Spatial congruence between carbon and biodiversity across forest landscapes of northern Borneo

Nicolas Labrière, Bruno Locatelli, Ghislain Vieilledent, Selly Kharisma, Imam Basuki, Valéry Gond, Yves Laumonier

ABSTRACT

Understanding how carbon and biodiversity vary across tropical forest landscapes is essential to achieving effective conservation of their respective hotspots in a global context of high deforestation. Whether conservation strategies aimed at protecting carbon hotspots can provide co-benefits for biodiversity protection, and vice versa, highly depends on the extent to which carbon and biodiversity co-occur at the landscape level. We used field measurements and easily accessible explanatory variables to model aboveground carbon density, soil carbon density and tree alpha diversity (response variables) over a mostly forested area of northern Borneo. We assessed the spatial relationships between response variables and the spatial congruence of their hotspots. We found a significant positive relationship between aboveground carbon density and tree alpha diversity, and an above-than-expected-by-chance spatial congruence of their hotspots. Consequently, the protection of areas of high aboveground carbon density through financial mechanisms such as REDD+ is expected to benefit tree diversity conservation in the study area. On the other hand, relationships between soil carbon density and both aboveground carbon density and tree alpha diversity were negative and spatial congruences null. Hotspots of soil carbon density, mostly located in peatlands, therefore need specific conservation regulations, which the current moratorium on peat conversion in Indonesia is a first step toward.

© 2016 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

HIGHLIGHTS

- We modeled tree diversity and carbon density using field measurements and accessible data.
- Aboveground carbon density and tree diversity were strongly correlated.
- High soil carbon density did not overlap with high aboveground carbon density or tree diversity.
- Protecting areas of high aboveground carbon density will benefit tree diversity.
- High soil carbon in peatlands must be protected with specific regulations.

ARTICLE INFO

Article history:
Received 23 December 2015
Accepted 23 January 2016

Keywords:
Ecosystem services
Spatialization
Spatial relationship
REDD+
Protected areas
Land-use planning
1. Introduction

Information on the nature, strength and extent of spatial relationships between multiple ecosystem services (ES) and biodiversity (that can be seen as a good, a final ES or a determinant of the delivery of other ES; see Mace et al., 2012) is crucial for sound ecosystem management and land-use planning (de Groot et al., 2010; Egoih et al., 2008; Naidoo et al., 2008), especially in tropical forest landscapes where conservation versus development goals are at stake (Malhi et al., 2014). Considering multiple ES together constitutes a major challenge in ecosystem management (Raudsepp-Hearne et al., 2010) but is necessary for managing trade-offs between ES and creating new funding opportunities for conservation by bundling co-occurring ES (Carpenter et al., 2006; Wendland et al., 2010).

Tropical forests have long received much attention for conservation, initially with respect to the extraordinary biodiversity that they host (15 of the 25 biodiversity hotspots sensu Myers include tropical forests; see Myers et al., 2000). The number of tree species in tropical forests, for example, is estimated to lie between 40,000 and 53,000 compared to only 124 tree species across temperate Europe (Slik et al., 2015). To date, the conservation of biodiversity has mostly relied on a network of protected areas. However, the efficiency and effectiveness of this strategy has been challenged because protected areas are too few, too small (at least many of them) and often lack sufficient funding, with only a small fraction of them shown to succeed in disrupting biodiversity erosion (Kramer et al., 1997; Laurance et al., 2012).

In the last decade, the development of REDD+, a United Nations initiative aimed at Reducing Emissions from Deforestation and forest Degradation, has shed new light on the necessity to protect tropical forests. Because tropical forests store large amounts of carbon, mostly in living woody biomass (Baccini et al., 2012) and soils (especially in peatlands; see Page et al., 2011), and are currently disappearing fast (Kim et al., 2015), their protection is particularly relevant for climate change mitigation (Pan et al., 2011). The REDD+ mechanism, which aims to provide financial incentives to maintain and enhance forest carbon stocks, appears to some as an unprecedented opportunity for biodiversity conservation provided strong safeguards are incorporated (Gardner et al., 2012; Paoli et al., 2010; Phelps et al., 2012). However, biodiversity monitoring as part of the social and environmental safeguards for REDD+ has not received much attention, with the focus clearly remaining on greenhouse gas emission estimations (Dickson and Kapos, 2012).

Schemes dedicated to biodiversity and carbon protection are not meant to be mutually exclusive. Yet many challenges (institutional, political, social, economic, etc.) need to be overcome for protection scheme optimization, for example, protected areas benefitting from the financial support of REDD+ despite seemingly lack of additionality (a project is ‘additional’ when emission reductions are linked to its implementation and would not have occurred without it; Macdonald et al., 2011), or REDD+ positively integrating biodiversity safeguards. Beyond scheme design and economic efficiency (that is necessary to tackle the high opportunity cost for conversion and relies for example on incentive structure; e.g. see Busch et al., 2012; Fisher et al., 2011), whether biodiversity conservation could benefit from climate change mitigation-oriented financial schemes, and vice versa, highly depends on the extent to which carbon and biodiversity co-occur at the landscape level (Strassburg et al., 2010).

So far, there is little agreement on the spatial relationship between carbon storage and biodiversity. At a global scale reports are contradictory, with one study pointing out an “overall lack of spatial concordance between biodiversity and ecosystem services” (including carbon storage; see Naidoo et al., 2008) while another found a “high congruence between species richness and biomass carbon” (Strassburg et al., 2010). The same seemingly conflicting results arose at national (e.g. see Egoih et al., 2009 vs. Locatelli et al., 2014) and local scales (e.g. see Kessler et al., 2012 vs. Ruiz-Jaen and Potvin, 2010).

Three key elements might shed some light on the apparent contradiction between these findings. First, as pointed out by Eigenbrod et al. (2010), proxies used to assess ES distribution may poorly fit primary data, impacting the quality of resulting ES maps and, therefore, the identification of ES hotspots and areas of spatial congruence (i.e. overlap) between multiple ES. Second, the nature, strength and extent of spatial relationships between carbon and biodiversity clearly depend on the metrics selected for carbon (e.g. aboveground carbon only vs. aboveground and soil organic carbon) and biodiversity (e.g. richness, threat, restricted range) in each study (Murray et al., 2015). Third, spatial covariance of ES provision can be influenced by data spatial resolution and study spatial extent (Anderson et al., 2009; Magnago et al., 2015; Murray et al., 2015).

There is a need for more primary data from data-scarce regions to characterize the spatial distribution of, and spatial relationship between, carbon storage and biodiversity so as to guide decision making on land-use planning, especially at the landscape level where most management decisions are made. Working in an area of northern Borneo that is mostly forested but facing clearance with the development of oil palm plantations, we assessed aboveground carbon density (ACD), soil carbon density (SCD, for the 0–20 cm soil layer) and tree alpha diversity (TAD, using Fisher’s α) in sampling plots scattered across the study area. The following research questions were addressed: (1) what are the relationships between carbon density and tree diversity across the study area; (2) do carbon and tree diversity hotspots overlap in the study area; and (3) to what extent can biodiversity conservation policies also protect carbon stocks and vice versa. We modeled ACD, SCD and TAD (response variables) with explanatory variables that are easily accessible at the landscape scale. We applied the resulting models over the whole study area and compared our predictions with existing maps for ACD, SCD and TAD. We then explored relationships between response variables and spatial congruence of their respective hotspots. We finally assessed threats and opportunities for carbon storage and tree diversity based on hotspot location, and quantified trade-offs in ES protection from different conservation strategies.
2. Materials and methods

2.1. Study area

Fieldwork was conducted in the Kapuas Hulu regency (hereafter referred to as the “study area”) in the Indonesian province of West Kalimantan. The study area spans ca. 31,000 km² of mostly forested land where altitude ranges from 0 to ca. 2000 masl. Mean annual precipitation varies from 2350 to 4300 mm with an average of ca. 3450 mm, and mean annual temperature varies from 17–27 °C with an average of ca. 25 °C (Hijmans et al., 2005). The majority of soils have developed on sedimentary parent material and belong to the Ultisols, Inceptisols and Histosols orders (RePPProT, 1990).

Two national parks (Betung Kerihun and Danau Sentarum) cover about 30% of the study area (Shantiko et al., 2013). In 2003, Kapuas Hulu was declared a conservation area by the local government, a commitment to improved natural resource management and ecosystem functionality preservation (Prasetyo et al., 2007).

2.2. Field sampling and index computation for carbon and tree diversity

We sampled the main vegetation types between 2011 and 2013. Of the 120 plots sampled, 85 were 100 × 20 m with natural or moderately disturbed vegetation, and 35 were 20 × 20 m with rubber gardens and/or secondary regrowth. Smaller plots were preferred for rubber gardens and secondary regrowth to ensure that plot vegetation was homogeneous despite landscape heterogeneity associated with swidden practices. In each plot, all trees with ≥ 10 cm diameter at breast height (1.3 m above ground) were measured, tagged and mapped, and their height estimated using a Blume-Leiss hypsometer (for details, see Walker et al., 2012). Leaf samples were collected and identified at the Herbarium Bogoriense in Bogor, Indonesia. In total, over the 18.4 ha of surveyed vegetation, 14,155 trees were measured and 4480 herbarium vouchers collected and identified.

Tree dry biomass was computed using a pantropical allometric model that includes tree diameter, tree height and wood specific gravity as explanatory variables (Chave et al., 2014), and carbon content was derived using the standard conversion factor of 0.47 (McGroddy et al., 2004). Wood specific gravity, a constitutive factor of the aforementioned equation, was obtained from the Global Wood Density Database (Zanne et al., 2009). When species were not found in the database, the genus-level average wood density was used instead. Any unresolved cases (ca. 1% of all specimens) were assigned the value of the mean wood density for tropical Southeast Asia (0.57 g cm⁻³; see Chave et al., 2009). ACD (in Mg ha⁻¹) was calculated by averaging total plot carbon content over plot area.

For mineral soils, composite soil samples were taken for topsoils (0–20 cm; collected using an auger at four different locations for each composite sample) in half of the vegetation plots. Samples were dried at 105 °C and further analyzed for carbon content (using Walkley and Black method; see Landon, 1984). Topsoil cores were also collected using 100 cm³ sample rings to measure dry bulk density (in g cm⁻³). SCD (in Mg ha⁻¹ for the 0–20 cm soil layer) was calculated using carbon content and dry bulk density. For peat soils, peat samples were collected for topsoils (0–15 cm) using a 203.4 cm³ Russian peat sampler. Samples were dried at 60 °C and weighted to measure dry bulk density. They were further analyzed for carbon content using a LECO TrueSpec induction furnace C analyzer (following Warren et al., 2012). As we only had a restricted number of soil samples in peatlands and swamp areas (n = 8), we also reviewed the literature in search of data of (1) soil carbon density or (2) carbon content and bulk density from sampling points geolocalized in the study area. We were only able to retrieve information about two sampling points from one study (Anshari et al., 2010).

We used Fisher’s α as an indicator to characterize TAD. Fisher’s α, a parametric index, is relatively independent of sample size and insensitive to the presence of rare species (Colwell, 2009; Parmentier et al., 2011). For details about sampling strategy, indices computation, and plot structural and compositional features, see Appendix A. Field estimates of ACD, SCD and tree diversity were compared with existing maps (Saatchi et al., 2011 and Baccini et al., 2012 for ACD; Wieder et al., 2014 for SCD; Slik et al., 2009, Raes et al., 2009; and Raes et al., 2013 for TAD; see Appendix B for details).

2.3. Potential explanatory variables to be tested for model building

We chose potential explanatory variables that are commonly used in the literature for this type of modeling and for which spatial data were available over the study area (Table 1). We worked at a 250 m grid cell resolution, which is the spatial resolution of MODIS MOD13Q1 (vegetation indices) and MOD44B (vegetation continuous fields) products. These products have been used widely for spatialization of various ground-measured vegetation-related response variables (e.g. see Nagler et al., 2007; Saatchi et al., 2008). All geographical information was projected using a Universal Transverse Mercator projection (zone 49 N, WSG 84 datum).

The MODIS MOD13Q1-derived vegetation indices (EVI and NDVI) have a temporal resolution of 16 days (23 images/year). We gathered images from 2011, 2012 and 2013 to cover the entire fieldwork period. For both EVI and NDVI, we computed a maximum, mean and standard deviation for each grid cell over the 3-year period (i.e. 3 × 23 = 69 images). The MODIS MOD44B-derived vegetation continuous fields (percent tree cover, percent non-tree vegetation and percent non vegetation) have a temporal resolution of one year. We gathered images from 2011 to 2013 and computed the average value for each of the three vegetation continuous fields.
### Table 1
Description of the 20 potential explanatory variables tested for ACD, SCD and TAD model building.


<table>
<thead>
<tr>
<th>Data origin</th>
<th>Source</th>
<th>Code</th>
<th>Description</th>
<th>Unit</th>
<th>Spatial resolution</th>
<th>Range over study area</th>
<th>Range over plot network</th>
<th>Example reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>MODIS</td>
<td>[1]</td>
<td>EVIMAX</td>
<td>Maximum EVI&lt;sup&gt;b&lt;/sup&gt;</td>
<td>ind&lt;sup&gt;d&lt;/sup&gt;</td>
<td>250 m</td>
<td>0.1617–1.0000</td>
<td>0.5549–0.9316</td>
<td>Parmentier et al. (2011)</td>
</tr>
<tr>
<td>MODIS</td>
<td>[1]</td>
<td>EVIMEAN</td>
<td>Mean EVI</td>
<td>ind</td>
<td>250 m</td>
<td>0.0298–0.6459</td>
<td>0.3567–0.5518</td>
<td>Parmentier et al. (2011)</td>
</tr>
<tr>
<td>MODIS</td>
<td>[1]</td>
<td>EVISD</td>
<td>Standard deviation EVI</td>
<td>ind</td>
<td>250 m</td>
<td>0.0447–0.2090</td>
<td>0.0727–0.1349</td>
<td>Parmentier et al. (2011)</td>
</tr>
<tr>
<td>MODIS</td>
<td>[1]</td>
<td>NDVIMAX</td>
<td>Maximum NDVI&lt;sup&gt;c&lt;/sup&gt;</td>
<td>ind</td>
<td>250 m</td>
<td>0.3417–0.9995</td>
<td>0.8868–0.9984</td>
<td>Parmentier et al. (2011)</td>
</tr>
<tr>
<td>MODIS</td>
<td>[1]</td>
<td>NDVIMEAN</td>
<td>Mean NDVI</td>
<td>ind</td>
<td>250 m</td>
<td>0.0471–0.8810</td>
<td>0.6221–0.8391</td>
<td>Parmentier et al. (2011)</td>
</tr>
<tr>
<td>MODIS</td>
<td>[1]</td>
<td>NDVISD</td>
<td>Standard deviation NDVI</td>
<td>ind</td>
<td>250 m</td>
<td>0.0460–0.3106</td>
<td>0.0850–0.2197</td>
<td>Oindo and Skidmore (2002)</td>
</tr>
<tr>
<td>MODIS</td>
<td>[1]</td>
<td>PNTV</td>
<td>Percent non tree vegetation</td>
<td>%</td>
<td>250 m</td>
<td>0–78</td>
<td>6–40</td>
<td>Asner et al. (2014)</td>
</tr>
<tr>
<td>SRTM&lt;sup&gt;a&lt;/sup&gt;</td>
<td>[2]</td>
<td>SLOPE</td>
<td>Slope</td>
<td>%</td>
<td>90 m → 250 m</td>
<td>0.0–48.5</td>
<td>0.1–16.3</td>
<td>Marshall et al. (2012)</td>
</tr>
<tr>
<td>WordClim</td>
<td>[3]</td>
<td>TEMPE</td>
<td>Mean annual temperature °C</td>
<td></td>
<td>1 km → 250 m</td>
<td>17.2–26.8</td>
<td>22.5–26.6</td>
<td>Parmentier et al. (2011)</td>
</tr>
<tr>
<td>WordClim</td>
<td>[3]</td>
<td>TEMPERANGE</td>
<td>Temperature range °C</td>
<td></td>
<td>1 km → 250 m</td>
<td>8.5–10.0</td>
<td>9.1–9.9</td>
<td>Slik et al. (2010)</td>
</tr>
<tr>
<td>WordClim</td>
<td>[3]</td>
<td>TEMPESEAS</td>
<td>Temperature seasonality °C</td>
<td></td>
<td>1 km → 250 m</td>
<td>1.3–4.3</td>
<td>1.9–3.5</td>
<td>Slik et al. (2010)</td>
</tr>
<tr>
<td>WordClim</td>
<td>[3]</td>
<td>PRECIP</td>
<td>Mean annual precipitation mm</td>
<td></td>
<td>1 km → 250 m</td>
<td>2346–4286</td>
<td>2818–4149</td>
<td>Parmentier et al. (2011)</td>
</tr>
<tr>
<td>Governmental</td>
<td>[4]</td>
<td>LANDALLOC</td>
<td>Land allocation</td>
<td>cat&lt;sup&gt;e&lt;/sup&gt;</td>
<td>nr</td>
<td>Area for other uses</td>
<td>Area for other uses</td>
<td>Van der Laan et al. (2014)</td>
</tr>
<tr>
<td>Governmental</td>
<td>[5]</td>
<td>MINDIST2</td>
<td>Minimum distance to disturbance source (either road, river or village)</td>
<td>m</td>
<td>nr</td>
<td>0–57,695</td>
<td>0–3579</td>
<td>Van der Laan et al. (2014)</td>
</tr>
</tbody>
</table>

<sup>a</sup> SRTM = Shuttle Radar Topographic Mission.

<sup>b</sup> EVI = Enhanced Vegetation Index.

<sup>c</sup> NDVI = Normalized Difference Vegetation Index.

<sup>d</sup> Ind = index, theoretically varying from 0 to 1.

<sup>e</sup> Cat = categorical variable.
Digital elevation data from the NASA Shuttle Radar Topographic Mission were obtained at a 90 m resolution (Jarvis et al., 2008) and further resampled to 250 m. Slope was computed from original data and subsequently resampled to match the designated study resolution.

We also used WorldClim dataset, which provides interpolated estimates of various bioclimatic variables for the 1950–2000 period with a 30 arc-second resolution (~1 km resolution at the equator; see Hijmans et al., 2005), and from which we extracted mean annual temperature, temperature seasonality, temperature range, mean annual precipitation, precipitation seasonality and precipitation range (all resampled to match our 250 m resolution).

Using ArcGis 10, we computed euclidean distances from three potential disturbance sources: roads (including logging roads), rivers and villages. A raster of minimum distance to potential disturbance source was created by selecting the minimum value between the three potential disturbance sources for each grid cell. Land allocation (i.e. the designated use of an area, e.g. production forest or national park) and soil data were also considered potential explanatory variables (obtained from the Ministry of Forestry and the Indonesian Soil Research Institute, respectively; see Table 1 for respective classes).

2.4. Explanatory variable selection and model building

We used “random forest” (a machine learning algorithm; see Breiman et al., 1984) for explanatory variable selection and response variable modeling. Random forests are collections of decision trees, each trained on a bootstrap sample of a full set of observations. Model accuracy is then evaluated for each tree using observations that had been left out of the corresponding bootstrap sample (Breiman et al., 1984). This modeling technique has already been used to predict aboveground biomass (e.g. Baccini et al., 2012) and tree alpha diversity (e.g. Parmentier et al., 2011).

In case several measurement plots belonged to the same 250 m grid cell, mean response variable values were computed. Sample size was reduced from 120 individual plots to 65 composite sample sites, these constituting our “full set of observations”. For each response variable, we grew a random forest of 1000 trees with the 20 potential explanatory variables. We then selected variables based on variable minimal depth and importance value (information available as model outputs; see Ishwaran and Kogalur, 2015). In addition, we eliminated the variables that showed unexplainable partial dependence behavior or whose distribution in observations was not representative of the entire study area. Details about random forest regression algorithm and explanatory variable selection are given in Appendix C.

Once explanatory variables were selected, new 1000–tree random forests were grown for each response variable and their performances in predicting dataset response values recorded. Random forests were finally used to predict response values over the whole study area. Because our measurements focused on tree-dominated vegetation lower than 900 masl, we excluded grid cells with either: (1) water, (2) non-forest vegetation (MODIS percent tree cover <50%; following Waring et al., 2006), or (3) altitude > 900 m. After mask application, 391,523 grid cells remained.

2.5. Relationships between response variables, spatial congruence of their hotspots and potential threat analysis

We used spatial regression models to study relationships between pairs of response variables based on our predictions. Soil type (mineral vs. peat) was explicitly incorporated in our models to test its influence on relationships between pairs of response variables. Analyses were performed at three different spatial resolutions (250 m, 1 km and 10 km) on a random subset of values (n = 250, i.e. ca. total number of non-empty grid cells at the coarsest resolution). At 1 km and 10 km spatial resolution, values were drawn among aggregated grid cells for which at least half initial grid cells had values. Analyses were done on original, square root- or log-transformed values, whichever format led to data distribution closest to normality. We tested for spatial autocorrelation on both initial values and residuals of a linear model using Moran’s I. In case residuals were still spatially correlated, we used the Lagrange Multiplier diagnostics for spatial dependence to determine the structure of the appropriate spatial regression model (i.e., spatial error model that accounts for error term correlation vs. spatial lag model that accounts for non-independence between observations; Haining, 1990).

Signs and significance of the resulting model coefficients were used to assess relationships between pairs of response variables. Only results at 250 m spatial resolution are presented here (see Appendix D for results at 1 km and 10 km spatial resolution).

The same method was used to study relationships between our predictions and those from existing maps, except soil type was not incorporated and we only worked at the spatial resolution of existing maps (see Appendix E for details).

For each response variable, hotspots were defined with three different thresholds, as the 10%, 20% or 30% grid cells with highest values. Spatial congruence between hotspots of response variables X and Y was evaluated using the proportion of grid cells that were hotspots for both response variables. This proportion can theoretically vary from 0% to 10% (20% or 30% for other hotspot thresholds, respectively) and would be 1% (4% or 9%, respectively) for two random distributions. The level of spatial congruence was used to infer the potential effect of conservation prioritization (i.e. spatial targeting of conservation measures to hotspots of a given response variable) on non-target response variables. Prioritizing hotspots of the response variable X for conservation was declared beneficial, neutral or detrimental to the non-target response variable Y when spatial congruence was higher, similar or lower than that found between two random distributions (i.e. 1%, 4% or 9% depending on threshold). We worked at a 10 km spatial resolution (i.e. the coarsest spatial resolution of existing maps) to jointly assess the spatial congruence of hotspots of response variables from our predictions and those...
of existing maps. Only results using the 10% threshold are presented here (see Appendix F for results using the 20% and 30% thresholds).

To assess potential threats to hotspots of response variables over the study area, we classified hotspot grid cells according to the land allocation under which they are situated, and recorded whether the corresponding areas were included in logging, mining or plantation concessions (concession data obtained from the Ministry of Forestry). We only performed this potential threat analysis on our predictions and therefore worked at the original (i.e. 250 m) spatial resolution of our predictions.

All statistical analyses were done using R (R Core Team, 2014). Most frequently used R packages include ‘vegan’ (Oksanen et al., 2015), ‘raster’ (Hijmans, 2015), ‘spdep’ (Bivand and Piras, 2015; Bivand et al., 2013), ‘randomForest’ (Liaw and Wiener, 2002) and ‘randomForestSRC’ (Ishwaran and Kogalur, 2015).

3. Results

3.1. Response variable predictions over the study area

Altitude was a key explanatory variable for all response variables, whereas percent tree cover was only important in explaining ACD and TAD (Fig. 1, Table 2). The percentage of variance explained by the models ranged from 50%–80% depending on response variable. Application of these models over the study area revealed differences in the broad pattern of response variable distribution, with high spatial variability for ACD, and the highest values for TAD and SCD in the outer and inner part of the study area, respectively (Fig. 2).

Our predictions of all response variables were significantly correlated to their values in existing maps but the fit was often poor (Fig. 3). When comparing our predictions of ACD with those of existing maps, we found that the slope was only a tenth to a fifth of the slope expected in case of perfect match. We also evidenced a saturation effect from both existing maps (i.e. none to few of their predicted values exceeded a threshold of about 175 Mg ha$^{-1}$; Fig. 3(a)–(b)). There was a significant positive relationship between our predictions of SCD and the top 30 cm soil carbon values in the Harmonized World Soil Database despite the limited number of unique values of the latter (Fig. 3(c)). Comparisons of our predictions of TAD with three existing maps showed mixed results, with relationships being either negative or positive depending on which map was used (see Fig. 3(d)–(f)).

3.2. Response variable relationships and hotspot spatial congruence

Despite soil type (mineral vs. peat) being consistently tested for inclusion in the models, the most parsimonious models did not necessarily include soil type and/or its interaction with other predictors (Table 3). The relationship between TAD and ACD was positive and the interaction between ACD and soil type – but not soil type itself – was significant and negative. In other words, TAD increased more rapidly with increasing ACD over mineral soils than over peat soils. Neither soil type nor its interaction with SCD was a significant predictor of TAD, and we found that TAD significantly decreased with increasing SCD. Soil type and SCD were both significant predictors of ACD, the latter being higher over peat soils for a given SCD and decreasing with increasing SCD. Analyzing the same relationships at 1 km and 10 km spatial resolution, we did not find any strong scale-dependent effects on our findings (see Appendix D).

Spatial congruence was very variable depending on the nature (ACD, SCD or TAD) and source (our predictions vs. predictions from existing maps) of response variables under consideration except for spatial congruence between hotspots of ACD and SCD that was consistently null (Fig. 4, Table 4). We found no spatial congruence between hotspots of SCD and TAD based on our predictions while spatial congruence could potentially be high based on predictions from some other sources, e.g. for SCD (Wieder’s) with TAD (Slik’s). According to our predictions, spatial congruence between hotspots of ACD and TAD was more than three times (3.5:1) higher than expected if hotspot spatial distributions were random, meaning that about a third of ACD hotspots were also hotspots of tree diversity. Conversely, most combinations of predictions of ACD and TAD resulted in spatial congruence null or close to that attained for random distributions. Using 20% or 30% thresholds for hotspot definition did not lead to markedly different results on spatial congruence between hotspots of ACD, SCD and TAD (see Appendix F).
4. Discussion

4.1. Ecological insight about the relationship between ACD and TAD

We found a significant positive relationship between our predictions of ACD and TAD over the study area. The only non-null spatial congruence – at least when considering our predictions – was found between hotspots of ACD and TAD. Using values computed from field measurements (those from which our predictions originated), we found a significant positive relationship between ACD and TAD (Fig. 5) despite high ACD variability for a given TAD value.

Part of that variability is inherently due to that we surveyed very different vegetation types. For example, TAD could be similar for some lowland natural forest and peat swamp plots (e.g. Fisher’s $\alpha = 25$; see Fig. 5), but ACD was much lower in the latter vegetation type. Soil nutrient content has been shown to affect spatial variations of aboveground biomass (and therefore carbon density) and tree diversity in Borneo (Cannon and Leighton, 2004; Paoli et al., 2008), with forests growing on oligotrophic soils (e.g. Kerangas forests on sandstone and peat swamp forests) containing lower aboveground biomass and fewer tree species than nearby forests on well-drained mineral soils (e.g. see Anderson, 1964 and Brünig, 1974). While we found that TAD was limited in most nutrient-poor natural forest plots, ACD was more variable and some values reached surprisingly high levels compared to those of lowland natural forest plots on mineral soils (see Fig. 5). This potentially originates from the relatively small size of our plots (0.04–0.2 ha).
Two – not necessarily mutually exclusive – hypotheses have gained attention in explaining how biodiversity might influence ecosystem properties such as carbon storage. The niche complementarity hypothesis states that higher levels of biodiversity lead to greater carbon storage because of increased resource use, whereas the mass ratio hypothesis holds that carbon storage is mostly driven by functional trait properties of the dominant species (Loreau and Hector, 2001). The significant positive relationship we found between ACD and TAD supports the niche complementarity hypothesis. Consistent with our findings, a study conducted in Panama found that species richness increased tree carbon storage (Ruiz-Jaen and
Fig. 3. Comparison of our predictions with values extracted from existing maps of (a) ACD (Baccini et al., 2012); (b) ACD (Saatchi et al., 2011); (c) SCD (Wieder et al., 2014); (d) TAD (Slik et al., 2009); (e) TAD (Raes et al., 2009); (f) TAD (Raes et al., 2013). Information about the best-fit significant model (selected among linear and spatial regression models) is displayed in each facet. Model coefficients (for spatial dependence and/or intercept and independent variable) are presented with information on statistical significance. Variable: ACD = aboveground carbon density; SCD = soil carbon density; TAD = tree alpha diversity; Selected model: lm = linear model; slm = spatial lag model; Significance: n.s. non-significant; *p < 0.05; **p < 0.01; ***p < 0.001.

Table 3

Relationships between ACD, SCD and TAD (response variables) over the study area depending on soil type (mineral vs. peat). Analyses were performed on a random subset of response variable values (n = 250) from our predictions at 250 m spatial resolution. We used original, square root- or log-transformed data, whichever format led to data distribution closest to normality. Depicted values are either test statistics (for Moran’s I, Lagrange Multiplier and the selected model) or model coefficients, and are presented along with information on statistical significance. Note that the most parsimonious model did not necessarily include soil type and/or its interaction.

<table>
<thead>
<tr>
<th>General model form</th>
<th>Y ~ X + PEAT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variable^a</td>
<td>Model 1</td>
</tr>
<tr>
<td>Data format</td>
<td>Y = TAD</td>
</tr>
<tr>
<td>Moran’s I</td>
<td>0.26***</td>
</tr>
<tr>
<td>On variable</td>
<td>0.07***</td>
</tr>
<tr>
<td>On linear model residuals</td>
<td>40.20***</td>
</tr>
<tr>
<td>Lagrange Multiplier (LM)</td>
<td>70.46***</td>
</tr>
<tr>
<td>LM^b</td>
<td>583.05***</td>
</tr>
<tr>
<td>Selected model type^b</td>
<td>slm</td>
</tr>
<tr>
<td>Model statistic^c</td>
<td>Intercept</td>
</tr>
<tr>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Model coefficients</td>
<td>PEAT</td>
</tr>
<tr>
<td>X:PEAT</td>
<td>-0.20**</td>
</tr>
<tr>
<td>Spatial dependence^d</td>
<td>0.94***</td>
</tr>
</tbody>
</table>

^a Variable: ACD = aboveground carbon density; SCD = soil carbon density; TAD = tree alpha diversity; PEAT = binary variable with mineral = 0 and peat = 1.

^b sem = spatial error model; slm = spatial lag model.

^c Wald statistic for spatial error or spatial lag model.

^d Rho in case of spatial lag model, Lambda in case of spatial error model.

n.s. non-significant; **p < 0.01; ***p < 0.001.

However, the study also showed that dominance was important in explaining tree carbon storage variations, thus highlighting the relevance of the mass ratio hypothesis. More work is required to test the mass ratio hypothesis with
Fig. 4. Spatial distribution of hotspots of response variables at 10 km spatial resolution using a 10% threshold: (a) ACD (our predictions); (b) SCD (our predictions); (c) TAD (our predictions); (d) ACD (Baccini et al., 2012); (e) SCD (Wieder et al., 2014); (f) TAD (Slik et al., 2009); (g) ACD (Saatchi et al., 2011); (h) TAD (Raes et al., 2009); (i) TAD (Raes et al., 2013).

Table 4
Spatial congruence between hotspots of ACD, SCD and TAD (response variables) at 10 km spatial resolution using a 10% threshold. Values (percentage of overlapping grid cells over the total number of grid cells) potentially range from 0 to 10%. Expected spatial congruence for two variables with random spatial distribution is 1%. For the sake of readability, values are highlighted with different colors depending on the nature of the response variables under consideration (yellow, green and blue for ACD-SCD, ACD-TAD and SCD-TAD, respectively).

<table>
<thead>
<tr>
<th></th>
<th>SCD (this study)</th>
<th>TAD (this study)</th>
<th>ACD (Baccini's)</th>
<th>SCD (Wieder's)</th>
<th>TAD (Slik's)</th>
<th>ACD (Saatchi's)</th>
<th>TAD (Raes 2009)</th>
<th>TAD (Raes 2013)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACD (this study)</td>
<td>0.0</td>
<td>3.5</td>
<td>0.4</td>
<td>0.0</td>
<td>0.0</td>
<td>2.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>SCD (this study)</td>
<td>0.0</td>
<td>0.0</td>
<td>6.3</td>
<td>4.3</td>
<td>0.0</td>
<td>0.0</td>
<td>0.8</td>
<td>0.0</td>
</tr>
<tr>
<td>TAD (this study)</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>5.1</td>
<td>0.4</td>
<td>0.0</td>
<td>2.0</td>
</tr>
<tr>
<td>ACD (Baccini's)</td>
<td>1.2</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>1.2</td>
<td>1.8</td>
<td>2.0</td>
<td>2.0</td>
</tr>
<tr>
<td>SCD (Wieder's)</td>
<td>7.1</td>
<td>0.0</td>
<td>0.0</td>
<td>2.0</td>
<td>2.0</td>
<td>0.0</td>
<td>2.4</td>
<td>0.4</td>
</tr>
<tr>
<td>TAD (Slik's)</td>
<td>0.0</td>
<td>2.4</td>
<td>0.0</td>
<td>4.4</td>
<td>2.8</td>
<td>2.0</td>
<td>2.4</td>
<td></td>
</tr>
<tr>
<td>ACD (Saatchi's)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TAD (Raes 2009)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

regard to our data and to weight, if necessary, the relative relevance of the two hypotheses in such a tropical forest context. Analyzing the compositional component of tree diversity and species turnover from site to site might also help gain new insights into ecosystem functioning and subsequent carbon storage.

4.2. Potential threats over carbon and tree diversity hotspots

Considering land allocation as an indicator of potential threats to carbon and tree diversity hotspots, we found that a very high proportion of ACD and TAD hotspots were located either in watershed protection forests or national parks (Table 5). Very few ACD and TAD hotspots were found in concessions, which legally avoid watershed protection forests or national parks (see Appendix G for more information about concessions extent and location). Overall, provided current land allocation is maintained and the law is enforced to ensure the integrity of protected areas, ACD and TAD hotspots appear to be under low threat.

In contrast, we found that a worryingly high proportion of SCD hotspots were located in areas under disturbance-prone land allocation (“area for other uses” and “limited production forest”) and overlapped logging and/or plantation concessions (Table 5). This raises concerns about the extent to which soil carbon stocks are secured in the study area. The vast majority
Fig. 5. Tree alpha diversity (TAD, using Fisher’s $\alpha$) against aboveground carbon density (ACD). Gray dots represent model predictions (10% random selection) over the study area. Colored dots correspond to field measurements. Information about the best-fit significant model between TAD and ACD (selected among linear and spatial regression models) is also displayed. Soil type (mineral vs. peat) and its interaction with ACD were tested for inclusion in the model. The most parsimonious model included the interaction term but not soil type itself. Model coefficients (for spatial dependence, intercept, independent variable and interaction between soil type and independent variable) are presented along with information on statistical significance. Vegetation type: HiF = hill natural forest on mineral soil; KeF = Kerangas forest (i.e. forest on sandstone); LoF = logged-over forest; LwF = lowland natural forest on mineral soil; SeV = secondary vegetation (either secondary regrowth or rubber gardens); SwF = swamp forest (either freshwater or peat swamp forests); Selected model: slm = spatial lag model; Significance: n.s. non-significant; $^* p < 0.05; ^{**} p < 0.01; ^{***} p < 0.001$. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 5

Distribution (%) of response variable hotspots depending on land allocation and presence of concession at 250 m spatial resolution using a 10% threshold. The expected repartition of hotspots of a response variable with random distribution is given for comparison. Land allocation columns are ordered from left to right along a gradient of decreasing likeliness of disturbance. The sum of each row across the six land allocation types is equal to 100%. Note that there can be more than one concession type on the same grid cell. A high proportion of SCD hotspots (compared to hotspots of other response variables) are situated in areas that are (most) likely to be disturbed.

<table>
<thead>
<tr>
<th>Hotspot of response variable</th>
<th>Land allocation</th>
<th>Concessions</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Area for other uses</td>
<td>Conversion forest</td>
</tr>
<tr>
<td>Random distribution</td>
<td>24</td>
<td>1</td>
</tr>
<tr>
<td>ACD</td>
<td>&lt;1</td>
<td>0</td>
</tr>
<tr>
<td>SCD</td>
<td>40</td>
<td>7</td>
</tr>
<tr>
<td>TAD</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

of SCD hotspots were situated in peatlands. For almost 4 years now, there has been a moratorium on the issuance of new licenses for logging and conversion to other land uses in natural forests and peatlands. Whether the moratorium is effective in reducing the deforestation rate is highly debated (Busch et al., 2015; Sloan et al., 2012). Nevertheless, the concessions we consider here were granted prior to 2012, and designated areas are therefore susceptible to being legally selectively logged and/or cleared at any time.

We did not rank concession types according to their impacts on ecosystems but do acknowledge that all activities are not equally damaging. There is a general consensus about the deleterious effect on biodiversity and ES of forest conversion – especially in peatlands – to monocrop plantations (e.g. see Koh and Wilcove, 2008; Savilaakso et al., 2014). In comparison, logged-over forests have been shown to maintain relatively high levels of services (Putz et al., 2012) provided harvesting techniques that reduce the negative impact of logging on biodiversity and ES (e.g. ‘reduced-impact logging’ techniques) are used (Edwards et al., 2014).

4.3. Implications for conservation and development

Significant positive relationships between ES do not necessarily imply high spatial congruence between their respective hotspots, and vice versa (Chan et al., 2006). Yet, in our case, we found a significant positive relationship between ACD and TAD, and an above-than-expected-by-chance spatial congruence of hotspots. Prioritizing hotspots of one of ACD or TAD for
conservation would therefore be beneficial to the other (Table 4). However, either type of prioritization would be detrimental to SCD, as would be SCD hotspot targeting for non-target response variables.

The spatial variations in total carbon stocks (i.e. carbon stored in aboveground biomass, belowground biomass, dead wood, litter and soil organic carbon) are mainly determined by variations in soil organic carbon, with stocks in peat forests largely outperforming those of forests on mineral soils (Page et al., 2011; Paoli et al., 2010). If REDD+ projects only try to maximize total carbon stock protection (i.e. focus on peatlands), additional gains for biodiversity conservation appear limited. We stress that REDD+ projects should target hotspots of ACD rather than SCD, which would also benefit TAD but would probably not be the most economically competitive option to counter the high opportunity cost for conversion (Fisher et al., 2011). In our case, preferential areas for REDD+ project development are in conservation areas or watershed protection forests. Yet, directing REDD+ funds toward conservation areas such as national parks is still controversial because of the seemingly lack of additionality (Macdonald et al., 2011). Gaining effective protection through the use of REDD+ funds would nonetheless de facto constitute additionality for those of the national parks where protection is currently missing (the so-called “paper parks”). Until the institutional framework is clarified, targeting high-ACD watershed protection forests that are adjacent to national parks for REDD+ project development could be a first step in hindering the general trend of national park isolation that has been increasing for the last few decades (DeFries et al., 2005).

SCD hotspot protection, on the other hand, could be achieved through the strict enforcement of the current moratorium that has been praised, despite other caveats, as succeeding in peatland protection (Edwards et al., 2012; Sloan et al., 2012). Re-evaluation of permits delivered prior to 2012, especially for mining and plantations that would lead to permanent conversion of peatlands, could provide further benefits to SCD hotspot protection. If old concessions on peatlands were re-allocated, logged-over forests on mineral soils should undoubtedly be spared and allowed to recover as they provide ES similar to those of natural forests (Labrière et al., 2015; Meijaard and Sheil, 2007). Development of plantation concessions should be directed to highly degraded lowland areas on mineral soils. Some plantation concessions (e.g. timber plantations) could even lead to ecological benefits beyond economic ones (Lamb et al., 2005). Plantations of the exotic species Acacia mangium have, for example, been used successfully to restore natural vegetation in degraded Imperata cylindrica grasslands (Kuusipalo et al., 1995).

The future role of oil palm plantations (the vast majority of plantation concessions granted before 2012) in the development of Kapuas Hulu cannot be overlooked. As the extent of oil palm plantations is predicted to triple in Kalimantan by 2020 (Carlson et al., 2013), it is essential that oil palm establishment is carefully planned and plantations well managed, two conditions necessary for plantations to provide not only goods but also services to a certain extent (Sayer et al., 2012). Plantations developed only on highly degraded lands and incorporated within a matrix of land uses related to the traditional swidden system (that produce more services than oil palm plantations; see e.g. Labrière et al., 2015) could allow for development in Kapuas Hulu that would not be detrimental to carbon and biodiversity conservation. However, sound development options will be limited due to the peculiar biophysical characteristics of the area (steep slopes and peats in the outer and inner parts of the study area, respectively) that restrict the range of potential non-harmful human activities.

4.4. Limitations of our predictions and those from existing maps, and perspectives on spatial congruence analysis

Despite the fact that we used the best available data, the main limitations of our predictions arise from: (1) the still restricted size of our measurement dataset, and (2) the reliability of explanatory variable data. We tried to sample at least 2 ha per main vegetation type while an inventory of 4–6 ha would for instance be recommended to assess biomass of the lowland and hill dipterocarp forests of Indonesia with an error margin not higher than 6%–8% (Laumonier et al., 2010). More vegetation and soil sampling will be required to challenge our predictions and gain a better knowledge of carbon and tree species distribution over the study area. Our predictions also depend on the choice and reliability of explanatory variable data. Though allowing for discrimination between open canopy disturbed forest and closed-canopy natural forests, percent tree cover (selected as explanatory variable in ACD and TAD models) alone would probably fail at capturing biomass variations between different forest types of equally high percent tree cover (Houghton et al., 2001). Yet, other variables included in the model, such as altitude, would have served as surrogates for forest type discrimination. WorldClim data, for example, are obtained by interpolating measurements collected over a vast network of weather stations worldwide. Yet, the density of weather stations was extremely low in our study area and over Borneo more generally (Hijmans et al., 2005).

The nature of relationships between our predictions and those from existing maps was highly variable. While significant positive relationships were found between our ACD predictions and those from Saatchi et al. (2011) and Baccini et al. (2012), we found a clear saturation of their predictions (see Fig. 3). Several studies recently challenged the accuracy of predictions from those maps, arguing that they fail to capture some forest carbon patterns both at regional (Mitchard et al., 2014) or pantropical (Avitabile et al., 2015) scale. The saturation effect might be due to poor performance of GLAS (the Geoscience Laser Altimeter System) height estimation in hilly terrain with slope over 10°–15° (about a third of our study area; Hilbert and Schmullius, 2012), asymptotical saturation of NDVI values in high biomass regions (Huete et al., 2002) and limited field sampling in Southeast Asian tropical forests that are structurally different from those of America and Africa (Slik et al., 2013).

While significant negative relationships were found between our predictions of TAD and those from Slik et al. (2009) and Raes et al. (2009), more recent predictions from Raes et al. (2013) are better aligned with ours. Species distribution models used in Raes et al. (2013) were obtained from a larger database of collection records and did not suffer from partial distribution modeling (Raes, 2012), as in earlier predictions.
While hotspot-based approaches are widely used to study spatial congruence between ecosystem services (e.g. see Egoh et al., 2008; Eigenbrod et al., 2010; Locatelli et al., 2014), such approaches might face limitations regarding spatial conservation prioritization. Conversely, the use of alternative tools such as Marxan (Ball et al., 2009) or Zonation (Moilanen et al., 2005), which allow for acquisition cost inclusion and fragmentation limitation for spatial conservation prioritization, has already proved valuable in helping identifying priority areas for joint biodiversity and ES conservation (e.g. see Chan et al., 2006 and Thomas et al., 2013) but exceeded the scope of our biophysical characterization of ES distribution in a data-scarce region. Further research could explore landscape and socioeconomic scenarios, using information about administratively meaningful planning units (e.g. districts or villages), tenure status, designated land allocation and the opportunity cost of conservation.

5. Conclusion

Mapping the distribution of biodiversity and ecosystem services is crucial to guide decision making on land-use planning. Using field measurements and easily accessible explanatory variables, we were able to predict aboveground carbon density, soil carbon density and tree alpha diversity (response variables) over a mostly forested area of northern Borneo. Analyses of the relationships between response variables, and the spatial congruence of and potential threats to their respective hotspots, enabled us to discuss implications of carbon and tree diversity spatial distributions for conservation and development. We stress that prioritizing hotspots of aboveground carbon – and not total carbon – for conservation through financial mechanisms such as REDD+ would be beneficial for tree diversity conservation in the study area even if it might not be the most economically competitive option to counter the high opportunity cost for conversion. The protection of the large carbon pool in peat soils should be achieved by conservation regulations, which the current moratorium on peat conversion in Indonesia is a first step toward. Re-allocating plantation concessions on highly degraded areas could help achieve economic development (and also ecosystem restoration in the case of timber plantations) in Kapuas Hulu without imperiling carbon and biodiversity conservation.

We believe that our approach can be used to predict the spatial distribution of other important ecosystem services in Kapuas Hulu (e.g. water flow regulation) or be transposed in other areas for the purpose of reaching an integrated ecosystem service based approach for land-use planning (Daily and Matson, 2008).

Acknowledgments

We are deeply grateful to villagers from Keluin, Nanga Dua, Belatong, Banua Tengah and Nanga Hovat for their hospitality and help in data collection. We warmly thank Bapak Engkamat for his precious help and support during fieldwork. We are thankful to Bapak Bundani from Polisi Kehutanan who joined the team and helped during surveys conducted in the Betung Kerihun National Park. We are much indebted to Bapak Ismail Rachman from the Herbarium Bogoriense for voucher identification. We are also sincerely grateful to Niels Raes and Ferry Slik for sharing data from previous work. This research received financial support from CIRAD, CoLUPSIA (EU financed project DCI), AusAid (Agreement 63650 with Center for International Forestry Research), ABIES (Ecole Doctorale Agriculture, Alimentation, Biologie, Environnements et Santé) and CRP-FTA (Consortium Research Program on Forests, Trees, and Agroforestry).

Appendix A. Supplementary data

Supplementary material related to this article can be found online at http://dx.doi.org/10.1016/j.gecco.2016.01.005.

References
