

# Continental-Scale Living Forest Biomass and Carbon Stock: A Robust Fuzzy Ensemble of IPCC Tier 1 Maps for Europe

Daniele de Rigo<sup>1,2</sup>, José I. Barredo<sup>1</sup>, Lorenzo Busetto<sup>1</sup>,  
Giovanni Caudullo<sup>1</sup>, and Jesús San-Miguel-Ayanz<sup>1</sup>

<sup>1</sup> European Commission, Joint Research Centre,  
Institute for Environment and Sustainability,  
via E. Fermi 2749, I-21027 Ispra (VA), Italy

<sup>2</sup> Politecnico di Milano, Dipartimento di Elettronica e Informazione,  
via Ponzio 34/5, I-20133 Milano, Italy

**Abstract.** Forest ecosystems play a key role in the global carbon cycle. Spatially explicit data and assessments of forest biomass and carbon are therefore crucial for designing and implementing effective sustainable forest management options and forest related policies. In this contribution, we present European-wide maps of forest biomass and carbon stock spatially disaggregated at 1km x 1km. The maps originated from a spatialisation improvement of the IPCC methodology for estimating the forest biomass at IPCC Tier 1 level (IPCC-T1). Using a categorical map of ecological zones within the mapping technique may originate boundary effects between the ecological zones. This may induce undue artifacts in the outcomes, as evident in previously published maps generated with the IPCC-T1 methodology. Here we present a novel method for IPCC-T1 biomass mapping which mitigates these artifacts. We propose the use of a fuzzy similarity map of the FAO ecological zones computed by estimating the relative distance similarity (RDS) of each grid-cells climate and geography with respect to the FAO ecological zones. A robust ensemble approach was used to merge an array of simple models with spatially distributed fuzzy set-membership. This allowed the boundary artifacts to be reduced, while mitigating the impact of model semantic extrapolation. The chain of semantically enhanced data-transformations is described following the semantic array programming paradigm. Preliminary results obtained from the application of this novel approach are presented along with a discussion of its impact on the derived maps.

**Keywords:** Ecological Zones, Living Forest Biomass, Living Forest Carbon Stock, IPCC Tier 1, Relative Distance Similarity, Robust Fuzzy Ensemble, Semantic Array Programming.

## 1 Introduction

Forest ecosystems play a key role in the global carbon cycle and are considered large and persistent carbon sinks [1]. Spatially explicit data and assessments

of forest biomass and carbon are therefore of great importance for the design and implementation of effective sustainable forest management options and forest related policies. In complex, wide scale problems, the effectiveness of the science-policy interface might depend on the availability, understandability and multiplicity of estimations, even for correctly assessing the extent and sources of critical uncertainties [2–7]. In this context, European-wide maps of forest biomass and carbon stock spatially disaggregated at  $1\text{ km} \times 1\text{ km}$  (LAEA grid) were recently developed [8], exploiting the IPCC Tier 1 level methodology described in [9]. This is a cost-effective approach for biomass and carbon mapping over large geographical regions where gaps in accurate biomass information exist. It is considered appropriate for continental or global studies [10–12]. According to this methodology [8], above ground living biomass  $B_c^{ag}$  for a given cell  $c$  of a generic regular grid may be estimated by aggregating the biomass contributions by forest type **typ** [13] – coniferous (*con*) and broadleaved (*bro*) – and management type **man** – semi-natural forest (*sn*) and plantations (*pl*):

$$B_c^{ag} = \sum_{\substack{\text{man} \in \{sn, pl\} \\ \text{typ} \in \{con, bro\}}} B_c^{ag, \text{man}, \text{typ}} \quad (1)$$

The contributions  $B_c^{ag, \text{man}, \text{typ}}$  are determined on the basis of two factors. First, the fraction  $f_c^{\text{man}, \text{typ}}$  of the cell's area covered by each combination  $\{\text{man}, \text{typ}\}$ . Second, the average above-ground biomass  $b_c^{ag, \text{man}}$  associated to semi-natural forest and plantations in the ecological zone (e.g. boreal, temperate, subtropical; mountain vs. steppe etc.; from hereafter, EZ) to which  $c$  belongs (see tab. 1):

$$B_c^{ag, \text{man}, \text{typ}} = b_c^{ag, \text{man}} \cdot f_c^{\text{man}, \text{typ}} \quad (2)$$

Below-ground biomass  $B_c^{bg}$  and carbon stock (above-  $C_c^{ag}$  and below-ground  $C_c^{bg}$ ) can be successively estimated as:

$$B_c^{bg} = \sum_{\substack{\text{man} \in \{sn, pl\} \\ \text{typ} \in \{con, bro\}}} B_c^{ag, \text{man}, \text{typ}} \cdot \lambda_c^{\frac{bg}{ag}} = B_c^{ag} \cdot \lambda_c^{\frac{bg}{ag}} \quad (3a)$$

$$C_c^{ag} = \sum_{\substack{\text{man} \in \{sn, pl\} \\ \text{typ} \in \{con, bro\}}} B_c^{ag, \text{man}, \text{typ}} \cdot \lambda_c^{\frac{C}{B}, \text{typ}} \quad (3b)$$

$$C_c^{bg} = \sum_{\substack{\text{man} \in \{sn, pl\} \\ \text{typ} \in \{con, bro\}}} B_c^{ag, \text{man}, \text{typ}} \cdot \lambda_c^{\frac{C}{B}, \text{typ}} \cdot \lambda_c^{\frac{bg}{ag}} = C_c^{ag} \cdot \lambda_c^{\frac{bg}{ag}} \quad (3c)$$

where  $\lambda_c^{\frac{bg}{ag}}$  is the ratio between the below-ground and the above-ground living forest biomass and  $\lambda_c^{\frac{C}{B}, \text{typ}}$  is the ratio between tons of living forest carbon and tons of biomass (carbon fraction). Approximately,  $\lambda_c^{\frac{C}{B}, \text{typ}}$  describes both above-ground and below-ground conditions, depending on the forest type. Both factors vary with the EZ to which  $c$  belongs (tab. 1).

For their pan-European IPCC-T1 application, [8] used as inputs the CORINE land cover map of 2006 [14] to determine the fraction of surface occupied by coniferous and broadleaved forests within  $1\text{ km} \times 1\text{ km}$  cells of a pan-European grid. Each cell was associated with an EZ, following the map of Global Ecological Zones for the Global Forest Reporting [15, 16]. Biomass and carbon stock for each cell were finally computed using eqs. 1–3. This method relies on the simplistic assumption that within a given EZ the coefficients of tab. 1 are constant.

**Table 1.** Above-ground forest biomass, ratio below- vs. above-ground biomass and carbon fraction of forest biomass in European semi-natural forests or plantations. Values from [17] (from [9] where marked with \*); d.u. denotes dimensionless units.

Ecological zone $ez$	$b^{ag,sn,ez}$ [ $\frac{\text{tonn}}{\text{ha}}$ ]	$b^{ag,pl,ez}$ [ $\frac{\text{tonn}}{\text{ha}}$ ]	$\lambda_{ag}^{bg,ez}$ d.u.	$\lambda_{B}^{C,con,ez}$ d.u.	$\lambda_{B}^{C,bro,ez}$ d.u.
Subtropical humid forest	180	180	0.24	0.47	0.47
Subtropical dry forest	130	90	0.28	0.47	0.47
Subtropical steppe	130*	$\begin{matrix} con : 20 \\ bro : 80 \end{matrix}$	0.28*	0.47	0.47
Subtropical mountain system	135	95 (40-150)	0.27*	0.47	0.47
Temperate oceanic forest	120	200	0.29	0.51	0.48
Temperate continental forest	120	200	0.29	0.51	0.48
Temperate steppe	120*	n/a	0.29	0.51	0.48
Temperate mountain system	130	200	0.29	0.51	0.48
Boreal coniferous forest	50	40	0.39	0.51	0.48
Boreal tundra woodland	50*	25	0.39	0.51	0.48
Boreal mountain system	45	40	0.39	0.51	0.48
Polar	50*	25*	0.39*	0.47*	0.47*

The variability of the estimated quantities within each EZ is therefore only driven by the varying percentage of forest area within the cells, and evident artefacts in the output maps are present on the ecozones' boundaries (see also fig. 3).

In order for these effects to be mitigated, we propose an innovative fuzzy similarity map of EZs as an alternative input dataset so as to associate the different areas of Europe with more consistent average biomass amounts and carbon fraction coefficients. This fuzzy set map was computed by estimating the relative distance similarity (RDS) [18] of each grid-cells climatic and geographic conditions from the distribution of values typically observed within the FAO EZs. An unsupervised robust ensemble approach was used to merge an array of simple models with spatially distributed fuzzy set-membership. This allowed the boundary artifacts to be reduced, while mitigating the impact of model semantic extrapolation. The chain of semantically enhanced data-transformation models (D-TM) is described following the *semantic array programming* paradigm [19, 20].

The theoretical basis and the preliminary results obtained from the application of this novel approach are presented along with a discussion of its impact on the derived maps.

## 2 Materials and Methods

The methodology to determine the fuzzy-set memberships of each  $1\text{ km} \times 1\text{ km}$  spatial cell (on a LAEA regular grid, INSPIRE [21, 22] compliant) to each EZ is briefly explained in the next subsection. It is based on a concise array-based mathematical notation for a series of D-TMs, implemented as array programming [23] modules in the GNU Octave/MATLAB programming language [24]. The adopted paradigm of semantic array programming [19, 20] is supported by the Mastrave library and is defined by a modularisation of autonomous tasks with an intense effort toward their most concise generalization and reusability. Each module is semantically constrained with array-oriented compact invariants, pre- and post-conditions. Some straightforward constraints are exemplified in the following with the notation **::constraint::** and an associated active link.

**From Crisp Ecological Zones to a Fuzzy Similarity Map.** FAO produced an expert-based EZ map considering not only bioclimatic variables, but also potential and national vegetation maps [16]. The FAO classification does not match exactly with the Holdridge one, based exclusively on bioclimatic data [25], and in general is not meant to be fully reproduced as a computational model.

FAO EZs embed a broad semantic understanding of climatic specificities such as the mass elevation effect (MEE), for which altitude influences vegetation less “on larger mountains than on smaller ones” [16]. Our method exploits the FAO crisp map as input for training a fuzzy EZ classification with a set of  $N^C$  bioclimatic covariates (tab. 2). Besides temperature and precipitation derived factors, geographic ones are also used – elevation range and solar irradiation – so as to better discriminate e.g. polar, alpine and boreal EZs. Evapotranspiration is not used [30]. MEE is considered with neighbourhood analysis (spatial average moving windows) in two covariates. For each ecozone  $ez$ , the fuzzy set-membership map  $\mu_c^{ez}$  is derived in three steps with intermediate D-TMs (eqs. 4a-c).

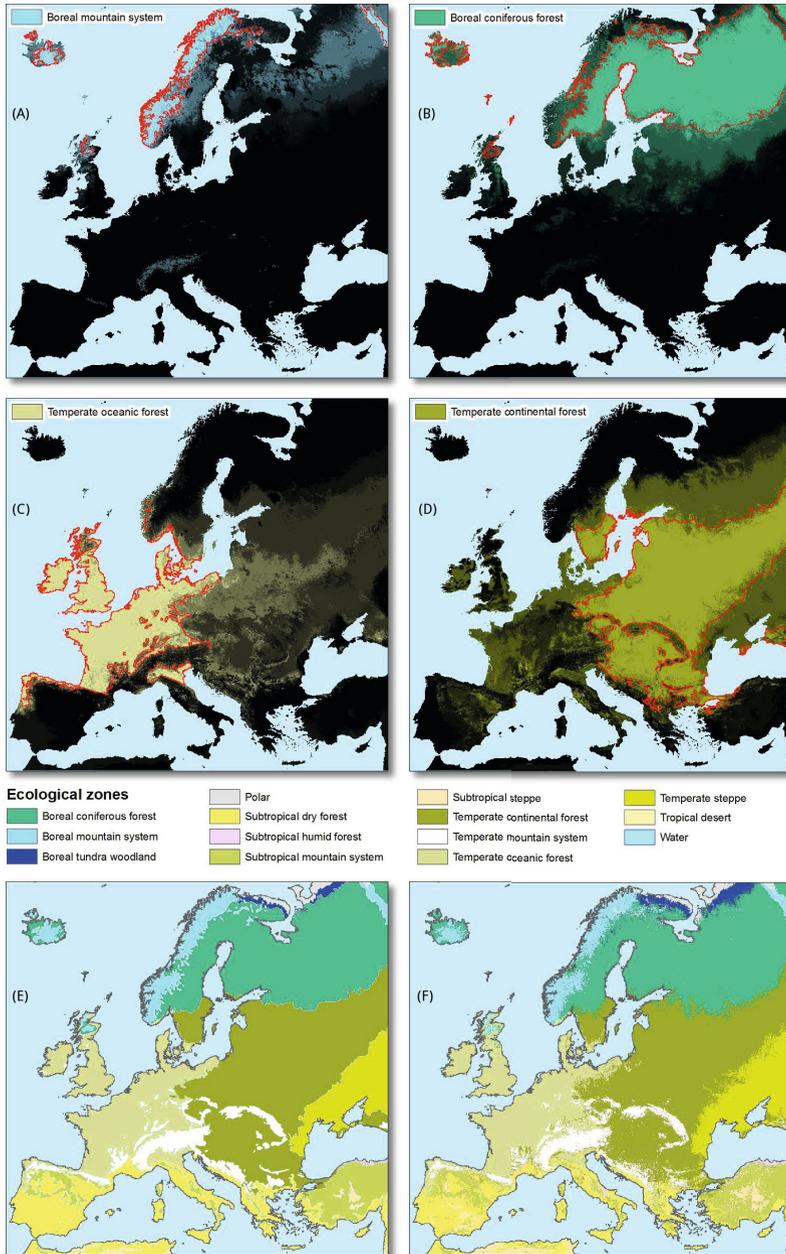
The spatial extent  $\mathcal{U}^S$  is partitioned in  $N^{EZ}$  crisp EZ sets  $\mathcal{S}^{ez}$ ,  $ez \in \mathcal{U}^{EZ} \equiv [1, N^{EZ}]$ . However, an unclassified spatial buffer of 2 km along the boundary between different EZs accounts for the rasterisation error and leads to the set of classified grid cells  $\tilde{\mathcal{U}}^S = \bigcup_{ez} \mathcal{S}^{ez} \subseteq \mathcal{U}^S$ . The RDS analysis [18] (Mastrave library) associates to each pair composed by a set  $\mathcal{S}^{ez}$  and a spatial cell  $c$  a similarity score  $RDS_c^{ez}$ . Given a training set  $\mathcal{U}^S \subseteq \tilde{\mathcal{U}}^S$  of classified cells,  $RDS_c^{ez}$  expresses the **::possibility::**<sup>1</sup>  $\in [0, 1]$  for  $c$  to belong to  $ez$  based on the similarity of each  $i$ -th **::nonnegative::**<sup>2</sup> bioclimatic covariate  $C_c^i$  from the covariate’s values in  $\mathcal{S}^{ez} \cap \mathcal{U}^S$ , aggregated  $\forall i \in [1, N^C]$  with the RDS operator  $\Omega$ :

$$RDS_c^{ez} = \max_{\nu \in \mathcal{S}^{ez}} \Omega_{i \in [1, N^C]} \frac{\min(C_c^i, C_\nu^i)}{\max(C_c^i, C_\nu^i)} \quad (4a)$$

The fuzzy set-membership  $\mu_c^{ez}$  of  $c$  to  $\mathcal{S}^{ez}$  is estimated by means of a robust ordinal algorithm. The algorithm has been designed so as for  $\mu_c^{ez}$  to be

<sup>1</sup> [http://mastrave.org/doc/mtv\\_m/check\\_is#SAP\\_possibility](http://mastrave.org/doc/mtv_m/check_is#SAP_possibility)

<sup>2</sup> [http://mastrave.org/doc/mtv\\_m/check\\_is#SAP\\_nonnegative](http://mastrave.org/doc/mtv_m/check_is#SAP_nonnegative)



**Fig. 1.** Preliminary maps of fuzzy set-membership (A-D) of four ecological zones EZ, and comparison between the original FAO map (E) and a defuzzified map obtained by selecting among the fuzzy EZ sets the dominant (the highest) value of set-membership (F), for each spatial cell. Red lines denote the crisp boundary of the original FAO EZs.

**Table 2.** Bioclimatic variables and derived covariates  $C_c^i$  used for computing the fuzzy set-membership  $\mu_c^{ez}$  of each spatial grid cell  $c$  to each ecological zone  $ez$ . Quantities averaged with a spatial moving window of  $3\text{ km}\times 3\text{ km}$  are denoted with  $^{3\times 3}$ . Temperature in degrees Celsius; precipitation in  $\text{mm}$ ; solar irradiations in  $\text{W h m}^{-2}$ ; elevation range in  $\text{m}$ .  $T_c$ : annual avg. temperature.  $T_c^m$ ,  $[T]_c^m$ ,  $[T]^m_c$ : avg., min. and max. temperature of the month  $m$ .  $P$ ,  $P_c^m$ : annual total precipitation and that of the month  $m$ .

Variable code	Description of the bioclimatic variable	Reference equation
$S_c^{\text{light}}$	Total potential solar irr. of the 6 lighter months [26]	
$S_c^{\text{dark}}$	Total potential solar irr. of the 6 darker months [26]	
$T_c^{\text{cold}}$	Average temperature of the coldest month	$\min_m T_c^m$
$T_c^{\text{warm}}$	Average temperature of the warmest month	$\max_m T_c^m$
$\Delta T_c$	Mean of monthly temperature range	$1/12 \sum_m [T]_c^m - [T]_c^m$
$T_c^{\text{tundra}}$	Nordenskiöld index [27]	$T_c^{\text{warm}} + 0.1T_c^{\text{cold}} - 9$
$P_c^{\text{dry}}$	Total precipitation of the driest month	$\min_m P_c^m$
$P_c^{\text{wet}}$	Total precipitation of the wettest month	$\max_m P_c^m$

Covariate code	Description of the bioclimatic covariate $C_c^i, i \in [1, 12]$	Reference equation
1. $[\Delta S]_c^{3\times 3}$	$3\text{ km}\times 3\text{ km}$ seasonal variation of pot. solar irr. (normalised by semestral minimum)	$\frac{S_c^{\text{light},3\times 3} - S_c^{\text{dark},3\times 3}}{S_c^{\text{dark},3\times 3}}$
2. $S_c$	Annual potential solar irradiation	$S_c^{\text{light}} + S_c^{\text{dark}}$
3. $eT_c^{\text{tundra}}$	Nordenskiöld exponential index	$\exp(T_c^{\text{tundra}})$
4. $T_c^{+100}$	Shifted average annual temperature	$T_c + 100\text{ }^\circ\text{C}$
5. $T_c^{\text{cold}+100}$	Shifted average temp. of the coldest month [16]	$T_c^{\text{cold}} + 100\text{ }^\circ\text{C}$
6. $[\Delta P]_c$	Seasonal variation of monthly precipitation (normalised by minimum: driest month)	$\frac{P_c^{\text{wet}} - P_c^{\text{dry}}}{P_c^{\text{dry}}}$
7. $[\Delta P]_c$	Seasonal variation of monthly precipitation (normalised by maximum: wettest month)	$\frac{P_c^{\text{wet}} - P_c^{\text{dry}}}{P_c^{\text{wet}}}$
8. $T^{100}P_c$	Shifted temperature per rain magnitude order	$T_c^{+100} / \log_{10}(P_t)$
9. $ER_c^{3\times 3}$	$3\text{ km}\times 3\text{ km}$ average elevation range	
10. $N_c^{m \geq 10\text{ }^\circ\text{C}}$	Number of months with $T_c \geq 10\text{ }^\circ\text{C}$ [16]	$\sum_m [T_c^m \geq 10]$
11. $N_c^{m \text{dry}}$	Number of dry months [16]	$\sum_m [P_c^m \leq 2T_c^m]$
12. $S\Delta T_c$	Potential irradiation by montly temp. range (simplified avg. monthly solar entransy flux [28, 29])	$S_c \cdot \Delta T_c$

invariant to any monotonically non-decreasing, rank-preserving transformation of the RDS scores. This means that the sole semantic requirement for the similarities  $RDS_c^{ez}$  is to be **::sortable::**<sup>3</sup>. The second D-TM implements this mathematical invariance by computing an intermediate quantity,  $\eta_c^{ez}$ . The EZs whose similarity is above the median are ranked and the ranking is scaled in  $[0, 1]$  so that the maximum among the scaled values  $\eta_c^{ez}$  is 1. This mathematically describes the semantic concept of ensuring each spatial cell  $c$  to belong with certainty to at least one EZ  $ez^{*,c}$  (i.e.  $c$  is in the kernel of  $\eta_c^{ez^{*,c}}$ ). In case two or more EZs in  $c$  share the maximum RDS score,  $c$  would belong to the kernel of each of them. The median threshold also ensures  $c$  not to belong to the farthest EZs  $ez_i^{\theta,c}$ . For example, it would otherwise be surprising for a spatial cell in the

<sup>3</sup> [http://mastrave.org/doc/mtv\\_m/check\\_is#SAP\\_sortable](http://mastrave.org/doc/mtv_m/check_is#SAP_sortable)

subtropical dry forest to be also substantially characterised by the polar zone. The described array transformation  $f^\eta : RDS \mapsto \eta$  is equivalent to the following:

$$\begin{bmatrix} \dots \\ \eta_c^{ez} \\ \dots \end{bmatrix} = f^\eta(RDS_c) = \frac{\begin{bmatrix} \dots \\ \max \left( 0, \sum_{\eta \in \mathcal{U}^{EZ}, \eta \neq ez} [RDS_c^{ez} \geq RDS_c^\eta] - \left\lfloor \frac{N^{EZ} - 1}{2} \right\rfloor \right) \\ \dots \end{bmatrix}}{\begin{bmatrix} N^{EZ} - 1 - \left\lfloor \frac{N^{EZ} - 1}{2} \right\rfloor \\ \dots \end{bmatrix}} \quad (4b)$$

where the Iverson bracket notation is used [31]. Depending on the number of EZs, the relative importance of the best score  $\eta_c^{ez^*}$  may be noticeably diluted by the remaining ones. For example, considering 12 EZs and the ranking  $r$  of the 6 best fitting ones (i.e. the above-median ones),  $\eta_c^{ez^*}$  would only account for 29% of the total ( $6 / \sum_{r=1}^6 r$ ). The third D-TM stabilises this behaviour by imposing the weight of the dominant EZ to be 50% with a simple power transformation:

$$\mu_c^{ez} = (\eta_c^{ez})^k \quad \text{such that} \quad 2 \cdot (N^{EZ})^k = \sum_{r=1}^{N^{EZ}} r^k \quad (4c)$$

**Table 3.** Comparison between the crisp FAO EZ  $\mathcal{S}^{ez}$  and their proposed extended fuzzy version  $\mu_c^{ez}$ . A subset  $\mathcal{U}^S$  of the classified grid cells has been used, uniformly distributed in  $\mathcal{U}^S$  with approximately  $1/16^{th}$  of the available cells. The number of cells of  $\mathcal{U}^S$  within (*true set*) and outside (*false set*) a given  $\mathcal{S}^{ez}$  is listed along with the mean square error (for all cells; only *true set* cells; only *false set* cells) between  $\mu_c^{ez}$  and  $\mathcal{S}^{ez}$ .

Ecological zone $ez$	Overall	Within the FAO EZ	Outside the FAO EZ		
	mse $c \in \mathcal{U}^S$	$\sum_{c \in \mathcal{U}^S} [c \in \mathcal{S}^{ez}]$	mse $c \in \mathcal{U}^S \cap \mathcal{S}^{ez}$	$\sum_{c \in \mathcal{U}^S} [c \notin \mathcal{S}^{ez}]$	mse $c \in \mathcal{U}^S \setminus \mathcal{S}^{ez}$
Subtropical humid forest	0.0072	748	0.009	566558	0.0072
Subtropical dry forest	0.026	63836	0.063	503470	0.022
Subtropical steppe	0.031	3246	0.0027	564060	0.032
Subtropical mountain system	0.039	30308	0.037	536998	0.04
Temperate oceanic forest	0.068	79124	0.018	488182	0.076
Temperate continental forest	0.11	164777	0.021	402529	0.15
Temperate steppe	0.041	31513	0.0031	535793	0.043
Temperate mountain system	0.042	26715	0.054	540591	0.041
Boreal coniferous forest	0.055	134513	0.014	432793	0.068
Boreal tundra woodland	0.029	3515	0.0056	563791	0.029
Boreal mountain system	0.039	24809	0.029	542497	0.04
Polar	0.015	4202	0.0089	563104	0.015

**Fuzzy Ecozones for Biomass and Carbon Stock Ensembles.** The final D-TM of eq.4c ensures  $\mu_c^{ez}$  to show a series of desirable properties for unsupervised fuzzy ensemble applications. First, the aforementioned semantic properties of  $\eta_c^{ez}$  also hold for  $\mu_c^{ez}$ . Second, a weighted median of quantities of choice (e.g.

EZ biomasses or carbon stocks), done using as weights the fuzzy scores  $\mu_c^{ez}$ , would always select the quantity corresponding to the dominant EZ in  $c$ . This provides a very intuitive and coherent way for estimating defuzzified (crisp) ensembles: selecting a very robust ensemble operator such as the weighted median automatically implies a crisp ensemble.

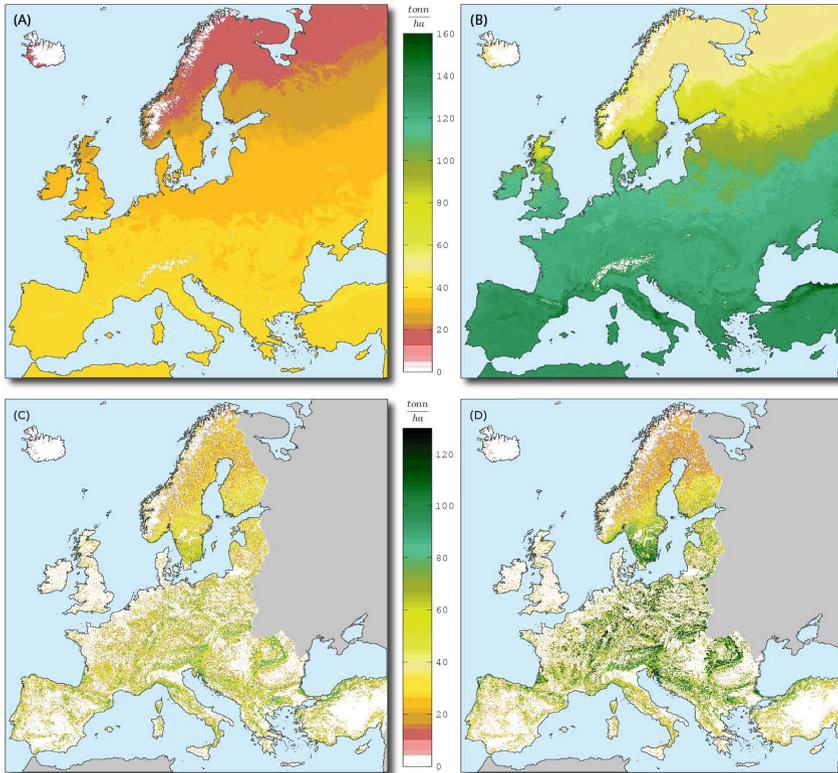
On the other hand, a weighted average would instead use all the above-median best fuzzy scores and would merge the corresponding associated quantities (but not the ones having worst scores, thus automatically behaving as a robust statistics). Classic intermediate robust statistics as the trimmed mean would show intermediate behaviour, as the intuition would expect. Furthermore, jack-knife resampling techniques might safely be applied by iteratively removing one fuzzy EZ per time. In the worst case, the removal would decrease the overall sum of remaining  $\mu_c^{ez}$  to not less than 50%, thus ensuring a quite robust statistical resampling. Fig. 2 shows the application of weighted average ensembles to estimate forest living biomass and carbon stock maps.

### 3 Results and Discussion

Maps of fuzzy set memberships obtained for four EZs are shown as an example in fig. 1 (A-D). The similarity with a given EZ can be nonnegligible even very far from the original FAO spatial domain of that zone (the red boundary in the maps). The reconstructed crisp EZ map obtained by selecting the dominant fuzzy EZ agrees with the FAO map where sudden spatial climatic changes occur (e.g. mountains) but highlights areas with smoother transitions (fig. 1 E and F). Results obtained by using the fuzzy EZ maps as the basis for the computation of forest biomass and carbon stock are shown in fig. 2. In particular, fig. 2 (A,B) show the average amount of below-ground and above-ground biomass per unit surface ( $b_c^{ag,sn}$ ) estimated across Europe for semi-natural forests (irrespective of the current  $f_c^{sn,typ}$ ) by transforming the fuzzy EZs into the corresponding fuzzy biomass models. It must be stressed that, without the use of the fuzzy EZ maps, computing the two maps with IPCC-T1 methods would have led to a piecewise constant approximation, with localised sudden spatial changes along the EZ boundaries.

Fig. 2 (C,D) show instead the total living carbon stock of forests ( $C_c^{ag} + C_c^{bg}$ ) estimated with eqs. 1–3 by taking into account the membership values for the different EZs. Due to the lack of gridded high-resolution information regarding the management type of forests in Europe, the maps represent two very different scenarios. The first scenario (fig. 2 C) assumes all forests to be semi-natural, while in the second one all forests are assumed as plantations (fig. 2 D). An example of the improvements introduced by the use of the fuzzy EZ maps for the computation of the IPCC-T1 biomass maps is finally shown in fig. 3.

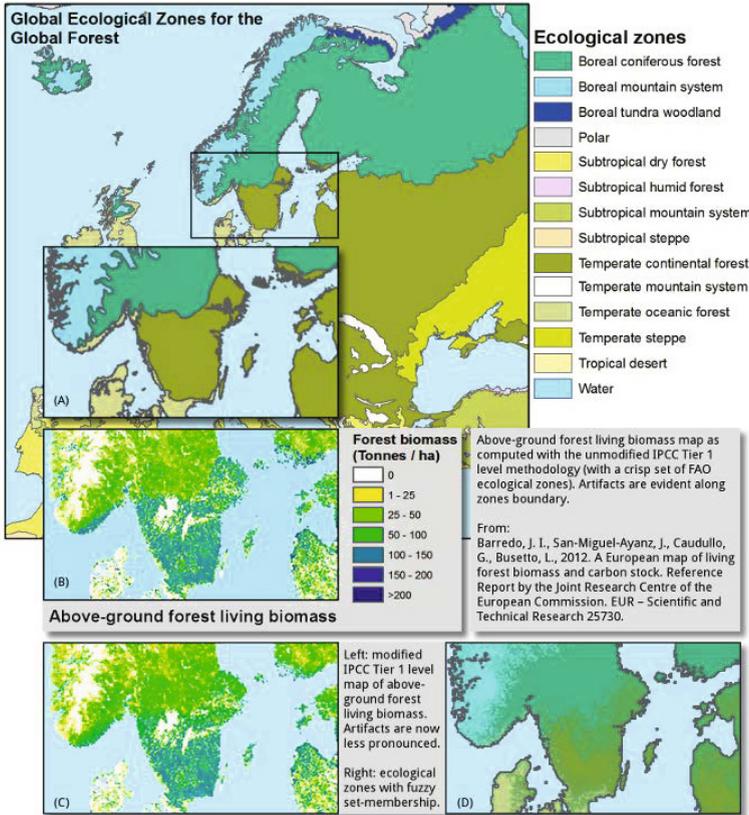
The detailed analysis referring to the southern part of Norway and Sweden highlights how the evident artefacts introduced in the above-ground forest living biomass by [8] and related to the crisp boundaries between EZs (see panel B) are substantially reduced in the improved map (see panel D). Although preliminary,



**Fig. 2.** Improved maps obtained exploiting the fuzzy EZ maps as an input for the IPCC TIER-1 mapping methodology. (A,B) Average amount of below ground ( $b_c^{bg,sn}$ ) and above-ground ( $b_c^{ag,sn}$ ) biomass per unit surface for semi-natural forests. (C,D) Total living carbon stock of forests ( $C_c^{ag} + C_c^{bg}$ ) estimated considering all forests as semi-natural (C), and all forests as plantations (D).

these results suggest the proposed methodology for deriving fuzzy EZ maps to be beneficial in mitigating undue artifacts in IPCC-T1 large-scale biomass and carbon mapping. It must however be remembered that, although the quantities are disaggregated at  $1\text{ km} \times 1\text{ km}$  spatial resolution in the resulting maps, the intrinsic limitations of the IPCC-T1 methodology (i.e., the use of constant coefficients to describe the typical characteristics of forests over huge spatial extents) suggest that local patterns should be considered as merely indicative.

Nevertheless, the method might prove helpful for interpolating spatial gaps in more detailed maps. The exemplified fuzzy ensemble modelling strategy – whose D-TM detailed computational description [32] is reproducible [33–35] and entirely based on free software [36] and data [37, 38] – lies within a family of semantically-enhanced ensemble techniques [39–42] which might also serve as reference for merging multiple partially independent estimations of biomass and carbon stock [43] and deriving new improved maps. Several factors, neglected in



**Fig. 3.** Example of improvements introduced with the use of the fuzzy EZ maps for the IPCC-T1 modelling of above-ground forest living biomass in Europe. The area surrounding the southern part of Norway and Sweden is highlighted as it provides an example of artifacts (B) generated along the sharp boundary between FAO ecological zones as a result of the implementation of the unmodified IPCC-T1 [8]. The artefacts are substantially reduced in the improved map (C), thanks to the fuzzy approach used for the determination of the several coefficients used for biomass computation. (Note: The combined spatial fuzzy set-membership is here approximately rendered (D) by merging the colours associated to the different EZs in the legend of the FAO map).

a simplified IPCC-T1 approach, might be considered [44–46]. The disaggregation by forest type might be improved by introducing information on pan-European forest tree species distribution [47]. While fragmentation is correlated with carbon stock alteration [50, 51], its relationship with forest cover is complex [48, 49]. The post-disturbance regeneration of vegetation might sometimes preserve the fractions  $f_c^{\text{man,typ}}$ , but it can alter biomass and carbon stock [50, 52]. In Europe, around 65000 wildfires per year burn approximately 500,000 hectares of wildland and forests [53]. The effect of fires and other disturbances also impacts on the canopy cover and thus on the non-living component of the carbon

stock – here not modelled – by means of soil erosion [54–56]. Recent advances in European soil data [57], soil erosion [58,59] and its relationship with wildfires might help to complement the analysis of changes in living and non-living forest carbon [41,42].

## 4 Conclusions

A pan-European similarity analysis has been performed – at 1x1 km spatial grid resolution – by comparing each grid-cell climate and geography with respect to the FAO ecological zones. A rank-based mapping of the relative distance similarity (RDS) for each spatial cell  $c$  to the eco-zones EZ in the crisp classification system of FAO led to a robust fuzzy ecological classification. The resulting array of simple models offered a spatially distributed fuzzy set-membership of the ecological zones which extends the FAO system in the European continent by moving from crisp categorical classes to a semantic fuzzy composition of them.

The use of the derived novel fuzzy EZ maps as an input dataset for the computation of pan-European forest biomass and carbon maps using the IPCC Tier 1 methodology highlighted their usefulness for modelling applications relying on EZ maps as categorical input datasets. A noticeable characteristic of the modelling approach for deriving the fuzzy EZ maps is also that it entirely relies on geographic and climate information routinely estimated by global and regional climate models. This makes the method also suitable for modelling applications in which the spatial shift of ecological zones and of eco-zone dependant quantities need to be assessed under climate change scenarios.

## References

1. Pan, Y., Birdsey, R.A., Fang, J., Houghton, R., Kauppi, P.E.: A Large and Persistent Carbon Sink in the Worlds Forests. *Science* 33(6045), 988–993 (2011)
2. Nabuurs, G.J., van Putten, B., Knippers, T.S., Mohren, G.M.J.: Comparison of uncertainties in carbon sequestration estimates for a tropical and a temperate forest. *Forest Ecology and Management* 256(3), 237–245 (2008)
3. Bonan, G.B.: Forests and climate change: Forcings, feedbacks, and the climate benefits of forests. *Science* 320(5882), 1444–1449 (2008)
4. Eurostat: Forestry in the EU and the world: a statistical portrait
5. Seebach, L., McCallum, I., Fritz, S., Kindermann, G., Leduc, S., et al.: Choice of forest map has implications for policy analysis: A case study on the EU biofuel target. *Env. Sci. & Policy* 22, 13–24 (2012)
6. de Rigo, D.: Behind the horizon of reproducible integrated environmental modelling at European scale: ethics and practice of scientific knowledge freedom. Submitted to F1000 Research (2013)
7. Food and Agriculture Organization of the United Nations: Global forest resources assessment 2010: Main report. (2010) ISBN: 9789251066546
8. Barredo, J.I., San-Miguel-Ayanz, J., Caudullo, G., Busetto, L.: A European map of living forest biomass and carbon stock. Reference Report by the Joint Research Centre of the European Commission. EUR – Sci. Tech. Res. 25730 (2012)

9. Ruesch, A., Gibbs, H. K.: New IPCC tier-1 global biomass carbon map for the year 2000. Carbon Dioxide Information Analysis Center, Oak Ridge National Lab (2008)
10. Koh, L.P., Ghazoul, J.: Spatially explicit scenario analysis for reconciling agricultural expansion, forest protection, and carbon conservation in Indonesia. *Proc. Natl. Acad. Sci. U.S.A.* 107(24), 11140–11144 (2010)
11. West, P.C., Gibbs, H.K., Monfreda, C., Wagner, J., Barford, C.C., et al.: Trading carbon for food: Global comparison of carbon stocks vs. crop yields on agricultural land. *Proc. Natl. Acad. Sci. U.S.A.* 107(46), 19645–19648 (2010)
12. Nelson, E., Sander, H., Hawthorne, P., Conte, M., Ennaanay, D., et al.: Projecting global Land-Use change and its effect on ecosystem service provision and biodiversity with simple models. *PLoS ONE* 5(12), e14327+ (2010)
13. Kempeneers, P., Sedano, F., Seebach, L.M., Strobl, P., San-Miguel-Ayanz, J.: Data fusion of different spatial resolution remote sensing images applied to Forest-Type mapping. *IEEE Trans. Geosci. Remote Sens.* 49(12), 4977–4986 (2011)
14. EEA: CORINE land cover technical guide Addendum 2000. European Environment Agency, Technical report No 40, Copenhagen (2000)
15. Food and Agriculture Organization of the United Nations: Global Ecological Zoning for the Global Forest Resources Assessment 2000 - Final Report. Food and Agriculture Organization of the United Nations, Forestry Department, Rome, Italy (2001)
16. Food and Agriculture Organization of the United Nations: Global ecological Zones for FAO forest reporting: 2010 update. *For. Resour. Assess. Work. Paper* 179 (2012)
17. Aalde, H., Gonzalez, P., Gytarsky, M., Krug, T., Kurz, W. A., et al.: Forest Land. IPCC Guidelines for National Greenhouse Gas Inventories, Prepared by the National Greenhouse Gas Inventories Programme. The Intergovernmental Panel on Climate Change (IPCC), vol. 4, ch. 4, p. 83 (2006)
18. de Rigo, D.: Relative distance similarity as multivariate supervised or unsupervised ensemble interpolation (in prep. 2013)
19. de Rigo, D.: Semantic Array Programming for Environmental Modelling: Application of the Mastrave Library. In: *Int. Congress on Environmental Modelling and Software. Managing Resources of a Limited Plant, Pathways and Visions under Uncertainty, Sixth Biennial Meeting*, pp. 1167–1176 (2012)
20. de Rigo, D.: Semantic array programming with Mastrave - introduction to semantic computational modelling (2012)
21. European Parliament: Directive 2007/2/EC of the European Parliament and of the Council of 14 March 2007 establishing an infrastructure for spatial information in the European community (INSPIRE). *Official J. Eur. Union* 50(L 108), 1–14 (2007)
22. European Commission: Commission regulation (EC) no 1205/2008 of 3 December 2008 implementing directive 2007/2/EC of the European Parliament and of the Council as regards metadata. *Official J. Eur. Union* 51(L 326), 12–30 (2008)
23. Iverson, K.E.: Notation as a tool of thought. *Commun. ACM* 23(8), 444–465 (1980)
24. Eaton, J.W.: GNU Octave and reproducible research. *J. Process Contr.* 22(8), 1433–1438 (2012)
25. Holdridge, L. R.: Life zone ecology. Tropical Science Center, Costa Rica. 206p. (1967)
26. Hofierka, J., Suri, M., Huld, T.: GRASS GIS manual: r.sun. In: *GRASS Development Team, 2013. GRASS GIS 6.4.3svn Ref. Manual. Open Source Geospat. Found.* (2007)
27. Nordenskiöld, O., Mecking, L.: The geography of the polar regions. *Amer. Geogr. Soc. Spec. Publ.* 8, 359 (1928)
28. Guo, Z.-Y., Zhu, H.-Y., Liang, X.-G.: Entransy A physical quantity describing heat transfer ability. *Int. J. Heat Mass Transfer* 50(13-14), 2545–2556 (2007)

29. Xu, M.: The thermodynamic basis of entransy and entransy dissipation. *Energy* 36(7), 4272–4277 (2011)
30. Milly, P.C.D., Dunne, K.A.: On the hydrologic adjustment of Climate-Model projections: The potential pitfall of potential evapotranspiration. *Earth Interact* 15(1), 1–14 (2011)
31. Knuth, D.E.: Two notes on notation. *Amer. Math. Monthly* 99(5), 403–422 (1992)
32. de Rigo, D., Barredo, J.I., Busetto, L., Caudullo, G., San Miguel-Ayanz, J.: Extending FAO Ecological Zones with Pan-European Fuzzy Geoclimatic Similarity. *Environmental Data and Data Transformation Models Review* ( in prep., 2013)
33. Mesirov, J.P.: Accessible reproducible research. *Science* 327(5964), 415–416 (2010)
34. Nature: Devil in the details. *Nature* 470(7334), 305–306 (2011)
35. Kleiner, K.: Data on demand. *Nature Climate Change* 1(1), 10–12 (2011)
36. Stallman, R.M.: Viewpoint: Why “open source” misses the point of free software. *Commun. ACM* 52(6), 31–33 (2009)
37. Murray-Rust, P.: Open data in science. *Serials Review* 34(1), 52–64 (2008)
38. Stodden, V.: Trust your science? open your data and code. *Amstat News* 21–22 (2011)
39. de Rigo, D., Guariso, G.: Rewarding Open Science: A Collaborative Review System for Semantically-Enhanced Free Software and Environmental Data Modelling ( in prep., 2013)
40. Bosco, C., de Rigo, D., Dijkstra, T., Sander, G., Wasowski, J.: Multi-Scale Robust Modelling of Landslide Susceptibility: Regional Rapid Assessment and Catchment Robust Fuzzy Ensemble. In: Hřebíček, J., Schimak, G., Kubásek, M., Rizzoli, A. (eds.) ISESS 2013. IFIP AICT, vol. 413, pp. 321–335. Springer, Heidelberg (2013)
41. de Rigo, D., Rodriguez-Aseretto, D., Bosco, C., Di Leo, M., San-Miguel-Ayanz, J.: An Architecture for Adaptive Robust Modelling of Wildfire Behaviour under Deep Uncertainty. In: Hřebíček, J., Schimak, G., Kubásek, M., Rizzoli, A. (eds.) ISESS 2013. IFIP AICT, vol. 413, pp. 367–380. Springer, Heidelberg (2013)
42. Di Leo, M., de Rigo, D., Rodriguez-Aseretto, D., Bosco, C., Petroliaqkis, T., Camia, A., San-Miguel-Ayanz, J.: Dynamic Data Driven Ensemble for Wildfire Behaviour Assessment: A Case Study. In: Hřebíček, J., Schimak, G., Kubásek, M., Rizzoli, A. (eds.) ISESS 2013. IFIP AICT, vol. 413, pp. 11–22. Springer, Heidelberg (2013)
43. Mitchard, E.: A Comparison of Tropical Carbon Maps. *Ecometrica, Natural Environment Research Council (NERC)* (2013)
44. Ciais, P., Schelhaas, M.J., Zaehle, S., Piao, S.L., Cescatti, A., et al.: Carbon accumulation in european forests. *Nature Geosci.* 1(7), 425–429 (2008)
45. Mascarelli, A.: Earth’s carbon sink downsized. *Nature* (2012)
46. Fiorese, G., Guariso, G.: Modeling the role of forests in a regional carbon mitigation plan. *Renewable Energy* 52, 175–182 (2013)
47. de Rigo, D., Caudullo, G., Amatulli, G., Strobl, P., San-Miguel-Ayanz, J.: Modelling tree species distribution in Europe with constrained spatial multi-frequency analysis (in prep.)
48. Estreguil, C., Caudullo, G., de Rigo, D., Whitmore, C., San-Miguel-Ayanz, J.: Reporting on European forest fragmentation: Standardized indices and web map services. *IEEE Earthzine* 5(2), 384031+ (2012); 2nd quarter theme: Forest Resource Information
49. Estreguil, C., Caudullo, G., de Rigo, D., San-Miguel-Ayanz, J.: Forest landscape in Europe: Pattern, fragmentation and connectivity. *EUR – Sci. Tech. Res. (JRC 77295)* (2013)
50. Malhi, Y., Baldocchi, D.D., Jarvis, P.G.: The carbon balance of tropical, temperate and boreal forests. *Plant, Cell and Environment* 22(6), 715–740 (1999)
51. Houghton, R.: Aboveground forest biomass and the global carbon balance. *Global Change Biology* 11(6), 945–958 (2005)
52. Myneni, R.B., Dong, J., Tucker, C.J., Kaufmann, R.K., Kauppi, P.E., et al.: A large carbon sink in the woody biomass of northern forests. *Proc. Natl. Acad. Sci. U.S.A.* 98(26), 14784–14789 (2001)

53. Rodriguez-Aseretto, D., de Rigo, D., Di Leo, M., Cortés, A., San-Miguel-Ayanz, J.: A data-driven model for large wildfire behaviour prediction in Europe. *Procedia Computer Science* 18, 1861–1870 (2013)
54. Lal, R.: Soil erosion and the global carbon budget. *Env. Intl.* 29(4), 437–450 (2003)
55. Lal, R.: Forest soils and carbon sequestration. *For. Ecol. Manage.* 220, 242–258 (2005)
56. Hurteau, M.D., Brooks, M.L.: Short- and long-term effects of fire on carbon in US dry temperate forest systems. *BioScience* 61(2), 139–146 (2011)
57. Panagos, P., Jones, A., Bosco, C., Senthil Kumar, P.S.: European digital archive on soil maps (EuDASM): Preserving important soil data for public free access. *Int. J. of Digital Earth* 4(5) (2011)
58. de Rigo, D., Bosco, C.: Architecture of a pan-european framework for integrated soil water erosion assessment. In: Hřebíček, J., Schimak, G., Denzer, R. (eds.) *ISESS 2011. IFIP AICT*, vol. 359, pp. 310–318. Springer, Heidelberg (2011)
59. Bosco, C., de Rigo, D., Dewitte, O., Poesen, J., Panagos, P.: *Modelling Soil Erosion at European Scale. Towards Harmonization and Reproducibility* (in prep.)