One Ecosystem 4: e28720
doi: 10.3897/oneeco.4.e28720

Research Article

# Uncertainty in ecosystem services maps: the case of carbon stocks in the Brazilian Amazon forest using regression analysis

Solen Le Clec'h[‡], Simon Dufour[§], Janic Bucheli[‡], Michel Grimaldi[|], Robert Huber[‡], Izildinha de Souza Miranda[¶], Danielle Mitja[#], Luiz Gonzaga Silva Costa[¶], Johan Oszwald[§]

‡ AECP group, ETH Zurich, Zürich, Switzerland
§ LETG Rennes, Rennes, France
| IRD, Bondy, France
¶ Universidade Federal Rural da Amazônia, Belem, Brazil
# IRD, Montpellier, France

Corresponding author: Solen Le Clec'h (solenle@ethz.ch)

Academic editor: Benjamin Burkhard

Received: 30 Jul 2018 | Accepted: 03 Jan 2019 | Published: 31 Jan 2019

Citation: Le Clec'h S, Dufour S, Bucheli J, Grimaldi M, Huber R, Miranda I, Mitja D, Silva Costa L, Oszwald J (2019) Uncertainty in ecosystem services maps: the case of carbon stocks in the Brazilian Amazon forest using regression analysis. One Ecosystem 4: e28720. https://doi.org/10.3897/oneeco.4.e28720

## Abstract

Ecosystem Service (ES) mapping has become a key tool in scientific assessments of human-nature interactions and is being increasingly used in environmental planning and policy-making. However, the associated epistemic uncertainty underlying these maps often is not systematically considered. This paper proposes a basic procedure to present areas with lower statistical reliability in a map of an ES indicator, the vegetation carbon stock, when extrapolating field data to larger case study regions. To illustrate our approach, we use regression analyses to model the spatial distribution of vegetation carbon stock in the Brazilian Amazon forest in the State of Pará. In our analysis, we used field data measurements for the carbon stock in three study sites as the response variable and various land characteristics derived from remote sensing as explanatory variables for the ES indicator. We performed regression methods to map the carbon stocks and calculated three indicators of reliability: RMSE-Root-mean-square-error, $R^2$-coefficient of determination - from an out-of-sample validation and prediction intervals. We obtained a map of carbon stocks and made explicit its associated uncertainty using a general indicator of reliability and a map presenting the areas where our prediction is the most uncertain. Finally, we highlighted the role of environmental factors on the range of uncertainty. The results have two implications. (1) Mapping prediction interval indicates areas where the map's reliability is the highest. This information increases the usefulness of ES maps in environmental planning and governance. (2) In the case of the studied indicator, the reliability of our prediction is very dependent on land cover type, on the site location and its biophysical, socioeconomic and political characteristics. A better understanding of the relationship between carbon stock and land-use classes would increase the reliability of the maps. Results of our analysis help to direct future research and fieldwork and to prevent decision-making based on unreliable maps.

## Keywords

Deforestation; ecosystem services; prediction intervals; reliability; statistical modelling; variability

## Introduction

Ecosystem services (ES) have progressively become an important concept in environmental planning and policy-making to bridge the science – policy interface in the management of ecosystems (Braat and de Groot 2012, Perrings et al. 2011, Groot et al. 2010). Mapping indicators of ES is one prominent approach to improve spatially explicit decision-making and land management (Burkhard et al. 2012). This ecosystem service approach can help policy-makers to target strategic areas, formulate new policies and/or evaluate impacts of previous policies (Burkhard et al. 2013, Maes et al. 2012, Martinez-Harms et al. 2015). ES maps are popular outreach and data visualisation products. They have profited from the increased availability and applications of tools such as GIS or remote sensing that helps to increase their production and distribution (Palsky 2013). Despite its popularity, mapping ES has its limitations (Eigenbrod et al. 2010van Oudenhoven et al. 2018). Specifically, spatial information used in mapping is rarely, if ever, completely accurate or verified (Heuvelink and Burrough 2002; Devendran and Lakshmanan 2014) and that spatial assessments are often based on coarse data, which decreases the confidence in the spatial products (Andrew et al. 2015; Hamel and Bryant 2017).

The importance of uncertainty in the use and analysis of spatial data has long been recognised in land-use and landcover change mapping (e.g. Dendoncker et al. 2008, Lavorel et al. 2017, Verburg et al. 2013, Verburg et al. 2011, Kuemmerle et al. 2013) and is increasingly acknowledged in the ES mapping community (Jacobs et al. 2017). The representation of uncertainties, related to missing or incomplete data, aggregation error, functional knowledge gaps or normative and value-laden indicators in these types of maps is key for the map user (Crossman et al. 2013, Jacobs et al. 2013). In principle, information on uncertainties should allow a critical look at the map and be an integral part of the decision-making process for which maps are used. As a matter of fact, there is an increasing number of scientific studies using ES maps addressing uncertainty or reliability (e.g. Bagstad et al. 2018, Brunner et al. 2017, Johnson et al. 2012, Boithias et al. 2016, Grêt-Regamey et al. 2013, Schulp et al. 2014). However, uncertainty in the valuation process still remains an exception and has not explicitly been accounted for in many of the studies.

Uncertainty is complex and there are many definitions or typologies of uncertainty in ES analyses (Hamel and Bryant 2017) and therefore there are many ways to address the issue. Most of the studies focus on the uncertainty stemming from the ecosystem complexity or on the social-ecological uncertainties from supply and demand in ES models. More recently, studies have highlighted the effects of the input data on the reliabilility of ES assessments (Kangas et al. 2018). There is, however, much less emphasis on the epistemic uncertainty in ES mapping due to the modelling process. In the field of ES mapping, there is specifically a lack in the representation of errors inherent to the extrapolation of point-based measurements to produce empirical-based ES maps. Mapping an ES indicator from point field measurements implies several causes of epistemic uncertainties, inherent in each step from measurement to extrapolation. For instance, because of the mismatch between the field sample, its GPS location information and pixel size, there are errors generated, as field and remote sensing data are linked (Réjou-Méchain et al. 2015) and there is also uncertainty generated by the application of statistical methods, i.e. related to the extrapolation process itself. A comprehensive analysis and transparent documentation of the latter are seldom provided in ES mapping approaches.

The objective of this research is to describe a simple approach to assess and spatially represent uncertainties associated with the extrapolation of measured field data using regression analysis to map an ES indicator (vegetation carbon stock). Our results express the degree of certainty that we have about our ES map and the confidence that policy-makers can have in the map. To this end, we firstly used field data measurements for vegetation carbon stock in the Brazilian Amazon forest as a response variable and various land characteristics derived from remote sensing as potential explanatory factors of the ES indicator. We implemented the mapping techniques from earlier works (Le Clec'h et al. 2017) and evaluated the variability of ES supply in space resulting from our extrapolation process. Secondly, we calculated and mapped an indicator based on the prediction interval as a measure of the uncertainty stemming from the regression. Thirdly, we represented the variation of the prediction interval within the land-use classes and within the three sites, to identify sources of uncertainty within our dataset. While such estimation and analysis of the uncertainty are common in some fields, such as geostatistics, it is still under-considered in the field of ES mapping. Therefore, in the field of ES assessment and mapping, this study can be seen as a first step in a systematic assessment of uncertainties related to ES mapping in which point-based measurements are used to construct ES maps. Systematic assessments of uncertainties should critically question the confidence and usability of ES maps. This should be critical for discussing and determining the potential use(s) of the map and/or for supporting policy and planning processes. As some studies recommend the application of the principle of parsimony in ES mapping (e.g. Jacobs et al. 2017), we favoured a simple approach in this paper.

## Material and methods

### *Site description*

We estimated and mapped uncertainty related to ES maps using the case of the Brazilian Amazon forest, a global hotspot of ES provision (Fearnside 2017). More specifically, we focused on the State of Pará that spans 1.25 million km$^2$, which is partly covered by tropical forest. Extensive forests, high biodiversity and rapid deforestation qualify Pará as indicative of contexts in which ES surveys are urgently required. The rapid deforestation observed in Pará is mainly caused by timber harvesting and subsequent cattle grazing, which prevents forest restoration (Fearnside 2017, Godar et al. 2012), with 1,887 km$^2$ deforested in 2014 (INPE, Instituto Nacional de Pesquisas Espaciais 2014). While deforestation has been slowing, the area deforested in Pará in 2013 still accounted for almost half of the total deforestation in Brazil.

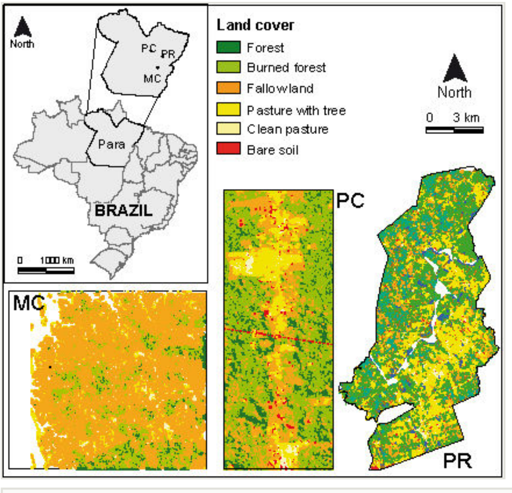


Figure 1.

Location of Pará State and the study sites of Maçaranduba (MC), Pacajá (PC) and Palmares II (PR). Coloured areas in the maps represent land use of the study sites in 2007.

Within Pará, we conducted field measures of ES within three local study sites that are representative of the regional variability in socio-economics, deforestation temporalities and ES change (Fig. 1, Oszwald et al. 2011). The first site, Maçaranduba belongs to the municipality of Nova Ipixuna. It covers 220 km$^2$ and has been deforested since the early 1970s. Although largely deforested, its remaining forest is relatively well preserved. The second site, named here Pacajá, belongs to the municipality of Pacajà. It covers 175 km$^2$ and is located on one of the 'fishbone' roads (*Traverssão* 338 South) of the Trans-Amazon highway. It has been deforested since the 1990s by agricultural colonisation that has resulted in the conversion of the landscape to pasture and annual and perennial crops. Despite the deforestation dynamics, forests still covered 63% of this site in 2007. The third site, Palmares II, is an *assentamento* (settlement) of 160 $km^2$ near the Carajas iron mine. Emblematic of the agrarian reform, this site was established following the conflict between Landless Movement workers, a *fazendeiro*, owner of a large farm, which was located on the site, the federal government and the Vale mining company. This conflict led to the fragmentation of the land and thus of the forest that only covered 23% of the site in 2007.

In a deforestation front context, because of the lack of law, public policies and environmental management, the study of ES is highly relevant. It allows us a better understanding of the impacts of deforestation and cultivation on the environment, in other words to evaluate the impacts of the past and current public policies (or their absence) on the ecosystems.

### *Data*

We used two different datasets to apply predictive statistical methodologies for the ES indicator: field data related to the ES indicator (response variable) and remote sensing data (explanatory variables – Table 1).

Table 1.

Description of the data used in this study.The data constitutes the input of the statistical model.

| | **Variable name** | **Source** | **Description** | **Unit/ range** |
|---|---|---|---|---|
| Response variable (ES indicator) | Vegetation carbon stock | Field measurement | Aboveground dry biomass of trees, bushes and herbaceous plants | Mg/ha |
| Explanatory variables (ES potential drivers) | Land cover | Landsat TM (30x30m) | Supervised maximum-likelihood classifications of six land-cover classes | 6 modalities |
| | Historical trajectory of land cover | | Classes of land-use trajectories (Oszwald et al. 2012), from a homogeneous forest structure (class 1) to an agricultural dynamic of extensive breeding (class 5) | 5 modalities |
| | NDVI | | Vegetation density (index) | -1;1 |
| | NDWI | | Water content into plants (index) | -1;1 |
| | Elevation | DEM Aster (30mx30m) | Elevation at every point | m |
| | Slope | | Altitude difference between two adjacent pixels | % |
| | Synthesis of topography | | Characterisation of the topographic context | 4 modalities |
| | Distance to water | | Buffers around the rivers | 5 modalities |
| | Site | | General location | 3 modalities |

### *Field data: carbon stocks*

According to the definition of ES proposed by Portela and Rademacher (2001), a way to map ES consists of defining spatial indicators of the biophysical processes that provide the services (Oszwald et al. 2014). We studied here an indicator of the service of climate regulation. The chosen indicator is the vegetation carbon stock whose analysis is relevant in a pioneer front context. Even though stocks are not *per se* an ES, we assumed this indicator to be relevant in this context, as their main driver of change is due to human activities. It was measured for 135 sampling points that were assessed during fieldworks in 2008 before the dry season. In each of the 27 farms (nine per site), five sampling points were equally spaced (around every 200 m) along a transect corresponding to the longest diagonal of the farm or to a north-south axis.

The carbon stock was estimated using a factor of 0.5 (Markewitz et al. 2004) in aboveground dry total biomass calculated from field data shown by Costa et al. (2012). At each sampling point, the dry biomass was calculated for four strata: upper, middle, lower (same strata where the vegetation cover was inventoried) and necromass. A plot of 10 m× 50 m (500 $m^2$) was established for the upper stratum inventory (individuals with DBH - diameter at breast height ≥ 10 cm), a subplot of 5 m× 50 m was established for the middle stratum (individuals with DBH < 10 cm and height ≥ 2.0 m) and finally 4 subplots of 1 m ×1 m were regularly distributed in the centre of the plot for the lower and necromass strata measured (individuals with 2.0 m > height ≥ 10 cm). In the upper and middle strata, dry biomass was estimated from allometric equations using diameter at breast height. In the primary forests, biomass was estimated according to Higuchi et al. (1998) using three classes of DBH: DBH > 20 cm, 5 cm < DBH < 20 cm and DBH < 5 cm. In secondary forest, biomass was calculated for Cecropia and all other trees, regardless of diameter, according to Nelson et al. (1999). In both forests, primary and secondary, biomass of lianas was estimated according to Gerwing (2002). The same equations were used in all plots to ensure a better comparison between them. Due to the low density of trees found in pastures and cultivation areas, the dry biomass of the upper and middle strata were calculated using the same formulas used in primary and secondary forests. In the lower stratum, the biomass was first calculated directly by fresh weight; thereafter, one sample was taken and also weighed; and finally, the sample was oven-dried at 70°C until a constant dry weight and the dry biomass of the sub-plot was calculated by cross-multiplication using the dry biomass of the sample. After the removal of plants from sub-plots, all remaining ground material (leaves, twigs, flowers, dead wood) was gathered to measure the necromass in a similar procedure to that used for measuring the biomass of the lower strata.

### *Remote sensing data*

We built and applied linear regression to extrapolate and map the ES indicator, using the plot-level measures and local high-resolution satellite imagery (Table 1). Since the vegetation carbon stock is strongly related to the biophysical characteristics of the landscape (Grimaldi et al. 2014), explanatory variables were chosen to characterise different aspects of the landscape. We used remote sensing data that provides information about vegetation cover and topography. They are derived from the processing of the 1986, 1996, 2001 and 2007 Landsat TM images (30 m x 30 m spatial resolution) carried out under ENVI and the processing of the Aster Digital Elevation Model (DEM - 30 m x 30 m spatial resolution). These data are known for the three study sites and they were extracted for the 135 sampling points.

We used remote-sensing data to characterise the land-cover of the three study sites for 2007. A Landsat TM image from the dry season (30 m spatial resolution) was used to build a supervised classification by maximum likelihood, to calculate two vegetation indexes (NDVI and NDWI) and to determine a historical trajectory of land-cover (Oszwald et al. 2012). The Landsat classifications characterise six land-cover classes (forest, burned forest, *juquira-capoeira* [fallow lands], grasslands with trees, clean grasslands and bare soils) for 2007. Training data, used for the supervised classification, were sampled during field campaigns using a GPS. Landsat images were radiometrically and geometrically corrected prior to classification to ensure comparability between the study sites. In addition, analysis and processing of land cover classification were also extended to all Landsat TM images of the dataset (from 1986 to 2007). Using the land-use maps obtained for the four dates, five classes of land-use trajectories were determined (Oszwald et al. 2012), ranging from a homogeneous forest structure (class 1) to an agricultural dynamic of extensive breeding (class 5). Landsat TM images were also used to calculate two vegetation indices giving information about vegetation density (NDVI) and water content into plants (NDWI).

Data also provided information about the elevation (in metres) at every point. Slopes synthesised the altitude difference between two adjacent pixels and are provided as a percentage. These two variables (elevation and slope) are quantitative and are treated as continuous raw data. The "topography" variable corresponds to a synthetic characterisation of the topographic context comprising four modalities: bottom of valleys, hilltops, zones of steep slopes and zones of low slopes. Finally, we deduced the hydrographic network from the DEM which was used to determine a distance to the rivers (0 to 100 m, 100 to 200 m, 200 to 300 m, 300 to 500 m and more than 500 m).

We also used a variable, named “Site” that corresponds to the identity of the study site to which each pixel belongs. This variable aims to (1) estimate the spatial auto-correlation of the sampling points and (2) take into account the inter-sites variability, due to homogeneous biophysical conditions and socioeconomic characteristics within each location.

#### *Statistical approach*

We aimed to evaluate and map the confidence we have in our prediction, as a way to represent the uncertainty related to the ES map. We based our approach on the implementation of statistical methodologies that link field (ES indicator) and remote sensing (explanatory) data. These statistical methodologies are used to (1) extrapolate field data using remote sensing to the three study sites. They also (2) give us information about the reliability of the resulting maps, through a general and a spatialised indicator (Fig. 2).

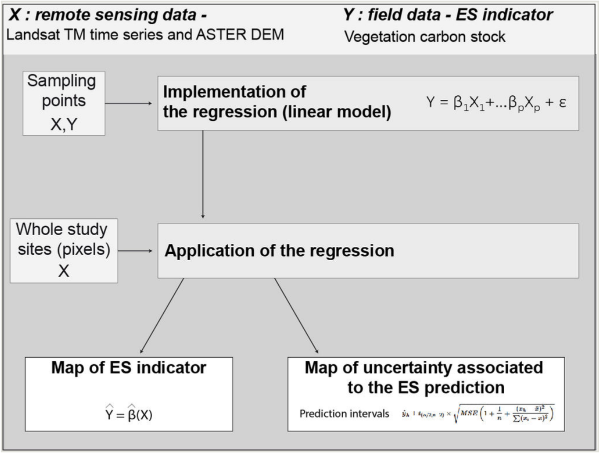


Figure 2.

General methodology used to map the ES indicator and the estimate the uncertainty associated to the map.

Regression methods represent one of the possible statistical approaches to map ES indicators from field and remote sensing data (Le Clec'h et al. 2017). Several regression methods can be computed to provide ES predictions based on both qualitative and quantitative inputs (Le Clec'h et al. 2013). Obtaining prediction intervals on the predictions can be quite challenging, in particular when no analytic expression of such intervals is available. For example, non-parametric methods, such as regression trees or random forests, may require bootstrap techniques to do so. However, linear modelling is interesting, for it allows the calculation of confidence intervals on predictions (Cornillon and Matzner-Lober 2011). In the following, we present results provided by linear regression methods.

We computed a linear model on the 135 sampling points. The model linked the ES indicator (carbon stocks - field data) and remote sensing data (related to the site, land cover and topography – Fig. 2). We manually performed a selection of variables, using Mallows' Cp Mallows 2000. The selection of variables aimed to (1) identify and select the best subset amongst the remote sensing variables and to (2) avoid overfitting. Thus, this selection allows us to obtain a model, which has good predictive properties. We called it, in this paper, “final model”, the model resulting from the selection of variables and that was used for the prediction of new ES values. This final model was based on two variables: the land-cover classification and the site (split into two classes: Palmares II and the two other sites, Table 2). The model can be considered as reliable ($R^2 = 0.75$).

We used the final model to extrapolate the vegetation carbon stocks to the whole study sites. To do so, the final model, trained on the 135 sampling points, was then applied to all the study sites to predict new carbon stocks, based on the land-cover data and the typology of the study sites (for more information, see Le Clec'h et al. 2017). Thanks to these predicted values, we obtained a map of the vegetation carbon stocks.

Table 2.

Outputs of the final linear model. Acronyms: $Site_{PR}$: Palmares II and Land $Cover_n$: class n of Land Cover (F: Forest; BF: Burned Forest; JC: Jucuira-Capoeira-Fallow lands; PL: Pasture with tree; CP: Clean Pasture and BS: Bare Soil).

| | | | | |
|---|---|---|---|---|
| **Linear model**: lm(formula = VCSt ~ $Site_{PR}$ + LC)<br>**Residuals**: | | | | |
| **Min** | **1Q** | **Median** | **3Q** | **Max** |
| -103.41 | -27.34 | -10.23 | 19.96 | 143.19 |
| **Coefficients**: | | | | |
| | **Estimate** | **Std. Error** | **t value** | **Pr(>\|t\|)** |
| **Intercept** | 161.57 | 13.474 | 11.99 | < 2e-16 *** |
| **$Site_{PR}$** | -37.77 | 9.82 | -3.85 | 0.000202 *** |
| **Land $Cover_{BF}$** | 60.64 | 16.51 | 3.67 | 0.000375 *** |
| **Land $Cover_{JC}$** | -89.37 | 15.33 | -5.83 | 5.86e-08 *** |
| **Land $Cover_{PT}$** | -143.46 | 16.78 | -8.55 | 9.07e-14 *** |
| **Land $Cover_{CP}$** | -143.06 | 16.65 | -8.592 | 7.20e-14 *** |
| **Land $Cover_{BS}$** | -131.70 | 17.58 | -7.49 | 1.99e-11 *** |
| Significativity codes: 0 '***'<br>Residual standard error: 47.99 on 108 degrees of freedom<br>Multiple $R^2$: 0.75, Adjusted $R^2$: 0.73<br>F-statistic: 52.62 on 6 and 108 DF, p-value: < 2.2e-16 | | | | |

In the second step, we used the same final model to estimate a simple indicator that gives information on the confidence we have in our prediction. In linear regression, one way to associate confidence/uncertainty to the prediction is the calculation of the prediction intervals on predictions (Cornillon and Matzner-Lober 2011). The prediction intervals of the prediction are a range that is likely to contain the mean response given specified settings of the predictors in our model. Prediction intervals assess the accuracy of the estimated ES indicator. Therefore, to characterise the confidence we have in our prediction, we evaluated and mapped an index based on the prediction intervals. This index allowed us to have spatial information about uncertainty of the estimated carbon stocks. Such information is useful to identify the areas where the prediction is very reliable or, on the contrary, where the predicted values should be considered carefully.

We proposed an index based on width of 95% prediction intervals. To do so, we calculated the difference between the upper and lower bounds of prediction confidence intervals. A high index characterised areas with high uncertainty around the prediction (up to 90 MgC/ha around the prediction). For each predicted value (pixel), we thus applied the final linear model to get (1) an estimated carbon stock and (2) the prediction interval around the predicted value. As we predicted an ES value for each pixel of the study sites, we also obtained an uncertainty index for each pixel. Thus, we can propose a spatial representation of the uncertainty. Finally, we represented the variations of the index within the two explanatory variables of the final model: the land cover and the site classifications.

## Results

### *Maps of carbon stocks and their global reliability*

The linear model based on the land-cover classification and the site classification was used to map vegetation carbon stock. The maps of vegetation carbon stock show the influence of land-cover changes on ES supply (Fig. 3).The highest values are located in forested areas, with the lowest values in deforested areas: farms and riversides in Maçaranduba, the main road in Pacajá and the southern part of Palmares II close to the city, influenced by the railway and the road. There are differences amongst study sites, because of local specificities such as differences in terms of biophysical conditions as human pressure, political and socio-economic factors that affect vegetation conditions and thus carbon stock. In Palmares II, the highest stock is lower than the one in the two other sites because the forests in Palmares II are highly degraded.

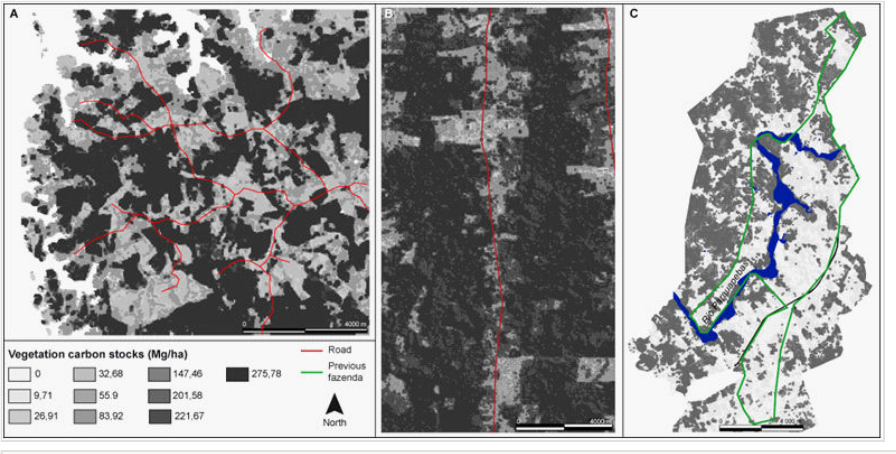


Figure 3.

Vegetation carbon stock in the three study sites: **A.** Maçaranduba; **B.** Pacajá; **C.** Palmares II.

### *Mapping uncertainty associated with the ES map*

We mapped the uncertainty index related to the application of a statistical method to predict values of an ES indicator for the vegetation carbon stock. The resulting map gives an overview of the areas where the map is and is not reliable. High prediction intervals express low confidence and can be associated with our inability to reliably predict carbon stocks, whereas high prediction intervals express our ability to estimate the vegetation carbon stocks with much greater certainty.

The index of uncertainty takes high values in highly anthropised areas (grazed landscapes and bare soils). Grazed grasslands have very heterogeneous profiles because they can be declined from bare soil to pasture with trees. Moreover, forest areas are associated with high carbon stock and quite high variability. Forests of the three study sites are at different stages of degradation and carbon stock is therefore relatively heterogeneous within these forests. Variability in transition areas can be explained by the nature of the class itself. Transition areas consist of secondary vegetation and fallow lands. This class is relatively homogeneous in Maçaranduba (Fig. 4) and in Pacajá, whereas it is more heterogeneous in Palmares II.

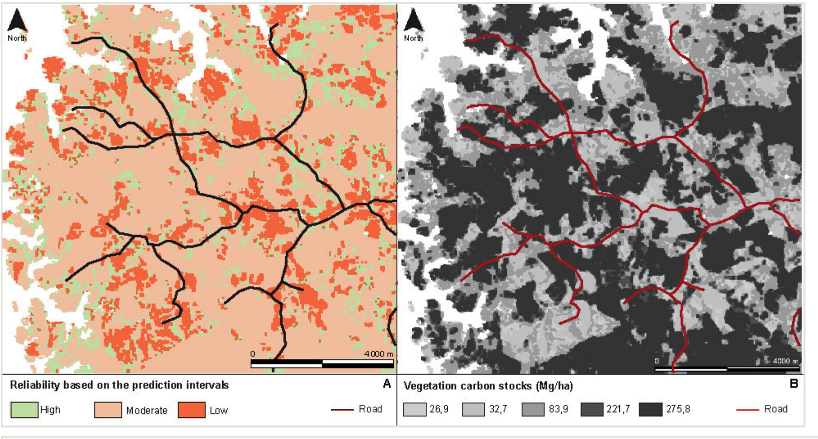


Figure 4.

Prediction intervals around the predicted vegetation carbon stock in Maçaranduba.

### *Role of land cover and location*

We analysed how the uncertainty index varies with the two explanatory variables of the final linear model: the land cover and the site classifications. This analysis helped us to better understand the level of uncertainty related to our assessments. To do so, we plotted the prediction intervals (1) within the six land-cover classes and (2) within the three study sites. Such analyses highlight the uncertainty related to the explanatory data that our knowledge of the study sites (socioeconomic, political and/or biophysical contexts and conditions) can explain. The range of the prediction intervals modelled, based on the final model (Table 2), varies amongst the land-cover type (Fig. 5A). Prediction intervals are low in degraded forests, even if this class is highly diversified, in terms of density, strata and type of vegetation, due to natural and human-based factors. Agricultural areas are very diverse, especially the bare soil class, which regroups two different realities that can be discriminated with the analysis of our satellite data. On the one hand, this class includes bare soils such as roads and tracks. On the other hand, it regroups as well annual crops that are actually bare soil without plant cover, during the dry season. The reliability of our prediction is also very dependent on the location, independently of the differences in land-cover between the sites (Fig. 5B). In Maçaraduba and Pacajá, our predictions are reliable with a low variability, whereas predicted values in Palmares are highly uncertain. For instance, in Maçaranduba, sharp contrasts in the landscape exist, between large areas of preserved forests and pastures. In Pacajá, despite the landscape gradient, there are also sharp contrasts in the landscape between well-conserved forests and already cultivated / pastured areas. In Palmares II, differentiation between land covers was less clear: pastures were abandoned then transformed into fallow lands and, sometimes, were transformed back into pastures. Consequently, for this site, continuity in land-cover types existed, leading to an absence of spatial contrasts that can influence our capacity to obtain reliable statistical ES predictions.

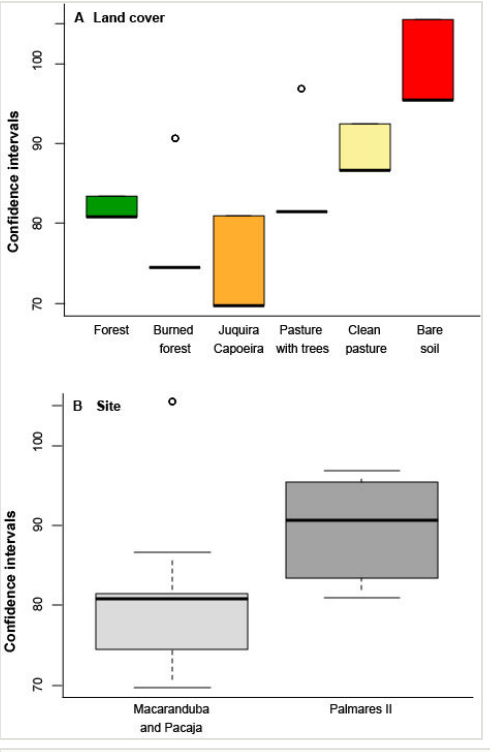


Figure 5.

Variation in the range of the prediction intervals due to (**A**) the land cover and (**B**) the geographical general location

## Discussion

A map is a generalisation or a schematisation of reality (Palsky 2013) and, accordingly, includes uncertainties or even errors. The issue of uncertainty can become particularly important when maps aim to support policies, for example, when they are used for land planning or to allocate budget to areas considered as ES hotspots (Bagstad et al. 2018, Eigenbrod et al. 2010, Palomo et al. 2018). In ES mapping processes, there is a plurality of uncertainties, inherent in each step of the modelling, for any methods that have been implemented. Scientific literature has increasingly acknowledged the need for considering uncertainties in the ES assessments (Bagstad et al. 2018, Grêt-Regamey et al. 2013, Schulp et al. 2014). There are, however, different types of uncertainties, from epistemic to scenarios inherent in mapping ES (Lavorel et al. 2017). Quantifiable and spatially explicit information about different types of uncertainty, as we presented here, help to qualify and critically assess ES maps and reveals the limits of the presented information, irrespective of the underlying methodological approach. However, the range of uncertainties produced in an ES assessment is much broader than the uncertainties induced by the statistical prediction (Hamel and Bryant 2017) and more systematic assessments are needed.

We proposed an exploratory approach to assess and communicate epistemic uncertainties related to ES mapping and valuation and thus specifically addressed a lack in the representation of errors inherent to the extrapolation of point-based measurements to produce empirical based ES maps. It is only one of the uncertainties produced in an ES assessment (Hamel and Bryant 2017), but it is also one that can be easily identified and communicated and/or reduced in order to create reliable outputs that can support policy-making or land management. We extrapolated measured field data using regression analysis to map carbon stocks for three study sites in the Brazilian Amazon and evaluated the variability of the ES provision across space. An indicator based on prediction intervals gave an overview of the spatially explicit reliability in the statistical procedure by enabling the identification of places where the prediction is reliable and where it has to be improved and cannot currently be used to base decisions.

Overall, the regression model provided a reliable map of carbon stocks that allowed us to highlight carbon hotspots and, at the same time, to identify areas where this prediction of carbon stocks is highly uncertain. The results showed that prediction intervals for bare soils and clean pastures had higher mean values but also a wider range of values and thus emerged as highly variable classes. The map's trustworthiness is thus partially a consequence of the uncertainty related to the explanatory data underlying these classes. From a scientific perspective, these findings have three implications. First, a better understanding of the drivers of ES provision could help to reduce the uncertainty stemming from the variability of the provision of ES per land-cover class. This aspect is very critical in the Amazon rainforest where the slowdown of deforestation rates leads to an increase in forest degradation. The consequences of forest degradation on ES provision in these areas are still largely unknown. Secondly, the land-cover classification should account for as many different classes as possible to reduce intra-class variability. This underlines the importance of studies on local scales with detailed information on land use/land cover to complement large scale maps with less accuracy in land-use classes (Foody 2015, Domac 2004). The availability of new satellite information such as PROBA-V might help to fill this gap given its relatively high spatial resolution and short processing time. Thirdly, despite this potential of new remote sensing information, data uncertainties highlight our limited ability to model some specific ES indicators. On the one hand, variations in the vegetation carbon stock, for instance, can be explained by modified properties of ecosystems such as land-cover changes. On the other hand, indicators related to soil or biodiversity are greatly influenced by inherent ecosystem properties such as biophysical processes that can rarely be inferred from remote-sensing data (Dominati et al. 2010). In addition, high uncertainty areas can vary from one ES to another and, for one specific ES, vary from one spatial scale to another (Le Clec'h et al. 2018). Thus, a systematic and transparent assessment of such uncertainties helps to reconsider and improve the methodological approach and the data used in the mapping process.

The use of prediction intervals to assess epistemic uncertainties in ES maps has also certain drawbacks. Not all methodologies allow a feedback between data and ES maps. The indicator proposed here, can only be applied in the case of certain statistical procedures. Even for some statistics, such as regression trees (CART algorithm - Breiman et al. 1984), the calculation of prediction intervals is not possible (Yohannes and Hoddinott 1999). Moreover, the indicator represents the variance in the prediction and does not contain information on the goodness of fit to the response variable. This point seems to be critical, as previous studies demonstrated that the reliability of the map in general could vary strongly from one indicator to another (Le Clec'h et al. 2017). Thus, we here used a more general indicator, such as the RMSE or the $R^2$, to complement the information on the prediction uncertainty. General information on the map´s reliability could help the cartographer to decide whether he/she should stop or proceed with the mapping process. However, it raises in turn the critical question of what an acceptable threshold of reliability is. Thus, the choice of threshold values should be transparent and be subject to sensitivity analysis. Finally, we are fully aware that the range of uncertainties in an ES assessment is much broader than the epistemic uncertainties induced by the statistical prediction (Hamel and Bryant 2017).

Uncertainty can become particularly important when maps aim to support policy development, for example, when they are used to analyse trade-offs and synergies (Bagstad et al. 2018, Eigenbrod et al. 2010, Palomo et al. 2018). Information about uncertainties in these maps can help to prevent unwanted or unintended consequences of policy measures that affect land-use decision or management (Jacobs et al. 2017). In addition, the provision of information about spatially explicit trustworthiness of ES maps helps to evaluate policies. Maps and information about their accuracy and flaws can support the spatially explicit identification of the relevant socio-economic and biophysical characteristics that lead to an (in-)effective policy implementation.

## Conclusions

Uncertainties are inherent and unavoidable in the assessment of environmental management in general and in particular in ES mapping. They exist in different stages of the assessments and some cannot integrally be reduced. We used here a linear model to describe the method because prediction intervals can be directly calculated. This approach could be extended to other methods, especially by the use of bootstrapping. However, their identification, acknowledgement and explanation should, at least, be systematic, to raise awareness and to determine the optimal use(s) of the maps, especially in the context of environmental policy-making and governance (Jacobs et al. 2017). Information related to uncertainty, such as the map presented in this paper, might be difficult to apprehend for non-specialists. Accordingly, further research should not only provide more holistic perspectives on uncertainties in ES mapping, but also offer insights on how to communicate such information. Further work should focus on what could be the best way (s) to map such information and how to adapt these way(s) to the purpose of the map. In that sense, our approach should be seen as a first step towards a systematic consideration and acknowledgement of uncertainties that would be included as a strong and integral component of ES mapping.

## Acknowledgements

This research was funded by the Institut des Amériques and by the French Agence Nationale de la Recherche through two grants: ANR AMAZ, coordinated by P. Lavelle and ANR AGES, coordinated by X. Arnauld de Sartre and by the Brazilian Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), also through two Grants: Processes 484990/2007-1 and 490649/2006-8.

## References

- Andrew ME, Wulder MA, Nelson TA, N. C. Coops (2015) Spatial data, analysis approaches, and information needs for spatial ecosystem service assessments: a review. GIScience & Remote Sensing 52: 3-344. https://doi.org/10.1080/15481603.2015.1033809
- Bagstad K, Cohen E, Ancona Z, McNulty S, Sun G (2018) The sensitivity of ecosystem service models to choices of input data and spatial resolution. Applied Geography 93: 25-36. https://doi.org/10.1016/j.apgeog.2018.02.005
- Boithias L, Terrado M, Corominas L, Ziv G, Kumar V, Marqués M, Schuhmacher M, Acuña V (2016) Analysis of the uncertainty in the monetary valuation of ecosystem services — A case study at the river basin scale. Science of the Total Environment 543: 683-690. https://doi.org/10.1016/j.scitotenv.2015.11.066
- Braat LC, de Groot R (2012) The ecosystem services agenda: bridging the worlds of natural science and economics, conservation and development, and public and private policy. Ecosystem Services 1: 1-4.
- Breiman L, Friedman J, Stone C, R. Olsen (1984) Classification and Tegression Trees (Boca Ratoned. USA
- Brunner SH, Huber R, Grêt-Regamey A (2017) Mapping uncertainties in the future provision of ecosystem services in a mountain region in Switzerland. Regional Environmental Change 17 (8): 2309-2321. https://doi.org/10.1007/s10113-017-1118-4
- Burkhard B, Kroll R, Nedkov S, Müller F (2012) Mapping ecosystem service supply, demand and budgets. Ecological Indicators 21: 17-29. https://doi.org/10.1016/j.ecolind.2011.06.019
- Burkhard B, Crossman N, Nedkov S, Petz K, Alkemade R (2013) Mapping and modelling ecosystem services for science, policy and practice. Ecosystem Services 4: 1-3. https://doi.org/10.1016/j.ecoser.2013.04.005

- Cornillon P-, Matzner-Lober E, (2011) Régression avec R. Springer, France https://doi.org/10.1007/978-2-8178-0184-1
- Costa S, Gonzaga L, Miranda IS, Grimaldi M, Silva ML, Mitja D, Lima TTS (2012) Biomass in different types of land use in the Brazil's 'arc of deforestation'. Forest Ecology and Management 278: 101-109. https://doi.org/10.1016/j.foreco.2012.04.007
- Crossman N, Burkhard B, Nedkov S, Willemen L, Petz K, Palomo I, Drakou E, Martín-Lopez B, McPhearson T, Boyanova K, Alkemade R, Egoh B, Dunbar M, Maes J (2013) A blueprint for mapping and modelling ecosystem services. Ecosystem Services 4: 4-14. https://doi.org/10.1016/j.ecoser.2013.02.001
- Dendoncker N, Schmit C, Rounsevell M (2008) Exploring spatial data uncertainties in land-use change scenarios. International Journal of Geographical Information Science 22 (9): 1013-1030. https://doi.org/10.1080/13658810701812836
- Devendran AA, Lakshmanan G (2014) A Review On Accuracy and Uncertainty of Spatial Data and Analyses with special reference to Urban and Hydrological Modelling. ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences 171-178. https://doi.org/10.5194/isprsannals-ii-8-171-2014
- Domac (2004) Increasing the accuracy of vegetation classification using geology and DEM. Middle East technical University, Ankara.
- Dominati E, Patterson M, Mackay A (2010) A framework for classifying and quantifying the natural capital and ecosystem services of soils. Ecological Economics 69: 1858-1868. https://doi.org/10.1016/j.ecolecon.2010.05.002
- Eigenbrod F, Armsworth P, Anderson B, Heinemeyer A, Gillings S, Roy D, Thomas C, Gaston K (2010) The impact of proxy-based methods on mapping the distribution of ecosystem services. Journal of Applied Ecology 47 (2): 377-385. https://doi.org/10.1111/j.1365-2664.2010.01777.x
- Fearnside P (2017) Deforestation of the Brazilian Amazon. Oxford Research Encyclopedia of Environmental Science https://doi.org/10.1093/acrefore/9780199389414.013.102
- Foody GM (2015) Valuing map validation: The need for rigorous land cover map accuracy assessment in economic valuations of ecosystem services. Ecological Economics 111: 23-28. https://doi.org/10.1016/j.ecolecon.2015.01.003
- Gerwing J (2002) Degradation of forests through logging and fire in the eastern Brazilian Amazon, Forest. Ecology and Management 157: 131-141. https://doi.org/10.1016/S0378-1127(00)00644-7
- Godar J, Tizado E, Pokorny B (2012) Who is responsible for deforestation in the Amazon? A spatially explicit analysis along the Transamazon Highway in Brazil, Forest. Ecology and Management 267: 58-73. https://doi.org/10.1016/j.foreco.2011.11.046
- Grêt-Regamey A, Brunner S, Altwegg J, Bebi P (2013) Facing uncertainty in ecosystem services-based resource management. Journal of Environmental Management 127: 145-154. https://doi.org/10.1016/j.jenvman.2012.07.028
- Grimaldi M, Oszwald J, Doledec S, Hurtado MdP, de Souza Miranda I, Arnauld de Sartre X, Assis WS, Castañeda E, Desjardins T, Dubs F, Guevara E, Gond V, Lima TTS, Marichal R, Michelotti F, Mitja D, Noronha NC, Delgado Oliveira MN, Ramirez B, Rodriguez G, Sarrazin M, Silva ML, d. Costa LGS, Souza SLd, Veiga I, Velasquez E, Lavelle P (2014) Ecosystem services of regulation and support in Amazonian pioneer fronts: Searching for landscape drivers. Landscape Ecology 29 (2): 311-328. https://doi.org/10.1007/s10980-013-9981-y

- Groot RSd, Alkemade R, Braat L, Hein L, Willemen L (2010) Challenges in integrating the concept of ecosystem services and values in landscape planning, management and decision making. Ecological Complexity 7 (3): 260-272. https://doi.org/10.1016/j.ecocom.2009.10.006
- Hamel P, Bryant B (2017) Uncertainty assessment in ecosystem services analyses: Seven challenges and practical responses. Ecosystem Services 24: 1-15. https://doi.org/10.1016/j.ecoser.2016.12.008
- Heuvelink GB, Burrough PA (2002) Developments in statistical approaches to spatial uncertainty and its propagation. International Journal of Geographical Information Science 16 (2): 11-113.
- Higuchi N, Santos J, Ribeiro RJ, Minette L, Biot Y (1998) Biomassa da parte aérea da vegetação da floresta tropical umida de terra-firme da Amazônia Brasileira. Acta Amazonica 28: 153-166. https://doi.org/10.1590/1809-43921998282166
- INPE, Instituto Nacional de Pesquisas Espaciais (2014) Monitoramento da Floresta Amazônica Brasileira por Satélite – Projeto Prodes. http://www.obt.inpe.br/prodes. Accessed on: 2017-10-01.
- Jacobs S, Keune H, Vrebos D, Beauchard O, Villa F, Meire P (2013) Ecosystem Service Assessments: Science or Pragmatism? Ecosystem Services. Ecosystem Services. Elsevier, Boston.
- Jacobs S, Verheyden W, Dendoncker N (2017) Why to map Mapping Ecosystem Services. Pensoft. In: Burkhard B, Maes J (Eds) Mapping Ecosystem Services. Pensoft, Sofia, Bulgaria. [ISBN 978-954-642-830-1].
- Johnson K, Polasky S, Nelson E, Pennington D (2012) Uncertainty in ecosystem services valuation and implications for assessing land use tradeoffs: An agricultural case study in the Minnesota River Basin. Ecological Economics 79: 71-79. https://doi.org/10.1016/j.ecolecon.2012.04.020
- Kangas A, Korhonen KT, Packalen T, Vauhkonen J (2018) Sources and types of uncertainties in the information on forest-related ecosystem services. Forest, Ecology and Management 427 https://doi.org/10.1016/j.foreco.2018.05.056
- Kuemmerle T, Erb K, Meyfroidt P, Müller D, Verburg PH, Estel S, Haberl H, Hostert P, Jepsen M, Kastner T, Levers C, Lindner M, Plutzar C, Verkerk PJ, der Zanden EHv, Reenberg A (2013) Challenges and opportunities in mapping land use intensity globally. Current Opinion in Environmental Sustainability 5 (5): 484-493. https://doi.org/10.1016/j.cosust.2013.06.002
- Lavorel S, Bayer A, Bondeau A, Lautenbach S, Ruiz-Frau A, Schulp N, Seppelt R, Verburg P, Teeffelen Av, Vannier C, Arneth A, Cramer W, Marba N (2017) Pathways to bridge the biophysical realism gap in ecosystem services mapping approaches. Ecological Indicators 74: 241-260. https://doi.org/10.1016/j.ecolind.2016.11.015
- Le Clec'h S, Oszwald J, Jegou N, Dufour S, Cornillon PA, Miranda I, Gonzaga L, Grimaldi M, Gond V, Arnauld de Sartre X (2013) Mapping carbon stocks in vegetation: prospects for the spatialization of an ecosystem service.". Bois Et Forets Des Tropiques 316: 35-47.
- Le Clec'h S, Jégou N, Decaens T, Dufour S, Grimaldi M, Oszwald J (2017) From Field Data to Ecosystem Services Maps: Using Regressions for the Case of Deforested Areas Within the Amazon. Ecosystems 21 (2): 216-236. https://doi.org/10.1007/s10021-017-0145-9

- Le Clec’h S, Sloan S, Gond V, Cornu G, Decaens T, Dufour S, Grimaldi M, Oszwald J (2018) Mapping ecosystem services at the regional scale: the validity of an upscaling approach. International Journal of Geographical Information Science 32 (8): 1593-1610. https://doi.org/10.1080/13658816.2018.1445256
- Maes J, Egoh B, Willemen L, Liquete C, Vihervaara P, Schägner JP, Grizzetti B, Drakou E, Notte AL, Zulian G, Bouraoui F, Paracchini ML, Braat L, Bidoglio G (2012) Mapping ecosystem services for policy support and decision making in the European Union. Ecosystem Services 1 (1): 31-39. https://doi.org/10.1016/j.ecoser.2012.06.004
- Mallows CL (2000) Some Comments onCp. Technometrics 42 (1): 87-94. https://doi.org/10.1080/00401706.2000.10485984
- Markewitz D, Davidson E, Moutinho P, Nepstad D (2004) Nutrient loss and redistribution after forest clearing on a highly weathered soil in Amazonia. Ecological Applications 14: 177-199. https://doi.org/10.1890/01-6016
- Martinez-Harms MJ, Bryan B, Balvanera P, Law E, Rhodes J, Possingham H, Wilson K (2015) Making decisions for managing ecosystem services. Biological Conservation 184: 229-238. https://doi.org/10.1016/j.biocon.2015.01.024
- Nelson B, Mesquita R, Pereira J, Garcia Aquino de Souza S, G. Teixeira Batista G, L. BC (1999) Allometric regressions for improved estimate of secondary forest biomass in the central Amazon, Forest. Ecology and Management 117: 149-167. https://doi.org/10.1016/S0378-1127(98)00475-7
- Oszwald J, Gond V, Doledec S, P. Lavelle (2011) Identification d'indicateurs de changement d'occupation du sol pour le suivi des mosaïques paysagères. Bois Et Forets Des Tropiques 307: 7-21. https://doi.org/10.19182/bft2011.307.a20484
- Oszwald J, Arnauld de Sartre X, Decaëns T, Gond V, Grimaldi M, Lefebvre A, De Araujo Fretas RL, Lindoso de Souza S, Marichal R, Veiga I, Velasquez E, Lavelle P (2012) Utilisation de la télédétection et de données socio-économiques et écologiques pour comprendre l'impact des dynamiques de l'occupation des sols à Pacajá (Brésil). Revue Française de Photogrammétrie et de Télédétection 198: 8-24.
- Oszwald J, Grimaldi M, Le Clec'h S, Dufour S (2014) Des processus biophysiques aux indicateurs de services écosystémiques: l'apport des approches paysagères. In: Arnauld de Sartre X, Castro M, Dufour S, Oszwald J (Eds) Political Ecology des Services Ecosystémiques. P. Lang, Bruxelles. [ISBN 978-2-87574-197-4].
- Palomo I, Willemen L, Drakou E, Burkhard B, Crossman N, Bellamy C, Burkhard K, Campagne CS, Dangol A, Franke J, Kulczyk S, Clec'h SL, Malak DA, Muñoz L, Narusevicius V, Ottoy S, Roelens J, Sing L, Thomas A, Meerbeek KV, Verweij P (2018) Practical solutions for bottlenecks in ecosystem services mapping. One Ecosystem 3: e20713. https://doi.org/10.3897/oneeco.3.e20713
- Palsky G (2013) Cartographie participative, cartographie indisciplinée. L'Information géographique 77: 10-25. https://doi.org/10.3917/lig.774.0010
- Perrings C, Duraiappah A, Larigauderie A, Mooney H (2011) The Biodiversity and Ecosystem Services Science-Policy Interface. Science 331 (6021): 1139-1140. https://doi.org/10.1126/science.1202400
- Portela R, Rademacher I (2001) A dynamic model of patterns of deforestation and their effect on the ability of the Brazilian Amazonia to provide ecosystem services. Ecological Modelling 143: 115-146. https://doi.org/10.1016/S0304-3800(01)00359-3
- Réjou-Méchain M, Tymen B, Blanc L, Fauset S, Feldpausch T, Monteagudo A, Phillips O, Richard H, Chave J (2015) Using repeated small-footprint LiDAR acquisitions to infer

spatial and temporal variations of a high-biomass Neotropical forest. Remote Sensing of Environment 169: 93-101. https://doi.org/10.1016/j.rse.2015.08.001

- Schulp CE, Burkhard B, Maes J, Vliet JV, Verburg P (2014) Uncertainties in Ecosystem Service Maps: A Comparison on the European Scale. PLoS ONE 9 (10): e109643. https://doi.org/10.1371/journal.pone.0109643
- van Oudenhoven AE, Aukes E, Bontje L, Vikolainen V, van Bodegom P, Slinger J (2018) 'Mind the Gap' between ecosystem services classification and strategic decision making. Ecosystem Services 33: 77-88. https://doi.org/10.1016/j.ecoser.2018.09.003
- Verburg P, Tabeau A, Hatna E (2013) Assessing spatial uncertainties of land allocation using a scenario approach and sensitivity analysis: A study for land use in Europe. Journal of Environmental Management 127: 132-144. https://doi.org/10.1016/j.jenvman.2012.08.038
- Verburg PH, Neumann K, Nol L (2011) Challenges in using land use and land cover data for global change studies. Global Change Biology 17 (2): 974-989. https://doi.org/10.1111/j.1365-2486.2010.02307.x
- Yohannes Y, Hoddinott J (1999) Classification and regression trees: an introduction. In: International Food Policy Research Institute (IFPRI) Technical Guide 3. Washington DC.