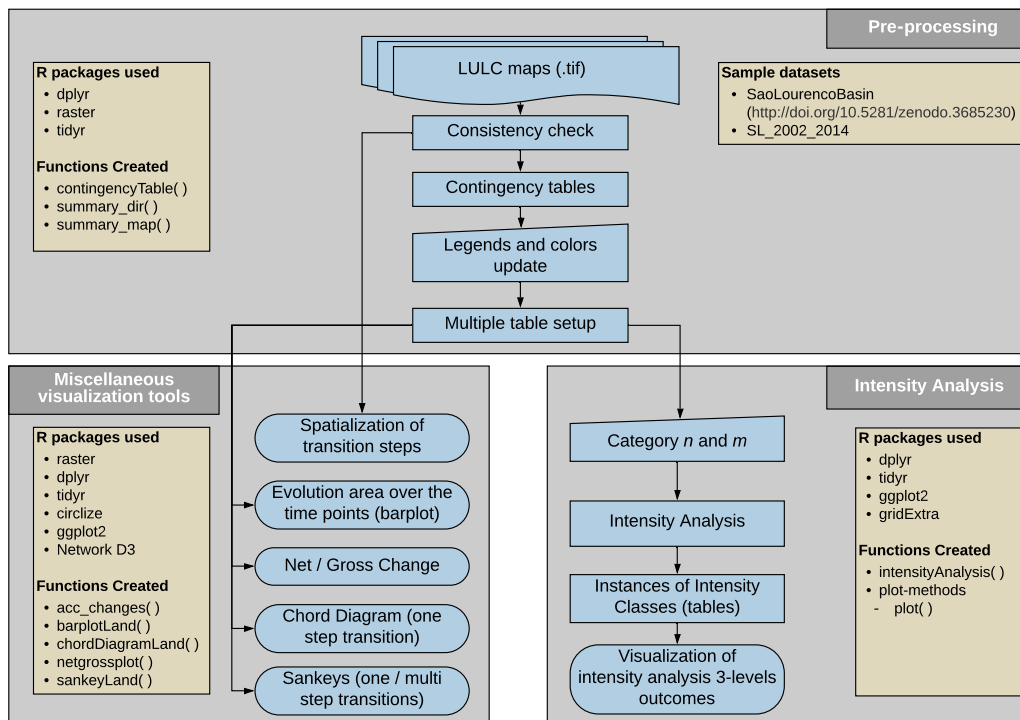
This highlights the need for flexible and comprehensive tools for multipurpose analyses of complex LUC time series data. This paper aims to demonstrate that we have developed a solution, available as freeware, which allows the agile consistency analysis, extraction, pre-processing, analysis, and visualization of multiresolution time series data in a straightforward workflow, thereby increasing productivity in information extraction for an improved understanding of LUCC processes at different spatial scales. Originating primarily from the application of the intensity analysis method, the formatted tabular representation of multiple transition steps can be a valuable tool to support the quick calculation of LUCC indicators.

## Conceptual overview

The software was developed to provide comprehensive support for LUCC analysis over time and implement the intensity analysis conceptual approach proposed by [Aldwaik and Pontius \(2012\)](#), and further described in [Aldwaik and Pontius \(2013\)](#).

The source language of the **OpenLand** package ([Exavier and Zeilhofer, 2020](#)) is R, and it was conceived using the integrated development environment (IDE) RStudio ([RStudio Team, 2016](#)). For full functionality, **OpenLand** depends on selected, third-party functions of the **raster**, **dplyr** and **tidyr** R packages ([Hijmans, 2019](#); [Wickham et al., 2019](#); [Wickham and Henry, 2019](#)) (Fig. 1). Visualization tools make use of **ggplot2**, **circlize**, **gridExtra** and **networkD3** R packages ([Allaire et al., 2017](#); [Auguie, 2017](#); [Gu et al., 2014](#); [Wickham, 2016](#)). Designed as an iterative workflow, the pre-processing of a raw LUC time series is followed by the extraction and visualization of single or multistep LUC transitions and/or a complete IA.

The main processing workflow should begin with a consistency check of the input files, which have to be a sequence of LUC maps in a tif format (Fig. 1). In the following step, a contingency table of LUC transitions for each time step can be calculated. Two complementary analysis workflows are then available. First, a set of miscellaneous non-spatial visualization tools allows for a quick screening of LUC dynamics. Second, a complete intensity analysis ([Aldwaik and Pontius, 2012, 2013](#)) can be implemented, including tools to visualize change intensities over time, category and transition levels. Ancillary functionalities include allow for extraction and spatial visualization of LUCC frequencies.



**Figure 1:** Conceptual overview of the OpenLand package showing reused R functionalities. The package is composed by three major blocks for *i*) pre-processing of raster time series and calculus of LUC transitions, *ii*) spatial and non-spatial visualization of LUCC and *iii*) intensity analysis including tailored visualization tools.

## Functionality and implementation

### The São Lourenço river basin example dataset

The **OpenLand** functionality is demonstrated using an LUC dataset of the São Lourenço river basin, of major importance to the Pantanal wetland into which it flows. The data is as provided in the 4<sup>th</sup> edition of the [Monitoring of Changes in Land cover and Land Use in the Upper Paraguay River Basin - Brazilian portion - Review Period: 2012 to 2014](#) (Instituto SOS Pantanal and WWF-Brasil, 2015), and the time series is composed of five LUC maps (2002, 2008, 2010, 2012 and 2014). The study area is located in the Cerrado savanna biome, in the southeastern corner of the Brazilian state of Mato Grosso, and covers approximately 22,400 km<sup>2</sup>. Some level of LUCC has occurred in about 12% of the area over the past 12 years, including some deforestation and intensification of existing agricultural uses. In order to be processed using the **OpenLand** package, the original multi-year shapefile was clipped to the extent of the São Lourenço basin, transformed into rasters and saved as a 5-layer RasterStack; it is available from a public repository (<https://doi.org/10.5281/zenodo.3685229>) as an .RDA file which can be loaded into R.

```
# Installing the released version of OpenLand from CRAN
install.packages("OpenLand")
```

```
# Loading the OpenLand package
library(OpenLand)
```

```
# downloading the SaoLourencoBasin multi-layer raster and make it available into R
url <- "https://zenodo.org/record/3685230/files/SaoLourencoBasin.rda?download=1"
```

```
temp <- tempfile()
download.file(url, temp, mode = "wb")
load(temp)
```

```
# looking on the metadata of the example dataset
SaoLourencoBasin
```

```
#> class      : RasterStack
#> dimensions : 6372, 6546, 41711112, 5 (nrow, ncol, ncell, nlayers)
#> resolution : 30, 30 (x, y)
#> extent     : 654007.5, 850387.5, 8099064, 8290224 (xmin, xmax, ymin, ymax)
#> crs        : +proj=utm +zone=21 +south +ellps=GRS80 +units=m +no_defs
#> names      : landscape_2002, landscape_2008, landscape_2010, ..., landscape_2014
#> min values :      2,      2,      2, ...,      2
#> max values :     13,     13,     13, ...,     13
```

To visualize the output of the LUC analysis in **OpenLand**, we simplified the legend of the original dataset and used the following 11 LUC categories for the basin: forest formation (FF), three Cerrado savanna formations (SF, SA, SG), anthropogenized vegetation (aa), i.e. mostly altered Cerrado formations used for grazing, composed of natural species, cattle farming (Ap), crop farming (Ac), mining areas (Im), urban areas (Iu), water bodies (Agua), and reforestation (R) (Table 1).

**Table 1:** The original legend of LUC classes and categories from [Instituto SOS Pantanal and WWF-Brasil \(2015\)](#) including colour coding.

Pixel Value	Legend	Class	Category	Colour
2	Ap	Anthropogenic	Cattle farming	#DDCC77
3	FF	Natural	Forest formation	#117733
4	SA	Natural	Park savanna	#44AA99
5	SG	Natural	Gramineous savanna	#88CCEE
7	aa	Anthropogenic	Anthropogenized vegetation	#CC6677
8	SF	Natural	Wooded savanna	#999933
9	Agua	Natural	Water bodies	#332288
10	Iu	Anthropogenic	Urban areas	#AA4499
11	Ac	Anthropogenic	Crop farming	#661100
12	R	Anthropogenic	Reforestation	#882255
13	Im	Anthropogenic	Mining areas	#6699CC

### Consistency check and data extraction from raster time series

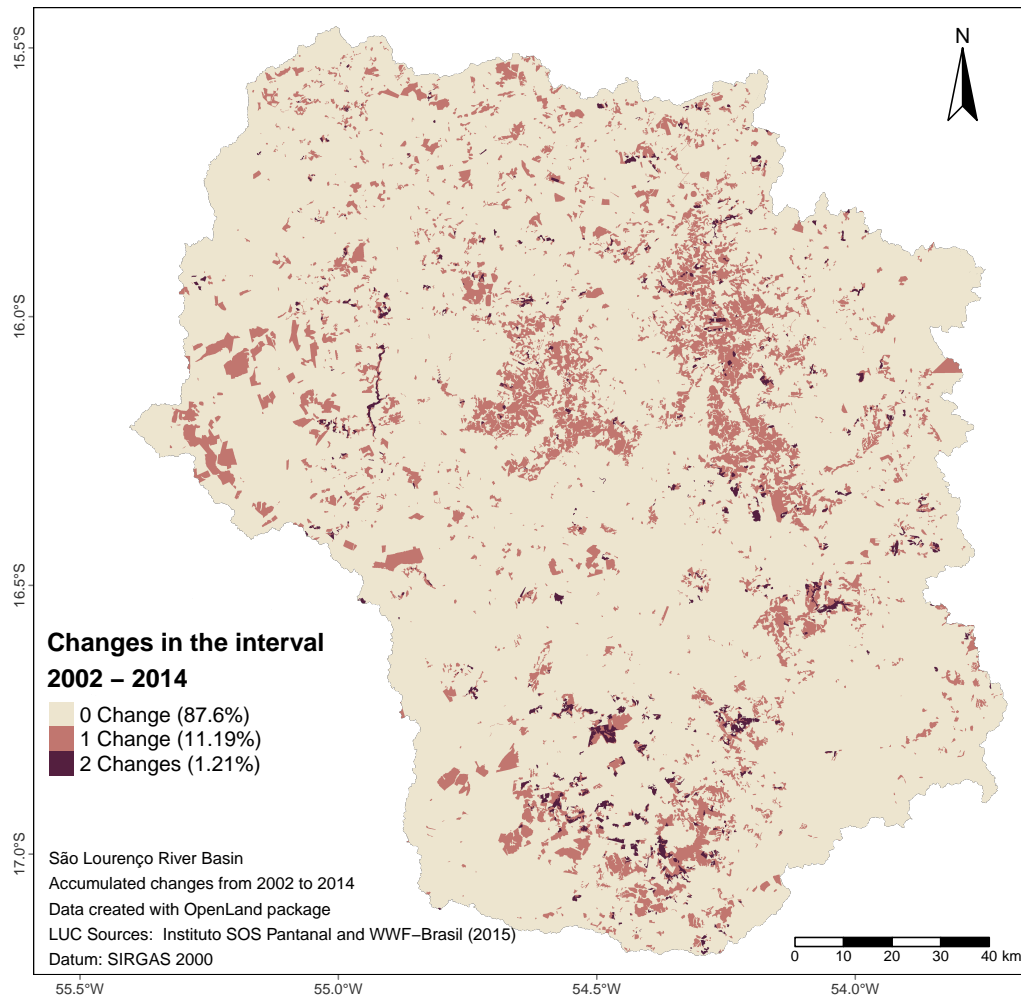
Two auxiliary functions allow users to check for consistency in the input gridded LUC time series, including extent, projection, cell resolution and categories. The `summary_map()` function returns the number of pixels in each category for each single raster layer, whereas the `summary_dir()` function lists the spatial extent, spatial resolution, cartographic projection and the category range of a set of LUC maps.

For the initial spatial screening of the time series, the `acc_changes()` function determines the number of LUC transitions during the entire time interval of the series. The results of percentage area by transition frequencies in the study area are stored in a table and a grid layer is generated, which can be plotted (Fig. 2) using for example the `tmap` package.

```
# the acc_changes() function, with the SaoLourencoBasin dataset
SL_changes <- acc_changes(SaoLourencoBasin)
SL_changes

#> [[1]]
#> class      : RasterLayer
#> dimensions : 6372, 6546, 41711112 (nrow, ncol, ncell)
#> resolution : 30, 30 (x, y)
#> extent     : 654007.5, 850387.5, 8099064, 8290224 (xmin, xmax, ymin, ymax)
#> crs        : +proj=utm +zone=21 +south +ellps=GRS80 +units=m +no_defs
#> names      : layer
#> values     : 0, 2 (min, max)
#>
#>
#> [[2]]
#> # A tibble: 3 x 3
#>   PxValue   Qt Percent
#>   <int>   <int> <dbl>
#> 1       0 21819779  87.6
```

```
#> 2      1 2787995 11.2
#> 3      2 301086  1.21
```



**Figure 2:** Number of LUCC between 2002 and 2014 at four time points (2002, 2008, 2010, 2012, 2014) in the São Lourenço river basin (Instituto SOS Pantanal and WWF-Brasil, 2015). LUCC occurred throughout the basin, with one-step changes concentrated in the central-northern regions and two-step changes in the central south.

As most deforestation occurred in the 20th century, unchanged areas totalled 87.6%. Approximately 35% of the watershed had already been deforested in 1985, versus 46% in 2014 (Project MapBiomass, 2019). 11.19% showed a unique alteration and 1.21% a two-folded alteration for the five time points of the input series considered. All further analytical and visualization tools are based on the `contingencyTable()` function, which builds a matrix of transitions between LUC categories according to the temporal resolution of the original time series. Multiple grid scanning by `contingencyTable()` returns 5 objects: `lulc_Multistep`, `lulc_Onestep`, `tb_legend`, `totalArea`, `totalInterval`. The first two objects are contingency tables; the first (`lulc_Multistep`) takes into account the grid cells of the entire time series, whereas the second (`lulc_Onestep`) calculates LUC transitions only between the first and last year of the series. The third object (`tb_legend`) is a table containing the category name associated with pixel values and a colour scheme. As category values and colours are initially created randomly, their values must be edited to produce meaningful plot legends and colour schemes. The fourth object (`totalArea`) is a table containing the extent of the study area in km<sup>2</sup> and in pixel units. The fifth table (`totalInterval`) stores the range of years between the first ( $Y_{t=1}$ ) and last year ( $Y_T$ ) of the series. Table 2 presents the fields created, together with their format, description and labelling as per a table output by the `contingencyTable()` function.

As mentioned, the `tb_legend` object must be edited with the real category names and colours associated with the category values. In our case, the category names and colours follow the conventions given by (Instituto SOS Pantanal and WWF-Brasil, 2015) (access document here, page 17) like the values in (Table 1).



**Table 2:** Structure of the contingency table that stores LUC transitions at chosen time intervals.

$[Y_t, Y_{t+1}]$	Category <sub>i</sub> Category <sub>j</sub>		$C_{ij}$ (km <sup>2</sup> )	$C_{ij}$ (pixel)	$Y_{t+1} - Y_t$	$Y_t$	$Y_{t+1}$
chr	int	int	dbl	int	int	int	int
Period of analysis from time point $t$ to time point $t+1$	A category at interval's initial time point	A category at interval's final time point	Number of elements in km <sup>2</sup> that transits from category $i$ to category $j$	Number of elements in pixel that transits from category $i$ to category $j$	Interval in years between time point $t$ and time point $t+1$	Initial Year of the interval	Final Year of the interval
<b>Period</b>	<b>From</b>	<b>To</b>	<b>km2</b>	<b>QtPixel</b>	<b>Interval</b>	<b>yearFrom</b>	<b>yearTo</b>

Users should be aware that `acc_changes()` and `contingencyTable()` process the entire input time series analyzing successive raster pairs. Processing time of `contingencyTable()`, the computationally most demanding function of **OpenLand** may range between 3 and 7 minutes for the SaoLourencoBasin dataset if using common desktop computers (8 MB Ram, i3-i7 processors, Windows 64bits versions).

```
## creating the contingency table
SL_2002_2014 <- contingencyTable(input_raster = SaoLourencoBasin,
                                pixelresolution = 30)

names(SL_2002_2014)

#> [1] "lulc_Multistep" "lulc_Onestep" "tb_legend" "totalArea"
#> [5] "totalInterval"

## editing the category names
SL_2002_2014$tb_legend$categoryName <- factor(c("Ap", "FF", "SA", "SG", "aa", "SF",
                                                "Agua", "Iu", "Ac", "R", "Im"),
                                              levels = c("FF", "SF", "SA", "SG", "aa", "Ap",
                                                         "Ac", "Im", "Iu", "Agua", "R"))

## adding the colours by the same order of the legend
SL_2002_2014$tb_legend$color <- c("#DDCC77", "#117733", "#44AA99", "#88CCEE",
                                   "#CC6677", "#999933", "#332288", "#AA4499",
                                   "#661100", "#882255", "#6699CC")
```

## Miscellaneous non-spatial visualization tools

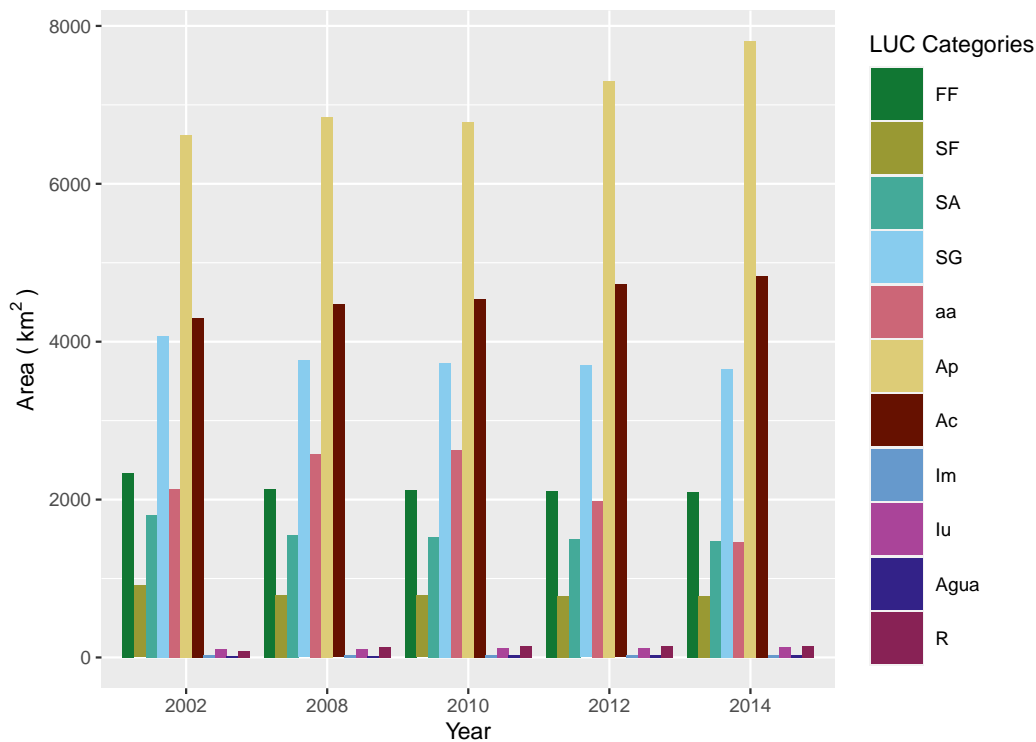
### Evolution of LUC areas

Exploratory data analysis based on the `contingencyTable()` function may begin with the visualization of the absolute and/or percentage area of each LUC category at each time point using a grouped bar plot (Fig. 3), in order to show the evolution of the LUC categories at each time point of the series.

```
barplotLand(dataset = SL_2002_2014$lulc_Multistep,
            legendtable = SL_2002_2014$tb_legend,
            xlab = "Year",
            ylab = bquote("Area ( ~ km^2~ ")"),
            area_km2 = TRUE)
```

### Net and gross changes

For the analysis of long time series with high temporal resolution, information extraction from evolution bar plots may become demanding and alterations between categories through time points are only as net changes. For a category-wise, simultaneous assessment of net and gross changes, multistep transitions can be balanced through a stacked bar chart (Fig. 4). In the São Lourenço river basin, all natural vegetation categories suffered an equal net and gross loss (FF, SF, SA, SG) between 2002 and 2014. In contrast, mining (Im) and urbanized (Iu) areas as well as waterbodies (Agua) and reforestation (R) had equal amounts of net and gross gains. Principally anthropogenized vegetation (aa) and pastures showed differences in net and gross changes, pointing to more complex underlying LUCC processes.



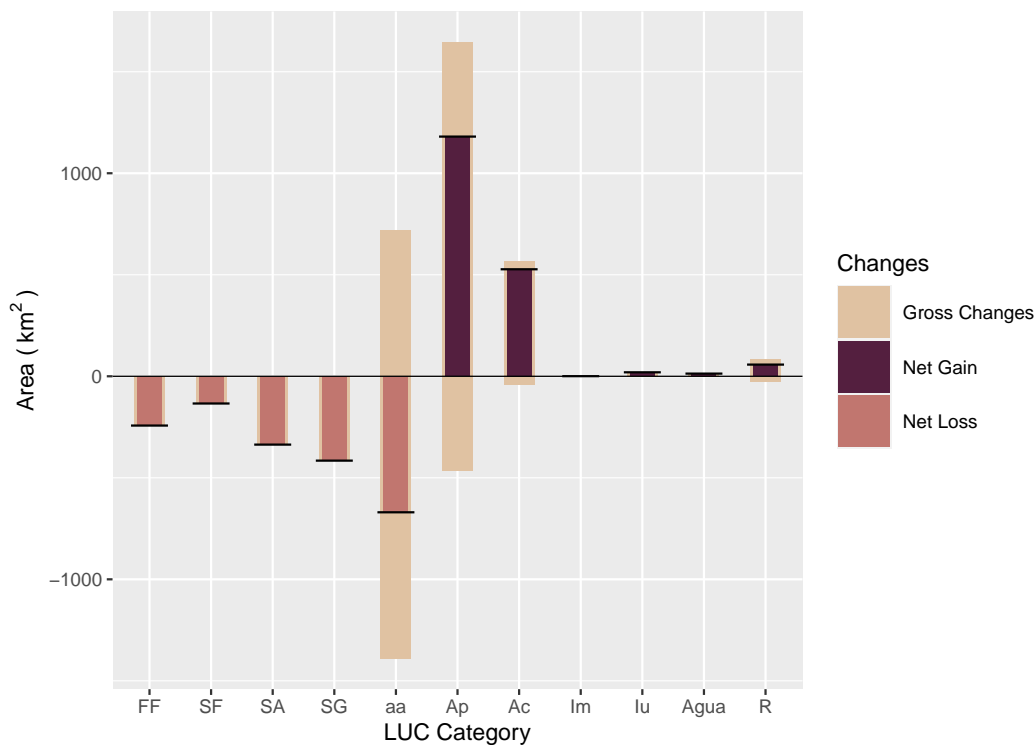
**Figure 3:** Evolution bar plot quantifying LUC categories at the five time points between 2002 and 2014, highlighting the increase of cattle farming (Ap) in the river basin.

```
netgrossplot(dataset = SL_2002_2014$lulc_Multistep,
             legendtable = SL_2002_2014$tb_legend,
             xlab = "LUC Category",
             ylab = bquote("Area ( ~ km^2 ~ ")"),
             changesLabel = c(GC = "Gross Changes", NG = "Net Gain", NL = "Net Loss"),
             color = c(GC = "#e0c2a2", NG = "#541f3f", NL = "#c1766f"),
             area_km2 = TRUE)
```

To further explore specifically those non-linear transitions, single-step Chord (Fig. 5) and Sankey diagrams (Fig. 6a and 6b) can be composed for each time point in the series. Considering the entire observation period, the major gross change was from anthropogenized areas (aa) to cattle farming (Ap). Both the Chord and Sankey diagrams show however that pastures did not directly gain from clear-cut deforestation, but from the previous degradation of natural vegetation categories principally until 2008 (FF, SF, SA, SG), and a subsequent transition from aa to Ap between 2010 and 2014.

#### Chord diagram (2002 – 2014)

```
chordDiagramLand(dataset = SL_2002_2014$lulc_Onestep,
                legendtable = SL_2002_2014$tb_legend,
                area_km2 = TRUE)
```



**Figure 4:** Combined stacked bar plot of gross changes, net gains and net losses of LUC categories (2002 – 2014). Anthropogenized vegetation (aa) and cattle farming (Ap) experimented both gross gains and losses, however net losses and net gains, respectively.

#### Single-step Sankey diagram (2002 – 2014)

```
sankeyLand(dataset = SL_2002_2014$lulc_Onestep,
           legendtable = SL_2002_2014$tb_legend)
```

#### Multistep sankey diagram

If between-category transitions have to be visualized simultaneously for the entire time series, a multistep version of the Sankey plot can be output (Fig. 6b).

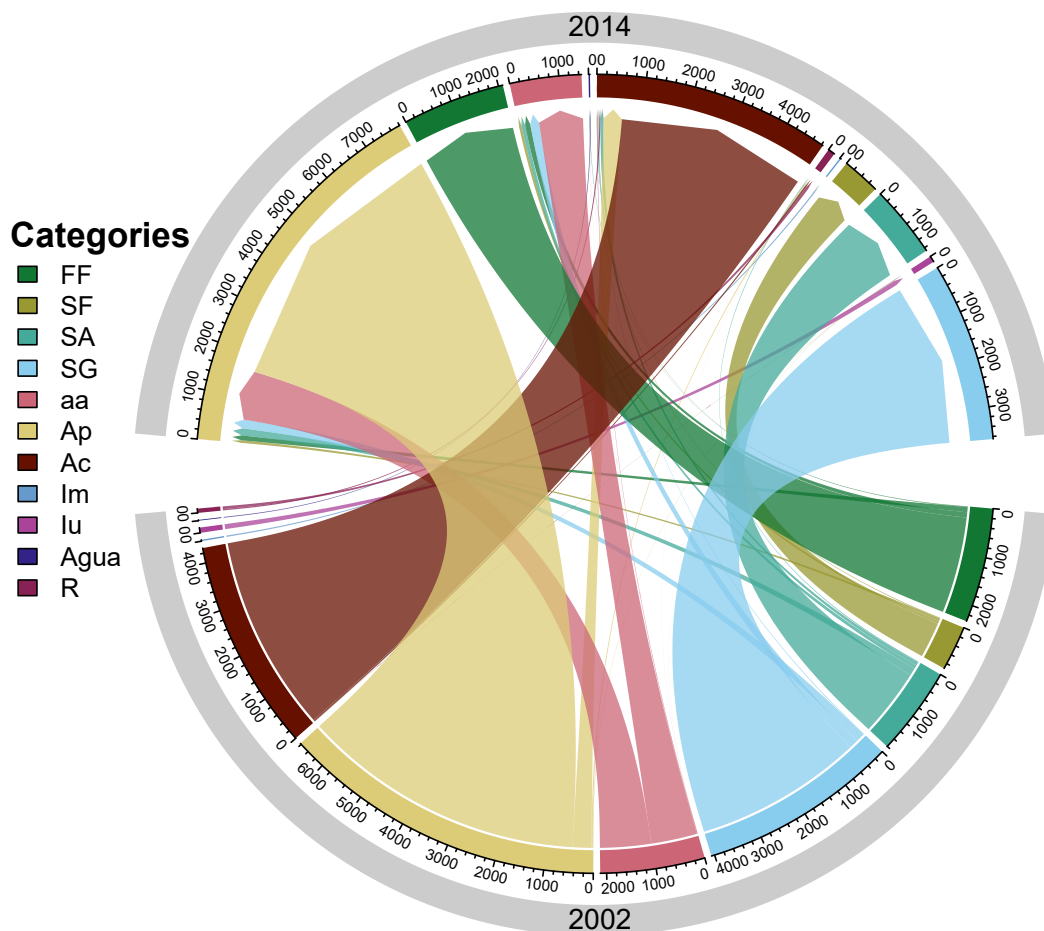
```
sankeyLand(dataset = SL_2002_2014$lulc_Multistep,
           legendtable = SL_2002_2014$tb_legend)
```

The `sankeyLand()` function returns html output, as it depends on the **networkD3** R package which uses such format as default.

#### Intensity analysis

Intensity Analysis (IA) is a quantitative method for the analysis of LUC maps over several time steps, using cross-tabulation matrices, where each matrix summarizes the LUC change for each time interval. IA evaluates the deviation between observed change intensity and hypothesized uniform change intensity in three levels. Thereby, each level details information given by the previous analysis level. First, the **interval level** indicates how the size and rate of change vary over time intervals. Second, the **category level** examines for each time interval how the size and intensity of gross losses and gross gains in each category vary across categories for each time interval. Third, the **transition level** determines the size and intensity of each transition from one category to another during each time interval. At each level, the method also tests for stationarity of patterns across time intervals (Aldwaik and Pontius, 2012). In the **OpenLand** package, the `intensityAnalysis()` function computes the three levels of analysis. It requires the object returned by the `contingenceTable()` function and that the user predefines two LUC categories  $n$  and  $m$ . Generally,  $n$  is a target category that experienced relevant gains and  $m$  a category with important losses.





**Figure 5:** Chord diagram of single-step transitions between LUC categories (2002 – 2014). The major change during the period was from anthropogenized vegetation (aa) to cattle farming (Ap).

```
testSL <- intensityAnalysis(dataset = SL_2002_2014, category_n = "Ap",
                           category_m = "SG", area_km2 = TRUE)

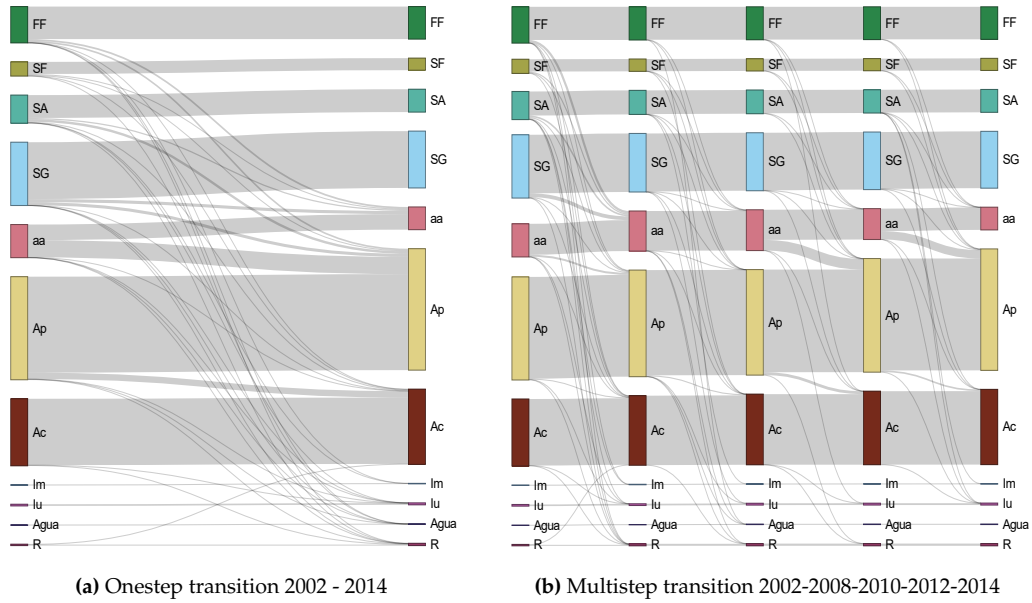
# it returns a list with 6 objects
names(testSL)

#> [1] "lulc_table"          "interval_lvl"        "category_lvlGain"
#> [4] "category_lvlLoss"   "transition_lvlGain_n" "transition_lvlLoss_m"
```

The `intensityAnalysis()` function returns 6 objects: `lulc_table`, `interval_lvl`, `category_lvlGain`, `category_lvlLoss`, `transition_lvlGain_n`, `transition_lvlLoss_m`. The object-oriented approach adopted here allowed us to set specific methods for plotting the intensity objects. Specifically, we used the S4 class, which requires the formal definition of classes and methods (Chambers, 2008). The first object is a contingency table similar to the `lulc_Multistep` object with the unique difference that the columns `From` and `To` are replaced by their appropriate denominations in the LUC legend.

The second object `interval_lvl` is an "Interval" object, the third `category_lvlGain` and the fourth `category_lvlLoss` are "Category" objects, while the fifth `transition_lvlGain_n` and the `transition_lvlLoss_m` are "Transition" objects.

An "Interval" object has one slot containing a table of the **interval level** results ( $(S_t$  equation (1) and  $U$  equation (2) values). A "Category" object has three slots: the first contains the colour associated with the legend item as name attribute, the second slot contains a table of the **category level** results ( $gain (G_{ij})$  equation (3) or  $loss (L_{ij})$  equation (4) values), and the third slot contains a table storing the results of a stationarity test. A "Transition" object also has three slots: the first contains the colour associated with the respective legend item defined as name attribute, the second slot contains a table of the **transition level** results ( $gain n (R_{ijn})$  equation (5) and  $W_{tm}$  equation (6) or  $loss m (Q_{tmj})$  equation (7) and  $V_{tm}$  equation (8) values). The third slot contains a table storing the results of a stationarity test. Aldwaik and Pontius (2012) consider a case stationary only when the intensities for all time intervals



**Figure 6:** Sankey plots of single- and multistep transitions between LUC categories (2002 – 2014). Major overall transition was from anthropogenized vegetation (aa) to cattle farming (Ap) (a), but aa received important contributions by the loss of natural vegetation classes (FF, SF, SA, SG) as well, principally from 2002 to 2008 (b).

are on one side of the uniform intensity, i.e. they are consistently either smaller or larger than the uniform rate over the entire period.

$$S_t = \frac{\sum_{j=1}^J [(\sum_{i=1}^J C_{tij}) - C_{tjj}]}{(Y_{t+1} - Y_t) (\sum_{j=1}^J \sum_{i=1}^J C_{tij})} \times 100\% \tag{1}$$

$$U = \frac{\sum_{t=1}^{T-1} \{ (Y_{t+1} - Y_t) \sum_{j=1}^J [(\sum_{i=1}^J C_{tij}) - C_{tjj}] \}}{(Y_T - Y_1) \sum_{t=1}^{T-1} [(Y_{t+1} - Y_t) (\sum_{j=1}^J \sum_{i=1}^J C_{tij})]} \times 100\% \tag{2}$$

$$G_{tj} = \frac{[(\sum_{i=1}^J C_{tij}) - C_{tjj}] / (Y_{t+1} - Y_t)}{\sum_{i=1}^J C_{tij}} \times 100\% \tag{3}$$

$$L_{ti} = \frac{[(\sum_{j=1}^J C_{tij}) - C_{tii}] / (Y_{t+1} - Y_t)}{\sum_{j=1}^J C_{tij}} \times 100\% \tag{4}$$

$$R_{tin} = \frac{C_{tin} / (Y_{t+1} - Y_t)}{\sum_{j=1}^J C_{tij}} \times 100\% \tag{5}$$

$$W_{tn} = \frac{[(\sum_{i=1}^J C_{tin}) - C_{tmm}] / (Y_{t+1} - Y_t)}{\sum_{j=1}^J [(\sum_{i=1}^J C_{tij}) - C_{tnj}]} \times 100\% \tag{6}$$

$$Q_{tmj} = \frac{C_{tmj} / (Y_{t+1} - Y_t)}{\sum_{i=1}^J C_{tij}} \times 100\% \tag{7}$$

$$V_{tm} = \frac{[(\sum_{j=1}^J C_{tmj}) - C_{tmm}] / (Y_{t+1} - Y_t)}{\sum_{i=1}^J [(\sum_{j=1}^J C_{tij}) - C_{tim}]} \times 100\% \tag{8}$$

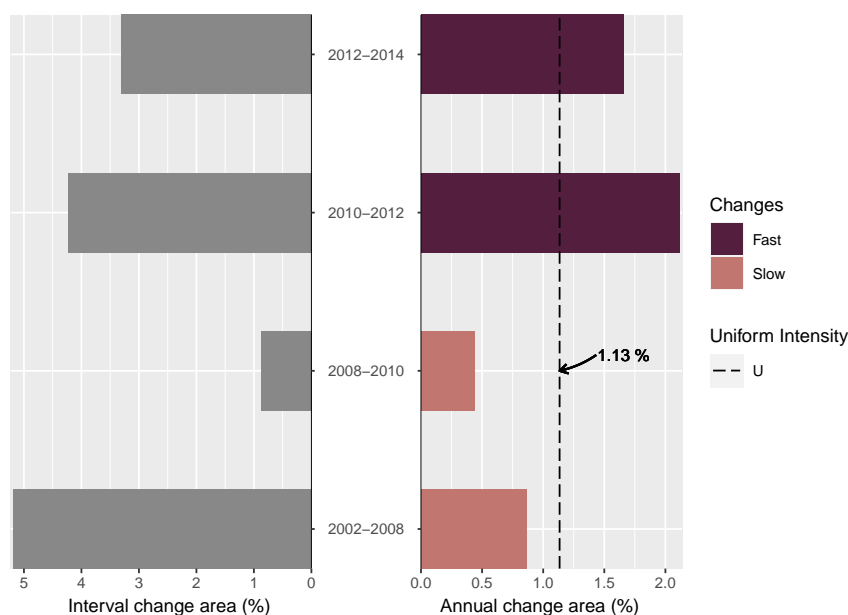
## Graphs (output of the intensity analysis)

Visualizations of the IA results are obtained with the `plot(intensity-object)` method. For more details on the function arguments, please see the documentation of the `plot()` method implemented.

### Interval Level

The IA interval level is a measure of the overall rate of LUC changes over consecutive time intervals in the series. The plot (Fig. 7) has two sides: on the left the percentage of change during the corresponding time step, and on the right the percentage of change per year showing the reference line of the Uniform rate.

```
plot(testSL$interval_lvl,
     labels = c(leftlabel = "Interval change area (%)",
                rightlabel = "Annual change area (%)"),
     marginplot = c(-8, 0), leg_curv = c(x = .3, y = .1),
     color_bar = c(fast = "#541f3f", slow = "#c1766f", area = "#888888"))
```



**Figure 7:** IA Interval Level plot of LUCC for the time intervals of the sample data set. Left-side bars represent LUCC during the time interval and right-side bars show the annualized rates. Strongest annual LUCC occurred between 2010 and 2012.

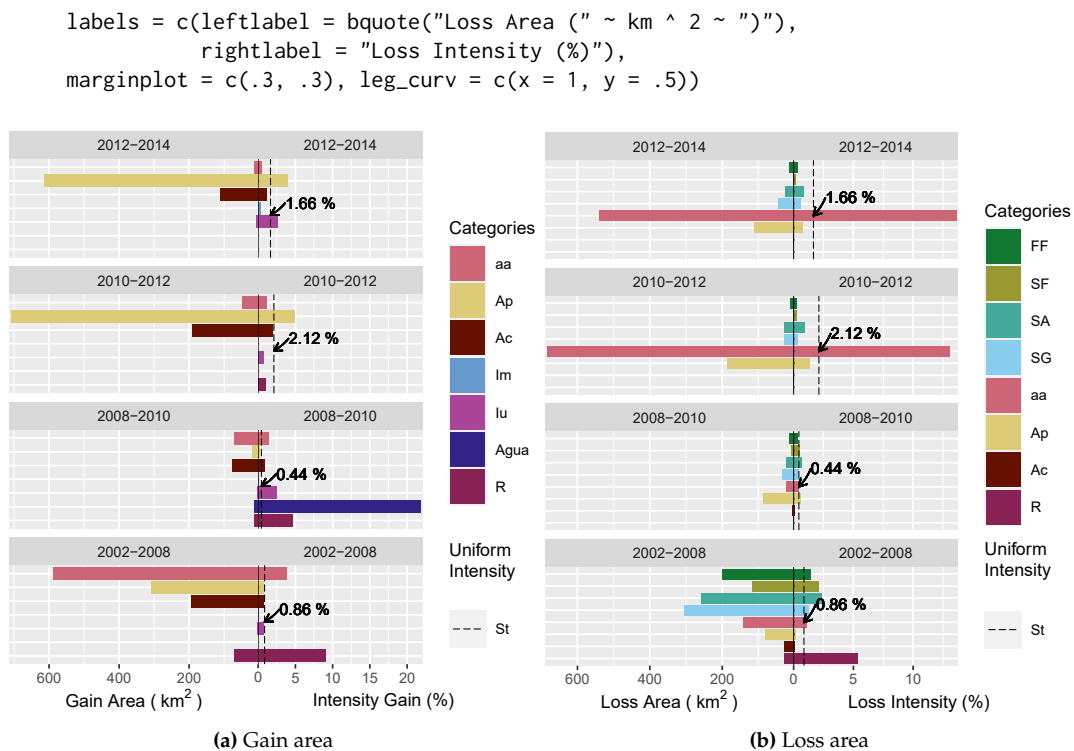
In the São Lourenço river basin, the interval level plot shows that LUC change has accelerated over the last decade, with a peak between 2010 and 2012, when the actual rate of change was almost double the Uniform rate. This recent acceleration in LUC in the Cerrado biome has been interpreted partially as a spillover effect of conservation efforts in the Amazon basin (Dou et al., 2018).

### Category Level (Gain and Loss Area)

After the analysis of LUC change intensity independently from LUC categories, the category level allows further examination into which land categories are relatively dormant versus active in a given time interval and whether this pattern is stable across time intervals (Aldwaik and Pontius, 2012).

```
# Gain area
plot(testSL$category_lvlGain,
     labels = c(leftlabel = bquote("Gain Area (" ~ km ^ 2 ~ ")"),
                rightlabel = "Intensity Gain (%)"),
     marginplot = c(.3, .3), leg_curv = c(x = 1, y = .5))
```

```
# Loss area
plot(testSL$category_lvlLoss,
```



**Figure 8:** IA Category level (a) gain and (b) loss plots. Cattle farming (Ap), for example, had the highest areal gain of all categories (2010 – 2012), whereas water bodies had a very high intensity gain between 2008 and 2010 due to the implantation of a large hydropower plant reservoir.

To facilitate legibility, we chose to split the category level plots (Fig. 8) into area gains (Fig. 8a) and losses (Fig. 8b). Area gains in land categories (R) and (aa) were more intense over the first six-year period (2002-2008) than they were at any other subsequent time point. The very intense area gain in water bodies during the second time interval (2008-2010) corresponds to the São Lourenço hydropower plant reservoir being filled. In the third and fourth intervals, the expansion of pasture areas (Ap) was more intense than during previous time steps. In parallel, the land category (aa) is in sharp decline.

### Transition level (gain of the category n “Ap” and loss of the category m “SG”)

In the transition level, the analysis focuses on the intensity of gain of a particular category n from all the individual categories in the landscape and/or on the intensity of loss of a particular category m to all the individual categories in the landscape for each time interval.

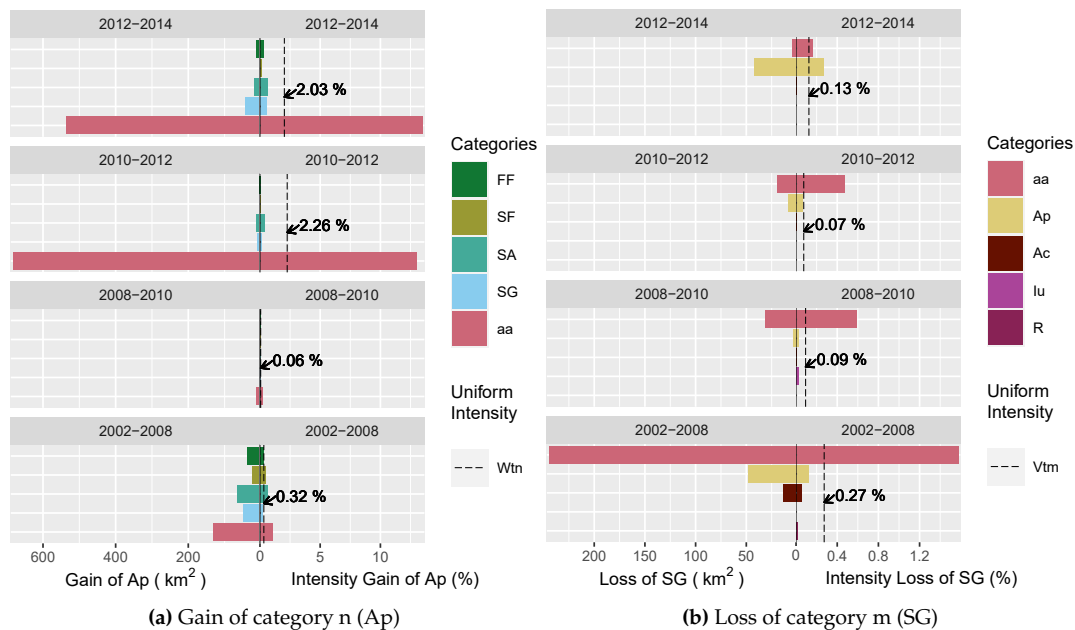
```

# Gain of the category `n` "Ap"
plot(testSL$transition_lvlGain_n,
     labels = c(leftlabel = bquote("Gain of Ap (" ~ km^2 ~ ")"),
               rightlabel = "Intensity Gain of Ap (%)"),
     marginplot = c(.3, .3),
     leg_curv = c(x = 1, y = .2))

# Loss of the category `m` "SG"
plot(testSL$transition_lvlLoss_m,
     labels = c(leftlabel = bquote("Loss of SG (" ~ km^2 ~ ")"),
               rightlabel = "Intensity Loss of SG (%)"),
     marginplot = c(.3, .3),
     leg_curv = c(x = .1, y = .4))

```

The area gains in the (Ap) category and losses in the SG category of the SaoLourencoBasin dataset are used here as an example. Results in (Fig. 9a) show that areas in (Ap) were principally gained from the aa category, and that this transition was particularly intense in the third and fourth time periods. Meanwhile, area losses in the (SG) savannah formation persisted over the entire span of the time series (Fig. 9b). However, there was a change in the land uses these areas were lost to: during the first three intervals, (SG) was lost principally to (aa), while in the last time interval, the loss was more intensely due directly to the (Ap) category.



**Figure 9:** IA Transition level plots of cattle farming (Ap) gains and Gramineous savanna (SG) losses. Left-side bars represent the areal gain/loss of a LUC category and right-side bars show the intensity of gain/loss. The SG category had the highest and most intense loss to anthropogenized vegetation (aa) category between 2002 and 2008.

## Conclusions and further research

In response to added and refined temporal, spatial and thematic dimensions increasing the volume of LUC data, the **OpenLand** package provides a comprehensive and integrated suite for the exploratory analysis of LUC changes. It offers seamless processing workflows beginning with time series consistency checking, data extraction, analysis and plotting of commonly used LUC metrics, as well as an implementation of Intensity Analysis, a state-of-the-art top-down hierarchical methodological framework to quantify the intensity of LUC changes. Regardless of the complexity of an LUC time series, all transitions and metrics are automatically extracted, quantified and stored as objects, without any need for further tabular data manipulation for analysis. Visualization tools create pre-formatted print-ready plots, which can be easily modified through function arguments.

Aldwaik and Pontius (2013) presented an extension of IA, which allows us to consider hypothetical classification errors in input LUC maps as part of the comparison between observed and uniform intensities in IA. The implementation of their method in **OpenLand** could further help users to assess the implications of errors on the strength of the evidence in the outputs of their Intensity Analysis and therefore, improve their understanding of LUC change processes.

## Acknowledgements

Reginal Exavier is supported by the Brazilian Funding Agency CAPES (Coordination for the Improvement of Higher Level Personnel) through a Master studentship (2018 - 2020) at the Department of Geography of the Federal University of Mato Grosso. The research received financial support from the Critical Ecosystem Partnership Fund - CEPF (Grant Agreement CEPF-104254) and FAPEMAT (Grant Agreement 219228/2015).

## Bibliography

- F. O. Akinyemi, R. G. Pontius, and A. K. Braimoh. Land change dynamics: insights from Intensity Analysis applied to an African emerging city. *J. Spat. Sci.*, 62(1):69–83, 2017. ISSN 14498596. doi: 10.1080/14498596.2016.1196624. URL <http://dx.doi.org/10.1080/14498596.2016.1196624>. [p1]
- S. Z. Aldwaik and R. G. Pontius. Intensity analysis to unify measurements of size and stationarity of land changes by interval, category, and transition. *Landsc. Urban Plan.*, 106(1):103–114,

2012. ISSN 01692046. doi: 10.1016/j.landurbplan.2012.02.010. URL <http://dx.doi.org/10.1016/j.landurbplan.2012.02.010>. [p1, 2, 8, 9, 11]
- S. Z. Aldwaik and R. G. Pontius. Map errors that could account for deviations from a uniform intensity of land change. *Int. J. Geogr. Inf. Sci.*, 27(9):1717–1739, 2013. ISSN 13658816. doi: 10.1080/13658816.2013.787618. URL <http://dx.doi.org/10.1080/13658816.2013.787618>. [p2, 13]
- J. J. Allaire, C. Gandrud, K. Russell, and C. J. Yetman. networkD3: D3 JavaScript Network Graphs from R, 2017. URL <https://cran.r-project.org/package=networkD3>. [p2]
- B. Auguie. *gridExtra: Miscellaneous Functions for "Grid" Graphics*, 2017. URL <https://cran.r-project.org/package=gridExtra>. [p2]
- A. P. Ballantyne, R. Andres, R. Houghton, B. D. Stocker, R. Wanninkhof, W. Anderegg, L. A. Cooper, M. DeGrandpre, P. P. Tans, J. B. Miller, C. Alden, and J. W. White. Audit of the global carbon budget: Estimate errors and their impact on uptake uncertainty. *Biogeosciences*, 12(8):2565–2584, 2015. ISSN 17264189. doi: 10.5194/bg-12-2565-2015. [p1]
- V. Brovkin, L. Boysen, V. K. Arora, J. P. Boisier, P. Cadule, L. Chini, M. Claussen, P. Friedlingstein, V. Gayler, B. J. Van den hurk, G. C. Hurtt, C. D. Jones, E. Kato, N. De noblet ducoudre, F. Pacifico, J. Pongratz, and M. Weiss. Effect of anthropogenic land-use and land-cover changes on climate and land carbon storage in CMIP5 projections for the twenty-first century. *J. Clim.*, 26(18):6859–6881, 2013. ISSN 08948755. doi: 10.1175/JCLI-D-12-00623.1. [p1]
- J. Chambers. *Software for Data Analysis*. Statistics and Computing. Springer New York, New York, NY, 2008. ISBN 978-0-387-75935-7. doi: 10.1007/978-0-387-75936-4. URL <http://link.springer.com/10.1007/978-0-387-75936-4>. [p9]
- G. Chaudhuri and N. B. Mishra. Spatio-temporal dynamics of land cover and land surface temperature in Ganges-Brahmaputra delta: A comparative analysis between India and Bangladesh. *Appl. Geogr.*, 68:68–83, mar 2016. ISSN 01436228. doi: 10.1016/j.apgeog.2016.01.002. URL <https://linkinghub.elsevier.com/retrieve/pii/S0143622816300029>. [p1]
- C. H. W. de Souza, W. R. Cervi, J. C. Brown, J. V. Rocha, and R. A. C. Lamparelli. Mapping and evaluating sugarcane expansion in Brazil’s savanna using MODIS and intensity analysis: a case-study from the state of Tocantins. *J. Land Use Sci.*, 12(6):457–476, nov 2017. ISSN 1747-423X. doi: 10.1080/1747423X.2017.1404647. URL <https://www.tandfonline.com/doi/full/10.1080/1747423X.2017.1404647>. [p1]
- Y. Dou, R. F. B. da Silva, H. Yang, and J. Liu. Spillover effect offsets the conservation effort in the Amazon. *J. Geogr. Sci.*, 28(11):1715–1732, nov 2018. ISSN 1009-637X. doi: 10.1007/s11442-018-1539-0. URL <http://link.springer.com/10.1007/s11442-018-1539-0>. [p11]
- R. Exavier and P. Zeilhofer. *OpenLand: Quantitative Analysis and Visualization of LUCC*, 2020. URL <https://CRAN.R-project.org/package=OpenLand>. R package version 1.0.1. [p2]
- Z. Gu, L. Gu, R. Eils, M. Schlesner, and B. Brors. circlize implements and enhances circular visualization in R. *Bioinformatics*, 30(19):2811–2812, 2014. [p2]
- M. Herold, J. S. Latham, A. Di Gregorio, and C. C. Schmullius. Evolving standards in land cover characterization. *J. Land Use Sci.*, 1(2-4):157–168, dec 2006. ISSN 1747-423X. doi: 10.1080/17474230601079316. URL <http://www.tandfonline.com/doi/abs/10.1080/17474230601079316>. [p1]
- R. J. Hijmans. *raster: Geographic Data Analysis and Modeling*, 2019. URL <https://cran.r-project.org/package=raster>. [p2]
- B. Huang, J. Huang, R. Gilmore Pontius, and Z. Tu. Comparison of Intensity Analysis and the land use dynamic degrees to measure land changes outside versus inside the coastal zone of Longhai, China. *Ecol. Indic.*, 89:336–347, jun 2018. ISSN 1470160X. doi: 10.1016/j.ecolind.2017.12.057. URL <https://linkinghub.elsevier.com/retrieve/pii/S1470160X1730849X>. [p1]
- G. C. Hurtt, L. P. Chini, S. Frolking, R. A. Betts, J. Feddema, G. Fischer, J. P. Fisk, K. Hibbard, R. A. Houghton, A. Janetos, C. D. Jones, G. Kindermann, T. Kinoshita, K. Klein Goldewijk, K. Riahi, E. Shevliakova, S. Smith, E. Stehfest, A. Thomson, P. Thornton, D. P. van Vuuren, and Y. P. Wang. Harmonization of land-use scenarios for the period 1500–2100: 600 years of global gridded annual land-use transitions, wood harvest, and resulting secondary lands. *Clim. Change*, 109(1-2):117–161, nov 2011. ISSN 0165-0009. doi: 10.1007/s10584-011-0153-2. URL <http://link.springer.com/10.1007/s10584-011-0153-2>. [p1]



- Instituto SOS Pantanal and WWF-Brasil. Monitoramento das alterações da cobertura vegetal e uso do Solo na Bacia do Alto Paraguai – Porção Brasileira – Período de Análise: 2012 a 2014. Technical report, Brasília, 2015. URL [https://d3nehc6yl9qzo4.cloudfront.net/downloads/publicacao\\_bap\\_relatorio\\_2012\\_2014\\_web.pdf](https://d3nehc6yl9qzo4.cloudfront.net/downloads/publicacao_bap_relatorio_2012_2014_web.pdf). [p3, 4, 5]
- E. F. Lambin. Modelling and monitoring land-cover change processes in tropical regions. *Prog. Phys. Geogr. Earth Environ.*, 21(3):375–393, sep 1997. ISSN 0309-1333. doi: 10.1177/030913339702100303. URL <http://journals.sagepub.com/doi/10.1177/030913339702100303>. [p1]
- E. F. Lambin and H. Geist, editors. *Land-Use and Land-Cover Change*. Global Change - The IGBP Series. Springer Berlin Heidelberg, Berlin, Heidelberg, 2006. ISBN 978-3-540-32201-6. doi: 10.1007/3-540-32202-7. URL <http://link.springer.com/10.1007/3-540-32202-7>. [p1]
- M. R. d. S. Melo, J. V. Rocha, V. D. Manabe, and R. A. C. Lamparelli. Intensity of land use changes in a sugarcane expansion region, Brazil. *J. Land Use Sci.*, 00(00):1–16, 2018. ISSN 17474248. doi: 10.1080/1747423X.2018.1499829. URL <https://doi.org/10.1080/1747423X.2018.1499829>. [p1]
- A. Morin, J. Urban, P. D. Adams, I. Foster, A. Sali, D. Baker, and P. Sliz. Shining Light into Black Boxes. *Science (80- )*, 336(6078):159–160, apr 2012. ISSN 0036-8075. doi: 10.1126/science.1218263. URL <http://www.sciencemag.org/cgi/doi/10.1126/science.1218263>. [p2]
- S. Moulds, W. Buytaert, and A. Mijic. An open and extensible framework for spatially explicit land use change modelling: the lulcc R package. *Geosci. Model Dev.*, 8(10):3215–3229, oct 2015. ISSN 1991-9603. doi: 10.5194/gmd-8-3215-2015. URL <https://www.geosci-model-dev.net/8/3215/2015/>. [p2]
- D. Müller and D. K. Munroe. Current and future challenges in land-use science. *J. Land Use Sci.*, 9(2):133–142, apr 2014. ISSN 1747-423X. doi: 10.1080/1747423X.2014.883731. URL <http://www.tandfonline.com/doi/abs/10.1080/1747423X.2014.883731>. [p1]
- H. Mwangi, P. Lariu, S. Julich, S. Patil, M. McDonald, and K.-H. Feger. Characterizing the Intensity and Dynamics of Land-Use Change in the Mara River Basin, East Africa. *Forests*, 9(1):8, dec 2017. ISSN 1999-4907. doi: 10.3390/f9010008. URL <http://www.mdpi.com/1999-4907/9/1/8>. [p1]
- G. Nelson, M. Rosegrant, A. Palazzo, and I. G. ... *Food security and climate change: Challenges to 2050 and beyond*. International Food Policy Research Institute, 2010. doi: 10.2499/9780896291874. URL <https://cgspace.cgiar.org/handle/10568/33400http://ebrary.ifpri.org/cdm/ref/collection/p15738coll12/id/6778>. [p1]
- R. D. Peng. Reproducible Research in Computational Science. *Science (80- )*, 334(6060):1226–1227, dec 2011. ISSN 0036-8075. doi: 10.1126/science.1213847. URL <http://www.sciencemag.org/cgi/doi/10.1126/science.1213847>. [p2]
- R. G. Pontius Jr. and S. Khallaghi. intensity.analysis: Intensity of Change for Comparing Categorical Maps from Sequential Intervals, 2019. URL <https://cran.r-project.org/package=intensity.analysis>. [p2]
- R. Prestele, P. Alexander, M. D. A. Rounsevell, A. Arneith, K. Calvin, J. Doelman, D. A. Eitelberg, K. Engström, S. Fujimori, T. Hasegawa, P. Havlik, F. Humpenöder, A. K. Jain, T. Krisztin, P. Kyle, P. Meiyappan, A. Popp, R. D. Sands, R. Schaldach, J. Schüngel, E. Stehfest, A. Tabeau, H. Van Meijl, J. Van Vliet, and P. H. Verburg. Hotspots of uncertainty in land-use and land-cover change projections: a global-scale model comparison. *Glob. Chang. Biol.*, 22(12):3967–3983, dec 2016. ISSN 13541013. doi: 10.1111/gcb.13337. URL <http://doi.wiley.com/10.1111/gcb.13337>. [p1]
- Project MapBiomass. Collection 3.1 of Brazilian Land Cover & Use Map Series, 2019. URL <https://mapbiomas.org/>. [p5]
- I. M. D. Rosa, S. E. Ahmed, and R. M. Ewers. The transparency, reliability and utility of tropical rainforest land-use and land-cover change models. *Glob. Chang. Biol.*, 20(6):1707–1722, jun 2014. ISSN 13541013. doi: 10.1111/gcb.12523. URL <http://doi.wiley.com/10.1111/gcb.12523>. [p2]
- RStudio Team. RStudio: Integrated Development Environment for R, 2016. URL <http://www.rstudio.com/>. [p2]
- X. P. Song, M. C. Hansen, S. V. Stehman, P. V. Potapov, A. Tyukavina, E. F. Vermote, and J. R. Townshend. Global land change from 1982 to 2016. *Nature*, 560(7720):639–643, 2018. ISSN 14764687. doi: 10.1038/s41586-018-0411-9. URL <http://dx.doi.org/10.1038/s41586-018-0411-9>. [p1]

- S. Subasinghe, R. Estoque, and Y. Murayama. Spatiotemporal Analysis of Urban Growth Using GIS and Remote Sensing: A Case Study of the Colombo Metropolitan Area, Sri Lanka. *ISPRS Int. J. Geo-Information*, 5(11):197, oct 2016. ISSN 2220-9964. doi: 10.3390/ijgi5110197. URL <http://www.mdpi.com/2220-9964/5/11/197>. [p1]
- O. G. Varga, R. G. Pontius, S. K. Singh, and S. Szabó. Intensity Analysis and the Figure of Merit's components for assessment of a Cellular Automata – Markov simulation model. *Ecol. Indic.*, 101(January):933–942, jun 2019. ISSN 1470160X. doi: 10.1016/j.ecolind.2019.01.057. URL <https://linkinghub.elsevier.com/retrieve/pii/S1470160X19300743><https://doi.org/10.1016/j.ecolind.2019.01.057>. [p1]
- P. H. Verburg, N. Crossman, E. C. Ellis, A. Heinimann, P. Hostert, O. Mertz, H. Nagendra, T. Sikor, K. H. Erb, N. Golubiewski, R. Grau, M. Grove, S. Konaté, P. Meyfroidt, D. C. Parker, R. R. Chowdhury, H. Shibata, A. Thomson, and L. Zhen. Land system science and sustainable development of the earth system: A global land project perspective. *Anthropocene*, 12(November):29–41, 2015. ISSN 22133054. doi: 10.1016/j.ancene.2015.09.004. URL <http://dx.doi.org/10.1016/j.ancene.2015.09.004>. [p1]
- H. Wickham. *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York, 2016. ISBN 978-3-319-24277-4. URL <http://ggplot2.org>. [p2]
- H. Wickham and L. Henry. *tidyr: Tidy Messy Data*, 2019. URL <https://cran.r-project.org/package=tidyr>. [p2]
- H. Wickham, R. François, L. Henry, and K. Müller. *dplyr: A Grammar of Data Manipulation*, 2019. URL <https://cran.r-project.org/package=dplyr>. [p2]
- H. Yang, S. Li, J. Chen, X. Zhang, and S. Xu. The Standardization and Harmonization of Land Cover Classification Systems towards Harmonized Datasets: A Review. *ISPRS Int. J. Geo-Information*, 6(5):154, may 2017. ISSN 2220-9964. doi: 10.3390/ijgi6050154. URL <http://www.mdpi.com/2220-9964/6/5/154>. [p1]
- J. Yu, W. Li, and C. Zhang. A framework of experimental transiogram modelling for Markov chain geostatistical simulation of landscape categories. *Comput. Environ. Urban Syst.*, 73:16–26, jan 2019. ISSN 01989715. doi: 10.1016/j.compenvurbsys.2018.07.007. URL <https://linkinghub.elsevier.com/retrieve/pii/S0198971517305653>. [p1]

Reginal Exavier

Department of Geography

Federal University of Mato Grosso

Avenida Fernando Corrêa da Costa, 2367 – Boa Esperança, Cuiabá – MT, 78060-900

ORCID: 0000-0002-5237-523X

[reginalexavier@rocketmail.com](mailto:reginalexavier@rocketmail.com)

Peter Zeilhofer

Department of Geography

Federal University of Mato Grosso

Avenida Fernando Corrêa da Costa, 2367 – Boa Esperança, Cuiabá – MT, 78060-900

[zeilhoferpeter@gmail.com](mailto:zeilhoferpeter@gmail.com)