

Disentangling the effects of vegetation and water on the satellite observations of soil organic carbon stocks in western European topsoils

Lixin Lin^{1,*}, Xixi Liu², Yuan Sun³

5 ¹School of Remote Sensing and Geomatics Engineering, Nanjing University of Information Science and Technology, Nanjing, 210044, China

²College of Information Science and Engineering, Key Laboratory of Grain Information Processing and Control Ministry of Education, Henan University of Technology, Zhengzhou, 450001, China

³College of Surveying and Geo-Informatics, Tongji University, Shanghai, 200092, China

10

Correspondence to: Lixin Lin (llxshxdhl8@nuist.edu.cn; llxshxdhl8@gmail.com)

Abstract. The performance of models based on satellite observations of soil organic carbon (SOC) stock in European soils is seriously limited by the complexity of natural land surfaces. Therefore, disentangling the SOC stock from other natural land surfaces including vegetation and water bodies has become a rather difficult but necessary task. This study proposed a novel and promising approach intended to resolve this frustrating problem. Based on a series of spectral narrowing, unchanging, and enlarging processes, 23,914,845 sets of SOC models were developed both for vegetation fuzzy disentangling and water fuzzy disentangling. The optimal model was obtained through comparison and was determined as the model that ultimately performed obviously better than **the model using unfuzzified spectra**. This model simulated the per-unit and total SOC stocks in western European topsoils as 99.742 t C ha⁻¹ and 9.373 Pg, respectively. In comparison with the results of previous studies, the gaps in the simulated per-unit SOC stocks across the western European countries were considerably narrower (83.673–104.334 t C ha⁻¹). The outstanding model performance and stable simulated per-unit values are the result of disentangling of the vegetation and water cover. This study proposed a valuable reference solution for disentangling the SOC stock from complex natural land cover.

25 1 Introduction

The amount of carbon stored in Earth's soils is greater than that stored in biomass and the atmosphere (Scharlemann et al., 2014). In recent years, many studies have suggested that environmental degradation and climate change are strongly associated with change of the soil organic carbon (SOC) stock (Crowther et al., 2016; Pries et al., 2017). Therefore, timely observation of SOC stock dynamics is increasingly recognized as of importance. Traditional laboratory-based SOC stock analysis is expensive, laborious, and time consuming. Owing to the advantages of real-time and dynamic tracking, the

30

method of soil remote sensing that exploits the spectral response of SOC has become a potential alternative for observing SOC stock (Thaler et al., 2019).

The implementation of SOC remote sensing depends on a SOC model developed using field SOC data and their corresponding spectral pixels, then we can use this model to estimate the SOC for all pixels of images (Koparan et al., 2022).

35 In recent years, many studies on SOC content/stock determination from spectral reflectance focused on optimal soil conditions, i.e., laboratory sieved and air-dried soil samples or small-area bare, dry, and smooth soils have been reported (Mueller et al., 2021; Liu et al., 2022). For example, Ward et al. (2020) tested SOC models using airborne hyperspectral remote sensing data and simulated satellite Environmental Mapping and Analysis Programme (EnMAP) data as input. Hutengs et al. (2019) examined in-situ spectroscopy and SOC estimation models using in-situ Mid-Infrared (MIR) spectra.

40 The Land Use/Cover Area frame statistical Survey (LUCAS) topsoil database, which is the largest expandable topsoil dataset for the European region (Orgiazzi et al., 2018), has supported several field SOC data for SOC spectroscopic modeling (Ward et al., 2020) and the combination of LUCAS and satellite data has provided tremendous potential for timely observation of SOC stock at the European scale. Based on LUCAS database, for example, Yigini and Panagos (2016) performed a digital soil mapping for European SOC, using climate, land cover, terrain, and soil covariates. Lugato et al.

45 (2014) used an agro-ecosystem SOC model to calculate European SOC stocks using soil/climate/land-use/management drivers, among others. However, natural land surfaces comprise heterogenous mixtures of vegetation, water, and soil, and the effects of surfaces such as vegetation and water have marked impact on the SOC spectral response, resulting to the low performance of models in these previous studies. Moreover, the areas of vegetation-free and water-free soil in satellite imagery are extremely limited. To adapt to soils with complex surface cover and to apply SOC remote sensing to large

50 geographical regions, the cover of both vegetation and water needs to be disentangled during SOC modeling.

Therefore, the focus of this study was the challenging but necessary task of how best to disentangle vegetation and water cover during satellite SOC modeling. Fuzzy deep learning has many advantages when facing uncertainty factors (Tscherko et al., 2007). A study by Lin and Liu (2022) indicated that changing the soil spectral reflectance of soil samples using corresponding soil moisture indexes was helpful for laboratory-based spectroscopic modeling of SOC content. Although

55 their study did not consider that the effects of soil moisture on each spectral band should be different, their adoption of the spectral fuzzy learning approach inspired our research. Hence, our study investigated the issue of “whether the satellite spectral indexes and spectral fuzzy learning (see Materials and methods) are helpful for disentangling the covers of vegetation and water during satellite SOC modeling.”

Take vegetation for example, the effects of vegetation on spectral reflectance can take on only three situations, namely

60 increasing, unchanging and decreasing. Assuming that we can modify the spectral reflectance according to these situations, the effect of vegetation will be disentangled. In this study, we attempted to modify the spectral reflectance of satellite imagery using our spectral fuzzy learning. This spectral fuzzy learning approach can disentangle the effects from vegetation and water through a series of fuzzy disentangling processes, based on the two indexes including normalized difference vegetation index (NDVI) and normalized difference water index (NDWI). Based on the fuzzified spectra, the SOC model

65 was developed, which produced better performance than **the model using unfuzzyfied spectra**. Then, this model was used to simulate the SOC stocks for the six western European countries (Belgium, France, Ireland, Luxembourg, Netherlands, and UK (excluding Monaco)).

2 Materials and methods

2.1 Input data

70 The inputs for SOC modeling included the measured SOC stocks and Landsat-8 images. LUCAS is the largest soil dataset at the European scale (van Deventer et al., 2019) with sampling density of 2×2 km. A total of five subsamples were composited for each location. Their soil data including SOC, and total contents of sand and clay were measured chemically after air-drying and sieving through a 2-mm sieve. The measured SOC stocks used in this study were derived using the LUCAS 2015 database and the following stock calculation function (Huang et al., 2019):

$$75 \quad SOC \text{ (t C ha}^{-1}\text{)} = SOC\% \times Db \text{ (g cm}^{-3}\text{)} \times depth \text{ (cm)} \quad (1)$$

where SOC (t C ha⁻¹) and SOC% are the measured SOC stock and the LUCAS 2015 SOC content, respectively, depth (cm) is the LUCAS sampling depth (20 cm), and Db (g cm⁻³) represents bulk density data calculated from the bulk density derived function, as shown in Eq. (2) (Hollis et al., 2012). Hollis et al. (2012) suggested that this bulk density derived function was useful for calculating the bulk density of European soils. According to Eq. (2), in addition to the LUCAS SOC content, the LUCAS sand and clay contents were also needed. A total of 754 LUCAS soil samples from the six European countries with sand and clay data were selected.

$$80 \quad Db \text{ (g cm}^{-3}\text{)} = 0.80806 + (0.823844 \times \text{EXP}(-0.27993 \times SOC\%)) \\ + (0.0014065 \times Total \text{ sand}\%) - (0.0010299 \times Clay\%) \quad (2)$$

In recent years, Castaldi (2021) extracted bare soil pixels using threshold values of NDVI of 0–0.35 and normalized burn ratio 2 (NBR2) of <0.125, and indicated that the bare soil pixel extraction technique was suitable for estimating the topsoil properties on cropland. Another study by Silvero et al. (2021) used threshold values of NDVI of 0–0.25 and NBR2 of <0.075 and suggested that using bare soil pixels was helpful for SOC content mapping in tropical regions. Applying this bare soil pixel extraction technique to our preprocessing might be of slight benefit for SOC modeling when considering large scales and complex natural land cover. Therefore, this study used threshold values of NDVI of 0–0.35 and NBR2 of <0.125 to extract bare soil pixels before SOC modeling. Moreover, organic non-clay (i.e., SOC > 20% and clay < 1%) was eliminated in this study. Through bare soil pixel extraction and organic non-clay elimination, 300 LUCAS soil samples were selected for our modeling. For these modeling samples, the average SOC stock of Belgium, France, Ireland, Luxembourg, the Netherlands and UK was 48.697, 103.749, 104.754, 75.927, 73.230, and 92.413 t C ha⁻¹, respectively. Moreover, the SOC

stock of the bare soil pixels was derived using the SOC model, and the SOC stock of the remaining pixels was interpolated using ordinary kriging, as performed by Yigini and Panagos (2016).

95 Landsat-8 images of the six western European countries were used as input satellite data. Landsat-8 is the eighth satellite of the Landsat program conducted by the National Aeronautics and Space Administration and the United States Geological Survey. The spatial resolution and revisit time for Landsat-8 are 30 m and 16 d, respectively. To minimize interference by clouds, images obtained during March–November 2015 with less cloud cover were (<10 %) were extracted. Spectral bands 1–7 used in this study correspond to the Landsat-8 bands of TM2–4 (visible), TM5 (near-infrared), TM6–7 (shortwave
100 infrared), and TM10 (thermal infrared), respectively. Many studies have indicated that first-order differential transformation is always useful for SOC spectroscopic modeling (Chen et al., 2020). Therefore, this study transformed the 7 original bands into differential bands before SOC modeling, and the 7 original bands and their corresponding 21 differential bands were used as spectral inputs.

2.2 Fuzzy disentangling

105 Natural land surfaces such as soil, vegetation, and water can affect the spectral reflectance in satellite imagery, and therefore the effects of vegetation and water cover must be considered in SOC stock modeling. If taking vegetation as an example and assuming that two pixels (pixels 1 and 2) correspond to the same SOC stock and different vegetation cover, then one of the following three situations will occur:

- (i) If vegetation can increase spectral reflectance, then the pixel with the higher vegetation cover will have higher spectral
110 reflectance and the predicted SOC stock of the two pixels will differ.
- (ii) If spectral reflectance is independent of vegetation, then the two pixels will have similar spectral reflectance and similar predicted SOC stock.
- (iii) If vegetation can decrease spectral reflectance, then their corresponding spectral reflectance and predicted SOC stock should also be different.

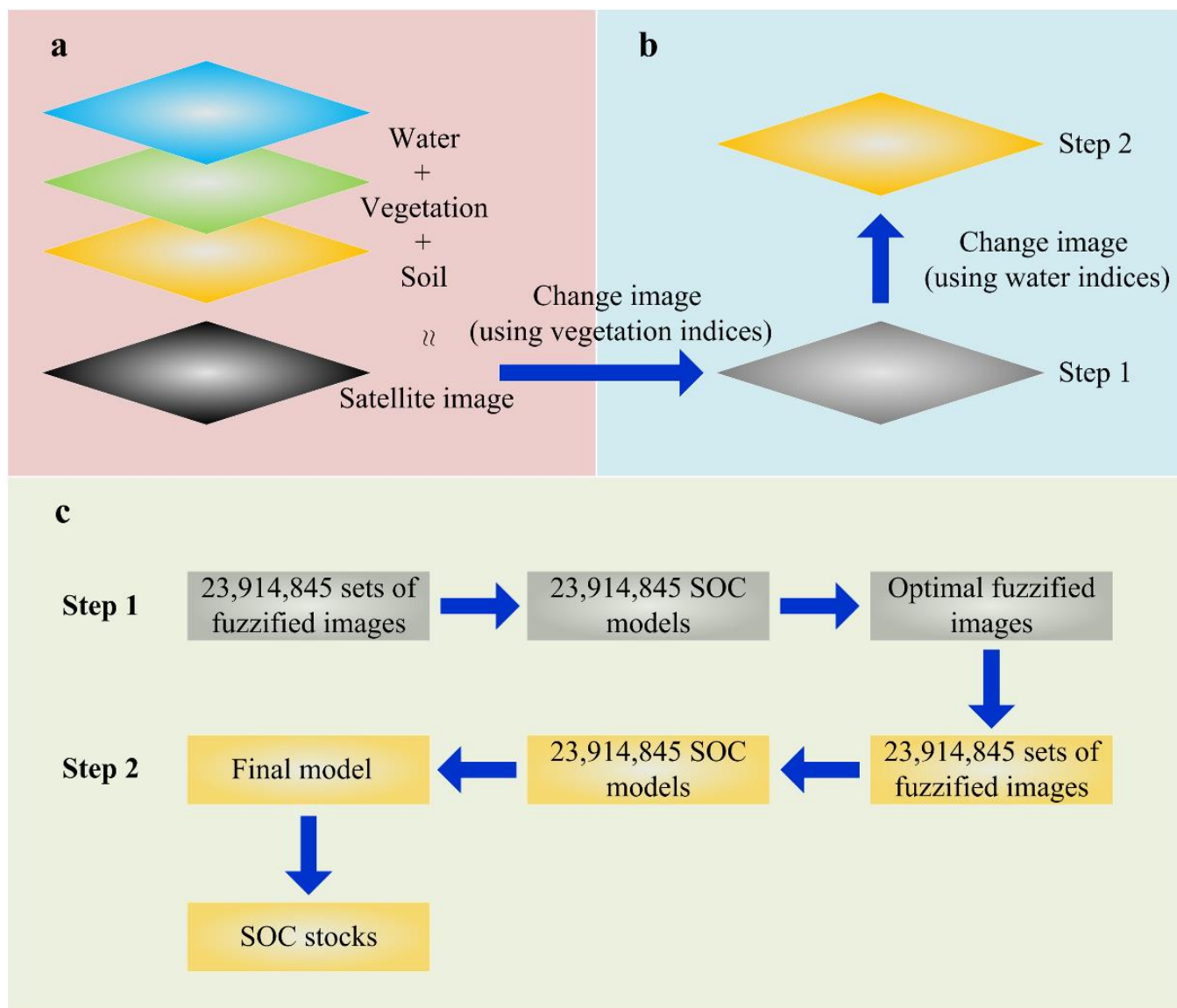
115 If we suppose that we could **modify** the spectral reflectance according to the above three situations, the spectral reflectance of satellite imagery could be optimized for the estimation of SOC stock. According to the three situations mentioned in the introduction, the fuzzy formula (f) is designed as follows (Figs. 1 and 2):

$$\begin{cases} R_f = R \times f \\ f = p^i \\ i = \text{normalized index}^j, j = 3^{-1}, 2^{-1}, 1, 2 \text{ and } 3 \\ p = 0.1 \text{ to } 1 \text{ step } 0.225, \text{ and } 1 \text{ to } 10 \text{ step } 2 \end{cases} \quad (3)$$

where R_f and R are the spectral reflectance after and before fuzzification, respectively, i is the NDVI/NDWI at different
120 power values (j), and p is the fuzzy parameter. In this equation, p and i are used for the changing pattern and the changing degree, respectively, and i is associated with either vegetation or water. In the interpretation of p : when $p < 1$, R_f will be

narrowed—for the situation (i); when $p = 1$, R_f will be unchanged—for the situation (ii); and when $p > 1$, R_f will be enlarged—for the situation (iii). The changing degree depends on the p and i value, tending to increase with increase of the p and i values. Moreover, many previous studies have suggested that NDVI and NDWI can be used for deriving land surface information of vegetation and water, and inspired by the study of Lin and Liu (2022), they used soil moisture indexes to ensure dynamic dewetting. Here, the normalized NDVI and NDWI were used instead of measured vegetation and water values. According to Eq. (3) and Fig. 2, smaller intervals of p and a larger upper boundary of p would produce too many SOC models during fuzzy disentangling. This study set the lower and upper boundaries of p as 0.1 and 10, respectively, with an interval of 0.225 for the 0.1–1 range and an interval of 2 for the 1–10 range. 0.1 to 1 step 0.225 and 1 to 10 step 2 meant 9 different degrees should be used for each band. Moreover, to enlarge the gaps of satellite spectral indexes and therefore further enhance the changing degrees, five different power values j were also used. The above designs meant that it was necessary to develop a total of $9^7 \times 5 = 23,914,845$ models for the vegetation fuzzy disentangling (Fig. 1). Through comparison, the optimal model from the 23,914,845 models of vegetation fuzzy disentangling could be obtained and its corresponding fuzzified spectra were treated as inputs for water fuzzy disentangling. Similarly, 23,914,845 models were also developed during water fuzzy disentangling, and the optimal model from this step was determined as the ultimate SOC model. The p and j values for the optimal models of the vegetation fuzzy disentangling and the water fuzzy disentangling are shown in the lower part of Fig. 2.

Wadoux et al. (2019) indicated that the random forest (RF) method represents a useful tool for estimating the SOC content of soil samples. Many researchers have suggested that the partial least square (PLS) can reduce the problem of multicollinearity, and thus PLS-RF has been proven a powerful combination tool in spectroscopic modeling (Cardelli et al., 2017; Ballabio et al., 2019; Wang et al., 2019). Therefore, this study used PLS-RF as the regression processing tool, with a principal component number of 6, and decision tree and split node numbers of 10 and 2, respectively. All models developed in this study were validated using the four estimation accuracy indexes of RPD, RMSE, R^2 , and MAE (Yigini and Panagos, 2016).



145 **Fig. 1. Diagram of the puzzle and its corresponding solution.** (a) Earth's natural land surfaces are the mixtures of vegetation, water, soil,
 150 etc, this means the other natural land surfaces including vegetation and water must make the satellite remote sensing of soil organic carbon
 (SOC) stocks difficult. (b) The solution works including the vegetation fuzzy disentangling (step 1) and the water fuzzy disentangling
 (step 2). (c) A total of 23,914,845 SOC models were developed during step 1, and based on the fuzzified spectra of the optimal model from
 these models, 23,914,845 sets of SOC models were also developed during step 2. The optimal model from step 2 was determined as the
 ultimate SOC model, and its corresponding simulated SOC stocks were also produced.

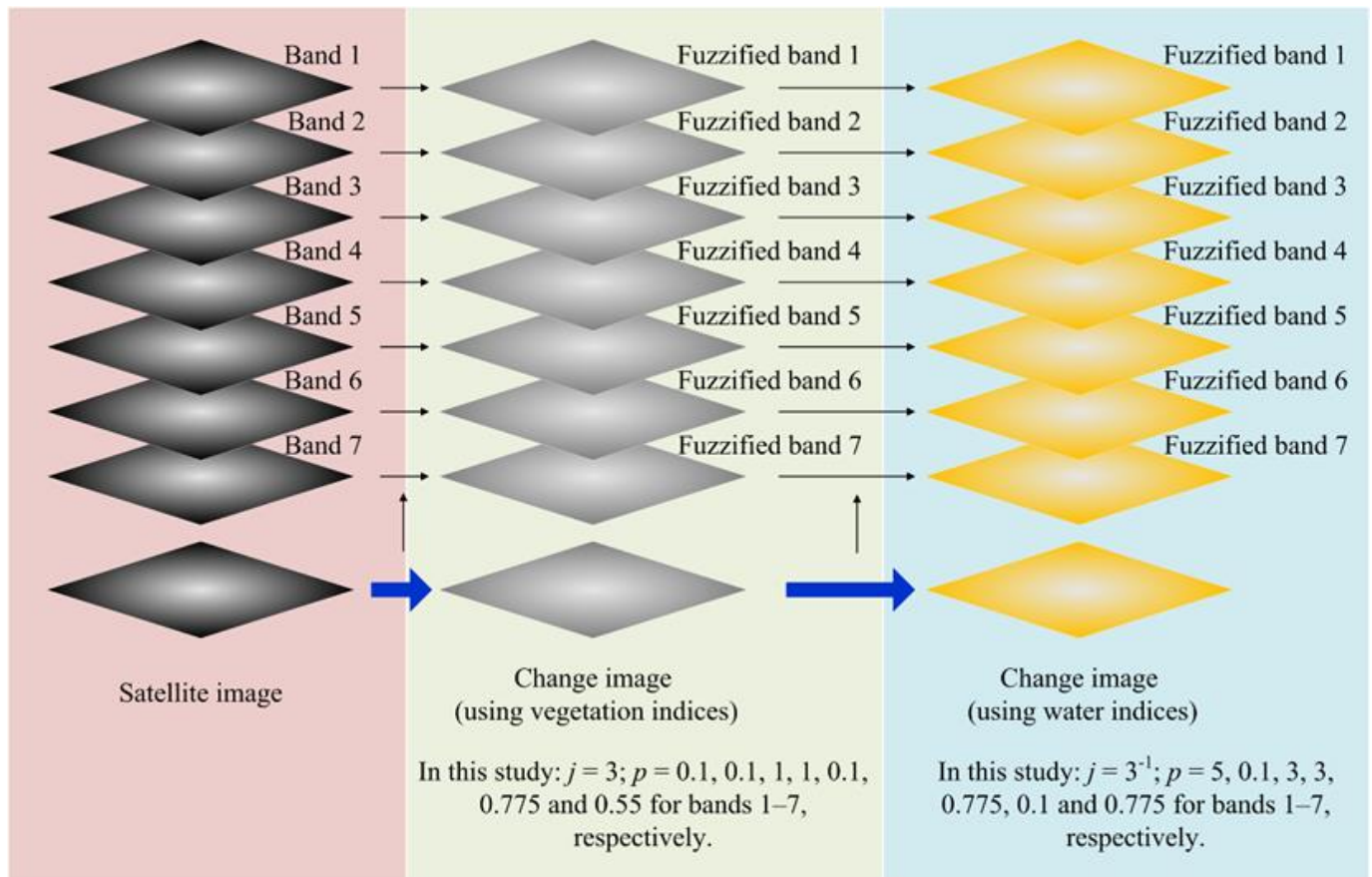


Fig. 2. Explanations of the model number. According to Eq. 1, nine different changing patterns (p) should be occurred on each band. In this study, seven Landsat-8 bands were used for modeling. Moreover, five different power values (j) were used to enlarge the gaps of satellite spectral indices. The above designs explained why a total of $9^7 \times 5 = 23,914,845$ models needed to be developed both for the vegetation fuzzy disentangling and the water fuzzy disentangling. p and j values for the optimal models from the vegetation fuzzy disentangling and water fuzzy disentangling were obtained through model comparisons.

3 Results and discussion

3.1 Simulated SOC stock in western Europe

Vegetation and water can affect the spectral reflectance in satellite imagery. In this study, the fuzzy formula (Eq. (2)) was used to modify the reflectance according to the two vegetation and water indices, and the SOC model based on the fuzzified spectra was developed for the satellite observations of SOC stocks.

The simulated per-unit topsoil (0–20 cm) SOC stock in each of the six western European countries is shown in Fig. 3a. It can be seen that the country with the highest simulated per-unit SOC stock was France ($104.334 \text{ t C ha}^{-1}$), followed by Ireland and the UK with a value of 94.646 and $94.335 \text{ t C ha}^{-1}$, respectively. The country with the lowest simulated per-unit SOC stock was the Netherlands, with a value of $83.673 \text{ t C ha}^{-1}$. It can be seen from Fig. 3b that France had the highest simulated total SOC stock (5.769 Pg) and that Luxembourg had the lowest value (0.024 Pg). The per-unit SOC stock for all six

countries was simulated as $99.742 \text{ t C ha}^{-1}$, corresponding to a total SOC stock of 9.373 Pg. Table 1 presents comparison of European SOC stock as determined by the two studies mentioned in the introduction (the simulated per-unit SOC stocks of Yigini and Panagos (2016) were calculated using the simulated total SOC stock in their published paper). Comparison with

170 Yigini and Panagos (2016) indicates that the per-unit and total stock SOC stocks for all six countries in our work were slightly higher (i.e., $99.742 \text{ t C ha}^{-1}$ versus $94.18 \text{ t C ha}^{-1}$ and 9.373 Pg versus 8.85 Pg). Interestingly, our simulated per-unit results across the six countries had obviously narrower gaps (i.e., $83.673\text{--}104.334 \text{ t C ha}^{-1}$ versus $58.45\text{--}154.90 \text{ t C ha}^{-1}$). The results from Lugato et al. (2014) indicate that their simulated per-unit SOC stocks in the Mediterranean regions and northeastern regions tended to be < 40 and $80\text{--}250 \text{ t C ha}^{-1}$, respectively. Many hotspot regions including Ireland, the

175 Netherlands, the UK, and Finland produced extremely high simulated per-unit SOC stocks with values of $> 250 \text{ t C ha}^{-1}$. Although their study only covered agricultural topsoil (0–30 cm), the comparison results further indicate the narrower simulated per-unit SOC gaps of our work. This is an interesting finding and deserves to be investigated further in the future. The simulated SOC stock distribution and the digital elevation model (DEM) of the six countries are shown in Fig. 3c–d. Over the entire western European region, the simulated SOC stocks tended to increase as the DEM value increased. Taking

180 France and the UK as examples, the central south, southeast, and south of France, the west of England, and most regions of Scotland and Wales had obviously high levels of simulated SOC stock. These high-elevation regions include the Central Plateau, Alps, and Pyrenees of France, the Pennine Hills of England, the Highlands and Southern Uplands of Scotland, and the Cambrian Mountains of Wales. It can also be seen from Fig. 4c that most regions of England and the Netherlands and the northwest of France had relatively low simulated SOC stocks, and Fig. 4d shows that these regions are generally low-lying

185 areas. Bojko et al. (2017) indicated that the distribution of SOC is associated with elevation. Prietzel and Christophel (2014) studied the forest soils in the German Alps and found that soils at high-elevation sites with low air temperature and high precipitation had obviously high SOC stock. The finding of our study of a general trend of high-elevation regions having high per-unit topsoil SOC stock supports the previous findings. Regions with elevation of $> 1000 \text{ m}$ were not sampled in the LUCAS campaigns; instead, the SOC stocks were derived by combining the SOC models with environmental predictors

190 (Yigini and Panagos, 2016) or satellite images (Zhou et al., 2021), as in previous studies. Therefore, this is an area that deserves further investigation in future work.

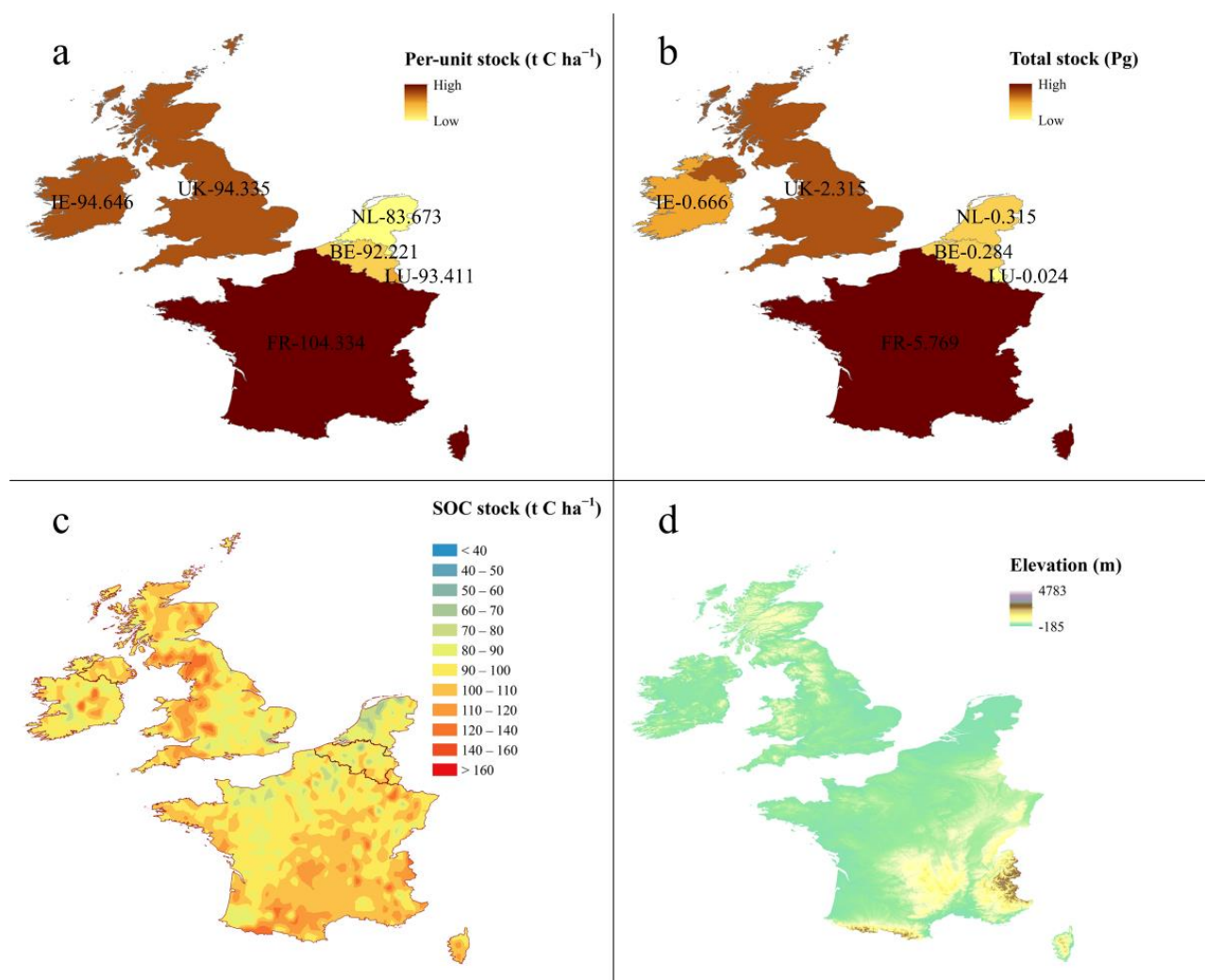


Fig. 3. Simulated soil organic carbon (SOC) stocks. Simulated (a) per-unit and (b) total topsoil (0–20 cm) SOC stocks in the six western European countries (Belgium, France, Ireland, Luxembourg, Netherlands and United Kingdom (UK)). (c) Simulated SOC stock distributions and (d) digital elevation model (DEM) in the six countries.

Table 1 Simulated soil organic carbon (SOC) stocks in this work and some previous studies

This work			Yigini and Panagos (2016)			Lugato et al. (2014)	
Region	Per-unit	Total	Region	Per-unit	Total	Region	Per-unit
Belgium	92.221	0.284	Belgium	58.45	0.18	Mediterranean regions	Tend to < 40
France	104.334	5.769	France	68.90	3.81	North-eastern Europe	80–250
Ireland	94.646	0.666	Ireland	154.90	1.09	Hotspot situations	> 250
Luxembourg	93.411	0.024	Luxembourg	77.84	0.02	Whole Europe	82.4
Netherlands	83.673	0.315	Netherlands	71.72	0.27	Netherlands	100.1
UK	94.335	2.315	UK	141.81	3.48		
Six countries	99.742	9.373	Six countries	94.18	8.85		

The study by Lugato et al. (2014) refers to the agricultural topsoils (0–30 cm), and the hotspot regions with the simulated per-unit SOC stocks $> 250 \text{ t C ha}^{-1}$ including Ireland, Netherlands, UK and Finland; This work and the study by Yigini and Panagos (2016) refer to all the land uses (0–20 cm).

200 **3.2 Model performance**

The first phase of this work was to develop the SOC model. Based on spectral narrowing, unchanging, and enlarging processes with different degrees, a total of 23,914,845 SOC models were developed during the water fuzzy disentangling. A selection of the models with a ratio of performance to deviation (RPD) of > 1.3 is shown in Fig. 4a–d. It can be seen that the RMSE, R^2 , and mean absolute error (MAE) of these models varied markedly (i.e., RMSE: $20.054\text{--}26.527 \text{ t C ha}^{-1}$, R^2 :
205 $0.534\text{--}0.708$, MAE: $17.078\text{--}21.893 \text{ t C ha}^{-1}$). Through comparison, the optimal model with values of RPD of 1.500, RMSE of $20.054 \text{ t C ha}^{-1}$, R^2 of 0.708, and MAE of $17.078 \text{ t C ha}^{-1}$ was obtained and determined as the ultimate SOC model (Fig. 4e). Moreover, the spectral inputs of the water fuzzy disentangling were obtained from the vegetation fuzzy disentangling, i.e., the corresponding fuzzified spectra of the optimal model from the 23,914,845 models obtained during the vegetation fuzzy disentangling. This model had values of RPD of 1.269, RMSE of $22.042 \text{ t C ha}^{-1}$, R^2 of 0.669, and MAE of 17.379 t C
210 ha^{-1} .

The SOC model developed based on unfuzzified spectra is shown in Fig. 4f. This model had values of RPD of 0.822, RMSE of $32.598 \text{ t C ha}^{-1}$, R^2 of 0.282, and MAE of $26.477 \text{ t C ha}^{-1}$, indicating comparable model performance with that reported by the previous studies at the European scale. For example, De Brogniez et al. (2015) and Meersmans et al. (2008) used generalized additive models to study the topsoil SOC content of European countries, and the R^2 value of each of their models
215 was 0.29 and 0.36, respectively. Similarly, Bell and Worrall (2009) mapped SOC content at the regional scale and realized a value of R^2 of 0.48. As already mentioned in the introduction, Yigini and Panagos (2016) investigated SOC stock in Europe and found reported an R^2 value of 0.40. They attributed the low performance of their model to many factors but especially to the effect of complex natural land cover. It can be seen from Fig. 4e that the SOC model with fuzzy disentangling provided better RPD, RMSE, R^2 and MAE values, and that its samples were closer to the 1:1 line than the samples of **the model using**
220 **unfuzzified spectra**. The evident better performance of the model with fuzzy disentangling must be the result of the fuzzy disentangling processes. The comparison results both support the previous assertion that complex natural land cover can substantially increase the difficulty of satellite-based modeling of SOC stock and demonstrate the potential of the fuzzy disentangling approach.

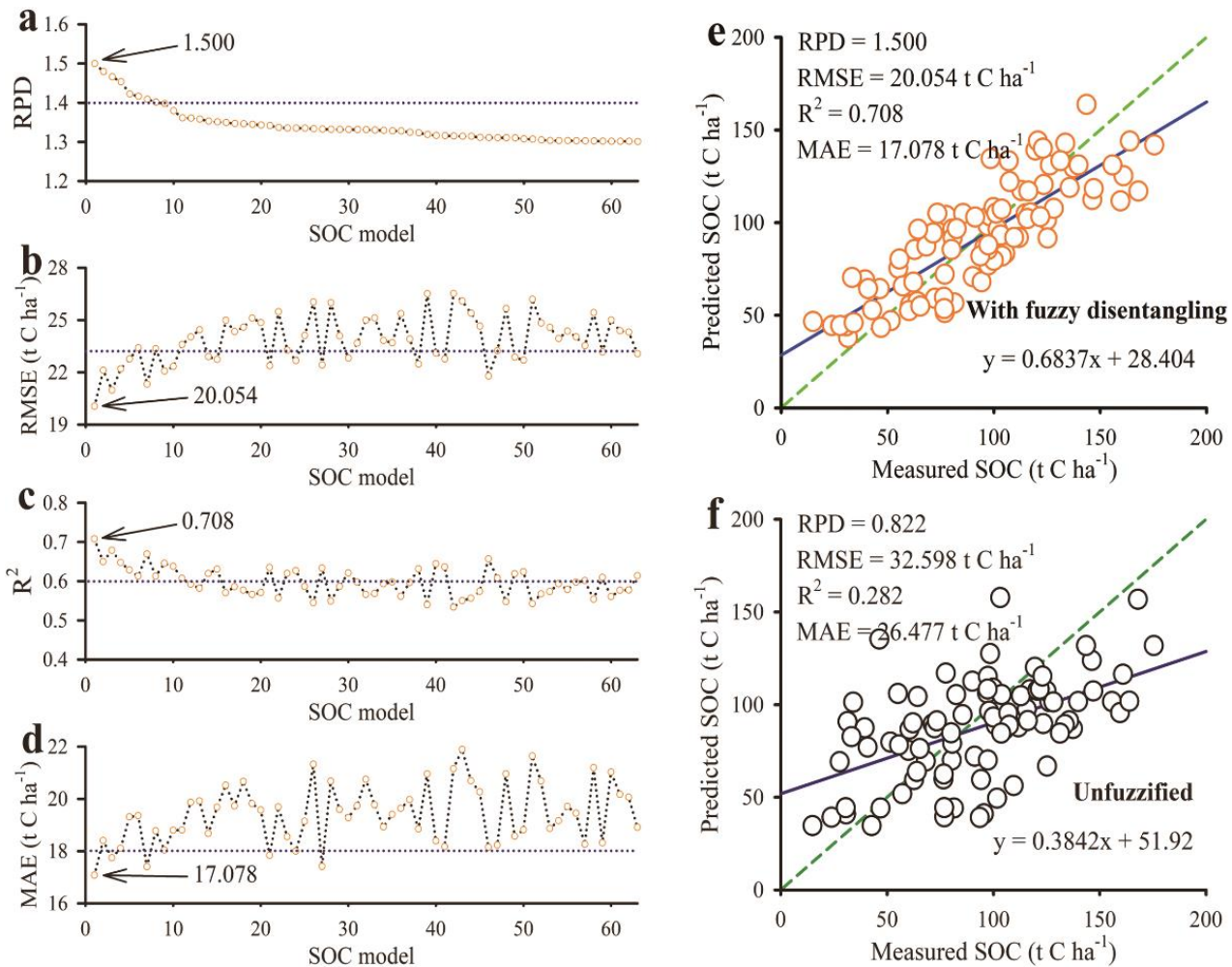


Fig. 4. Model performance. (a) Ratio of performance to deviation (RPD), (b) root mean squared error (RMSE), (c) R-squared (R^2) and (d) mean absolute error (MAE) of the SOC models during fuzzy disentangling. Validation results of (e) the SOC model with fuzzy disentangling versus (f) the model without fuzzy disentangling.

4 Conclusions

This study attempted to disentangle the SOC stock from complex vegetation and water cover, and the two indexes including NDVI and NDWI were used. Based on the processes of vegetation fuzzy disentangling and water fuzzy disentangling, the SOC model gave better results than the model using unfuzzified spectra because of its fuzzy disentangling processes. This model then was used to simulate the SOC stocks in western European topsoils. This study proposed a potential solution for disentangling SOC stock from complex vegetation and water cover, represents a valuable reference for addressing the difficulties of satellite-based observation of continental SOC stock.

The contact author has declared that none of the authors has any competing interests.

Acknowledgments

The research was funded by the National Natural Science Foundation of China (grant no. 42107491). We also thank the European Soil Data Centre, European Space Agency, National Aeronautics and Space Administration, and United States Geological Survey for supporting the research data. We thank James Buxton, MSc, from Liwen Bianji (Edanz) (www.liwenbianji.cn/), for editing the English text of a draft of this manuscript.

References

- Ballabio, C., Lugato, E., Fernandez-Ugalde, O., Orgiazzi, A., Jones, A., Borrelli, P., Montanarella, L., Panagos, P.: Mapping LUCAS topsoil chemical properties at European scale using Gaussian process regression, *Geoderma*, 355, 1–14, doi: 10.1016/j.geoderma.2019.113912, 2019.
- Bell, M.J., Worrall, F.: Estimating a region's soil organic carbon baseline: the undervalued role of land-management, *Geoderma*, 152, 74–84, doi: 10.1016/j.geoderma.2009.05.020, 2009.
- Bojko, O., Kabala, C., Mendyk, L., Markiewicz, M., Pagacz-Kostrzewa, M., Glina, B.: Labile and stabile soil organic carbon fractions in surface horizons of mountain soils-relationships with vegetation and altitude, *J. MT. Sci*, 14, 2391–2405, 2017.
- Cardelli, V., Weindorf, D.C., Chakraborty, S., Li, B., De Feudis, M., Cocco, S., Agnelli, A., Choudhury, A., Ray, D.P., Corti, G.: Non-saturated soil organic horizon characterization via advanced proximal sensors, *Geoderma*, 288, 130–142, doi: 10.1016/j.geoderma.2016.10.036, 2017.
- Castaldi, F.: Sentinel-2 and Landsat-8 Multi-Temporal Series to Estimate Topsoil Properties on Croplands, *Remote Sen*, 13, 1–15, doi: 10.3390/rs13173345, 2021.
- Chen, Y., Li, Y.Q., Wang, X.Y., Wan, J.L., Gong, X.W., Niu, Y.Y., Liu, J.: Estimating soil organic carbon density in Northern China's agro-pastoral ecotone using vis-NIR spectroscopy, *J. Soil. Sediment*, 20, 3698–3711, doi: 10.1007/s11368-020-02668-2, 2020.
- Crowther, T.W., et al.: Quantifying global soil carbon losses in response to warming, *Nature*, 540, 104–108, 2016.
- De Brogniez, D., Ballabio, C., Stevens, A., Jones, R.J.A., Montanarella, L., van Wesemael, B.: A map of the topsoil organic carbon content of Europe generated by a generalized additive model, *Eur. J. Soil Sci.*, 66, 121–134, doi: 10.1111/ejss.12193, 2015.
- Hollis, J.M., Hannam, J., Bellamy, P.H.: Empirically-derived pedotransfer functions for predicting bulk density in European soils, *Eur. J. Soil Sci.*, 63, 96–109, 2012.

- Huang, J.Y., Hartemink, A.E., Zhang, Y.K.: Climate and Land-Use Change Effects on Soil Carbon Stocks over 150 Years in Wisconsin, USA, *Remote Sens.*, 11, 1–20, doi: 10.3390/rs11121504, 2019.
- Hutengs, C., Seidel, M., Oertel, F., Ludwig, B., Vohland, M.: In situ and laboratory soil spectroscopy with portable visible-to-near-infrared and mid-infrared instruments for the assessment of organic carbon in soils, *Geoderma*, 355, 1–11, doi: 10.1016/j.geoderma.2019.113900, 2019.
- Koparan, M.H., Rekabdarkolaei, H.M., Sood, K., Westhoff, S.M., Reese, C.L.: Estimating soil organic carbon levels in cultivated soils from satellite image using parametric and data-driven methods. *Int. J. Remote Sens.* 43, 3429–3449. doi: 10.1080/01431161.2022.2093144, 2022.
- Lin, L.X., Liu, X.X.: Soil-moisture-index spectrum reconstruction improves partial least squares regression of spectral analysis of soil organic carbon, *Precis. Agric.*, 1–13, doi: 10.1007/s11119-022-09905-3, 2022.
- Liu, J., Zhang, D.X., Yang, L., Ma, Y.X., Cui, T., He, X.T., Du, Z.H.: Developing a generalized vis-NIR prediction model of soil moisture content using external parameter orthogonalization to reduce the effect of soil type, *Geoderma*, 419, 1–9, 2022.
- Lugato, E., Panagos, P., Bampa, F., Jones, A., Montanarella, L.: A new baseline of organic carbon stock in European agricultural soils using a modelling approach, *Global Change Biol.*, 20, 313–326, doi: 10.1111/gcb.12292, 2014.
- Meersmans, J., De Ridder, F., Canters, F., De Baets, S., Van Molle, M.: A multiple regression approach to assess the spatial distribution of soil organic carbon (SOC) at the regional scale (Flanders, Belgium), *Geoderma*, 143, 1–13, doi: 10.1016/j.geoderma.2007.08.025, 2008.
- Mueller, C.W., Steffens, M., Buddenbaum, H.: Permafrost soil complexity evaluated by laboratory imaging Vis-NIR spectroscopy. *Eur. J. Soil Sci.*, 72, 114–119, 2021.
- Orgiazzi, A., Ballabio, C., Panagos, P., Jones, A., Fernández-Ugalde, O.: LUCAS Soil, the largest expandable soil dataset for Europe: A review, *Eur. J. Soil Sci.*, 69, 140–153, doi: 10.1111/ejss.12499, 2018.
- Pries, C.E.H., Castanha, C., Porras, R.C., Torn, M.S.: The whole-soil carbon flux in response to warming, *Science*, 355, 1420–1422, 2017.
- Prietz, J., Christophel, D.: Organic carbon stocks in forest soils of the German Alps, *Geoderma*, 221, 28–39, doi: 10.1016/j.geoderma.2014.01.021, 2014.
- Scharlemann, J.P.W., Tanner, E.V.J., Hiederer, R., Kapos, V.: Global soil carbon understanding and managing the largest terrestrial carbon pool, *Carbon Manag.*, 5, 81–91, 2014.
- Silvero, N.E.Q., Dematte, J.A.M., Amorim, M.T.A., dos Santos, N.V., Rizzo, R., Safanelli, J.L., Poppiel, R.R., Mendes, W.D., Bonfatti, B.R.: Soil variability and quantification based on Sentinel-2 and Landsat-8 bare soil images: A comparison, *Remote Sens. Environ.*, 252, 1–19, 2021.
- Thaler, E.A., Larsen, I.J., Yu, Q.: A New Index for Remote Sensing of Soil Organic Carbon Based Solely on Visible Wavelengths, *Soil Sci. Soc. Am. J.*, 83, 1443–1450, doi: 10.2136/sssaj2018.09.0318, 2019.
- Tscherko, D., Kandeler, E., Bardossy, A.: Fuzzy classification of microbial biomass and enzyme activities in grassland soils, *Soil Biol. Biochem.*, 39, 1799–1808, 2007.

- van Deventer, H., Cho, M.A., Mutanga, O.: Multi-season RapidEye imagery improves the classification of wetland and dryland communities in a subtropical coastal region, *ISPRS J. Photogramm*, 157, 171–187, doi: 10.1016/j.isprsjprs.2019.09.007, 2019.
- Wang, J.J., Zareef, M., He, P.H., Sun, H., Chen, Q.S., Li, H.H., Ouyang, Q., Guo, Z.M., Zhang, Z.Z., Xu, D.L.: Evaluation of matcha tea quality index using portable NIR spectroscopy coupled with chemometric algorithms, *J. Sci. Food Agr.*, 99, 5019–5027, doi: info:doi/10.1002/jsfa.9743, 2019.
- Ward, K.J., Chabrillat, S., Brell, M., Castaldi, F., Spengler, D., Foerster, S.: Mapping Soil Organic Carbon for Airborne and Simulated EnMAP Imagery Using the LUCAS Soil Database and a Local PLSR, *Remote Sens.*, 12, 1–20, doi: 10.3390/rs12203451, 2020.
- Wadoux, A.M.J.C., Brus, D.J., Heuvelink, G.B.M., 2019. Sampling design optimization for soil mapping with random forest, *Geoderma*, 355, 1–11.
- Yigini, Y., Panagos, P. Assessment of soil organic carbon stocks under future climate and land cover changes in Europe, *Sci. Total Environ.*, 557, 838–850, doi: 10.1016/j.scitotenv.2016.03.085, 2016.
- Zhou, T., Geng, Y.J., Lausch, A. Prediction of soil organic carbon and the C:N ratio on a national scale using machine learning and satellite data: A comparison between Sentinel-2, Sentinel-3 and Landsat-8 images. *Sci. Total Environ.* 755, 1–16, doi: 10.1016/j.scitotenv.2020.142661, 2021.