

# SCIENTIFIC REPORTS



OPEN

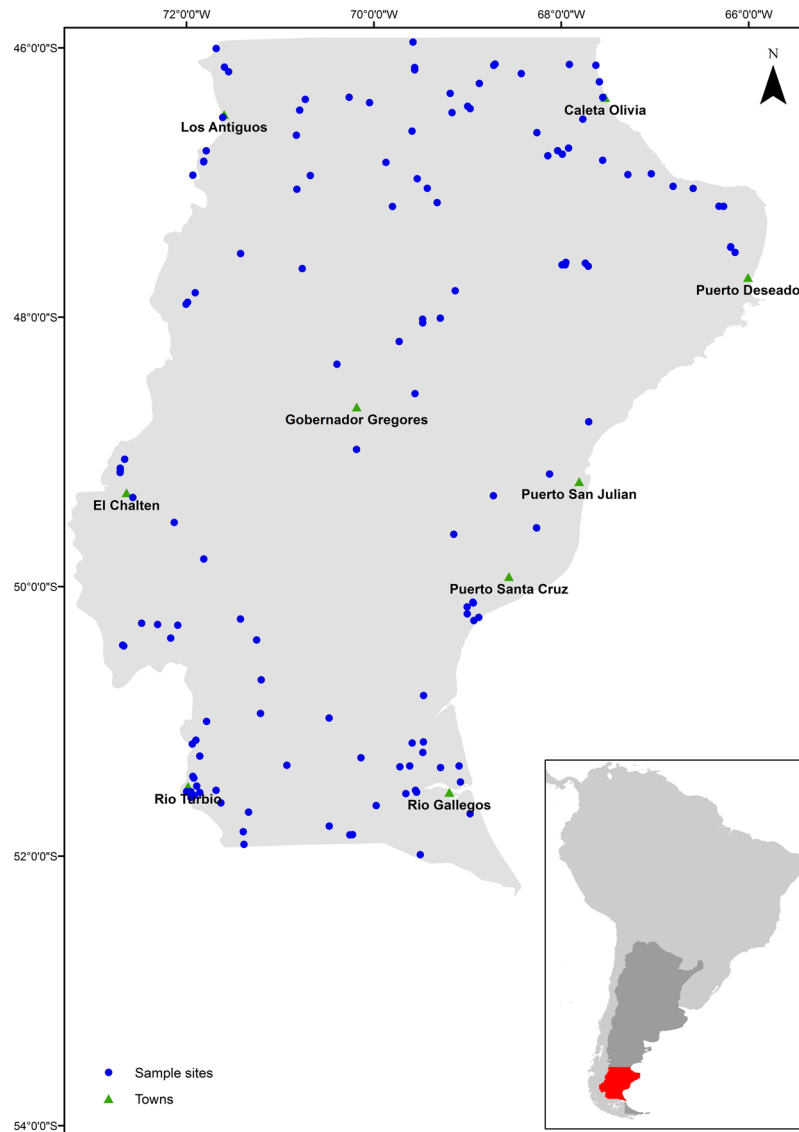
## Soil carbon is a useful surrogate for conservation planning in developing nations

Pablo L. Peri<sup>1,2</sup>, Romina G. Lasagno<sup>1</sup>, Guillermo Martínez Pastur<sup>3</sup>, Rachel Atkinson<sup>4</sup>, Evert Thomas<sup>4</sup> & Brenton Ladd<sup>5,6</sup> 

Defining the optimal placement of areas for biodiversity conservation in developing nations remains a significant challenge. Our best methods for spatially targeting potential locations for biodiversity conservation rely heavily on extensive georeferenced species observation data which is often incomplete or lacking in developing nations. One possible solution is the use of surrogates that enable site assessments of potential biodiversity values which use either indicator taxa or abiotic variables, or both. Among the plethora of abiotic variables, soil carbon has previously been identified as a potentially powerful predictor for threatened biodiversity, but this has not yet been confirmed with direct observational data. Here we assess the potential value of soil carbon for spatial prediction of threatened species using direct measurements as well as a wide range of GIS derived abiotic values as surrogates for threatened plant species in the PEBANPA network of permanent plots in Southern Patagonia. We find that soil carbon significantly improves the performance of a biodiversity surrogate elaborated using abiotic variables to predict the presence of threatened species. Soil carbon could thus help to prioritize sites in conservation planning. Further, the results suggest that soil carbon on its own can be a much better surrogate than other abiotic variables when prioritization of sites for conservation are calibrated on increasingly small sets of observation plots. We call for the inclusion of soil carbon data in the elaboration of surrogates used to optimize conservation investments in the developing world.

While achieving conservation goals requires a landscape level approach, protected areas play an important role and an optimal protected area network should contain representative examples of the ecosystems and biodiversity present in a given region. Unfortunately protected area networks rarely achieve this aim. Ecosystems and landscapes with little to no productive value are more often than not over represented<sup>1–3</sup>. In response to this problem, numerous solutions have been proposed to identify a minimum set of areas that are representative. Examples include the concept of the biodiversity hotspots<sup>4</sup>, ecoregions<sup>5</sup>, important bird areas<sup>6</sup>, the approach used by the alliance for zero extinction sites, as well as site selection based on level of threat (irreplaceability and vulnerability). These have all led to suggestions for a more systematic approach to define priorities for conservation. There has also been discussion around the need to incorporate stakeholders into the decision making process<sup>7</sup>, to determine the relative costs of different strategies<sup>8</sup>, and to consider the fact that neither biodiversity nor threats are static in space or time<sup>9</sup>. However, as many countries with high levels of biodiversity often lack suitable information for identifying important areas for biodiversity conservation there has been a shift toward ecosystem-based planning for the expansion of protected area networks<sup>10,11</sup>. Perhaps the most promising approach towards more objective conservation planning in data poor regions is the development of biodiversity surrogates which have been developed using both abiotic variables<sup>12,13</sup> and indicator taxa. The incorporation of abiotic factors that describe environmental variability in particular have shown promise for surrogate elaboration<sup>13,14</sup>. However despite the significant progress there is a need for further development and refinement of biodiversity surrogates in conservation planning<sup>13</sup>.

<sup>1</sup>Instituto Nacional de Tecnología Agropecuaria (INTA), 9400 Río Gallegos, Río Gallegos, Argentina. <sup>2</sup>Universidad Nacional de la Patagonia Austral (UNPA)-CONICET, 9400 Río Gallegos, Río Gallegos, Argentina. <sup>3</sup>Laboratorio de Recursos Agroforestales, Centro Austral de Investigaciones Científicas (CADIC CONICET), 9410, Ushuaia, Argentina. <sup>4</sup>Biodiversity International, c/o CIP Avenida La Molina, 1895, La Molina, Lima, 12, Peru. <sup>5</sup>School of Biological, Earth and Environmental Sciences, University of New South Wales, Sydney, 2052, Australia. <sup>6</sup>Escuela de Agroforestería, Universidad Científica del Sur, Lima, 33, Peru. Correspondence and requests for materials should be addressed to B.L. (email: [bladd@cientifica.edu.pe](mailto:bladd@cientifica.edu.pe))

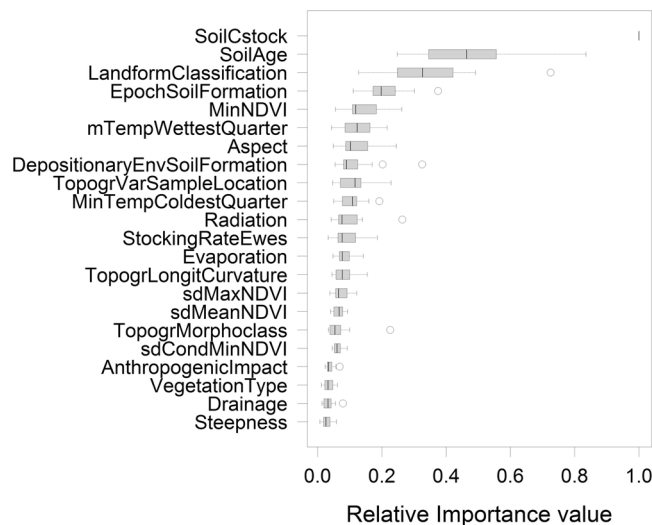


**Figure 1.** Locations of sample sites in the PEBANPA network of permanent plots in Southern Patagonia.

An important variable that has previously been overlooked or missed in the elaboration of biodiversity surrogates is soil carbon. Locations with rich fertile soil in which soil carbon is abundant are an attractive target for conversion to agriculture and this has been the case since the onset of the Neolithic farming revolution<sup>15</sup>. As a result locations with fertile, carbon rich soils are poorly represented in existing reserve networks<sup>2</sup> and the biodiversity that specializes in high-carbon soils may have highly reduced and fragmented ranges<sup>5</sup>. In addition, areas with carbon rich soils are likely subject to more intense anthropogenic threats related to factors such as eutrophication, invasive species, over harvesting, and land use change. We thus hypothesize that species that exist in areas where soil carbon is relatively abundant may face a larger number of more intense extinction threats than species that occur in habitats where soil carbon stocks are relatively low. Here we use observational data from the PEBANPA network of long term biodiversity plots in Southern Patagonia to assess the relationship between soil carbon and threatened plant species. This in turn allows us to test the potential value of soil carbon as a variable that could be used to improve the efficacy of surrogates used for biodiversity conservation planning.

## Methods

To assess for a possible relationship between threatened plant biodiversity and soil carbon we analyzed data from the PEBANPA network of long term biodiversity plots (Biodiversity and Ecological long-term plots in Southern Patagonia)<sup>16</sup> (Fig. 1). The PEBANPA plots were established over 12 years ago and encompass native forest, grassland, shrub-land and wetlands. For each plot comprehensive botanical survey data exists as well as detailed information on climate, topography, land form, soil properties etc.<sup>16</sup>. To investigate possible links between soil organic carbon (SOC) stocks and threatened plant biodiversity, we extracted plant survey data for all 145 sites in the PEBANPA network.

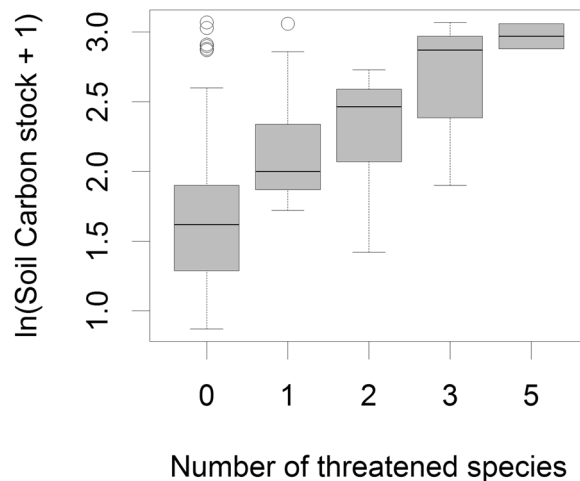


**Figure 2.** Boxplots of relative importance values of the environmental variables based on mean decrease in Gini index across twenty random runs of random forest models. Soil Age = number of years before present that the soil on site formed, Landform Classification = the landform classification of Meybeck *et al.*<sup>23</sup>, EpochSoilFormation = Epoch when the soil on site was formed, MinNDVI = minimum Normalized Difference Vegetation Index, mTempWettestQuarter = mean Temperature of the Wettest Quarter, DepositionaryEnvSoilFormation = Depositionary Environment during Soil Formation, TopogrVarSampleLocation = Topographic Variability in a 500 m diameter spatial window around the Sample Location, mTempcoldestQuarter = mean Temperature of the Coldest Quarter. More detailed description of variable labels of lesser important is given in Peri *et al.*<sup>29</sup>.

All PEBANPA plots are permanently marked and assessed at least once during the flowering period (spring-summer) for accurate plant identification. At each sampling location, plant diversity is measured in a 20 m × 50 m quadrat (1000 m<sup>2</sup>). This plot size enables regional comparisons in diversity-associated factors for the broad vegetation types (e.g. grasslands, shrublands and forests). Species were classified according to origin (native, endemic, exotic), life-form (herb, graminoids, tussock grass, fern, shrub, dwarf shrub, tree), life-span (perennial, annual, biennial), and location of the plant's growth-point (meristem) based on Raunkiaer classification system (geophyte, chamaephyte, phanerophyte, hemicryptophytes, cryptophytes, therophytes).

We used the PlanEAR website (<http://www.lista-planear.org/>) to identify species at risk of extinction in the PEBANPA plots. To assess the importance of different environmental variables in explaining presence or absence of these threatened species in the sample plots we used Random Forests. Given that the majority of plots contained zero or one threatened species, we used presence-absence as a response variable. The explanatory variables we used were soil C stock to 30 cm and the GIS variables described in Peri *et al.*<sup>17</sup> that relate to climate, topography and land use. We reduced collinearity of variables through iterative calculations of variance inflation factors (VIF), retaining only variables with VIFs smaller than 3. To quantify variability in variable importance scores we developed 20 random forest models based on random selections of 70% of the PEBANPA plots. Importance values were based on the mean decrease in the Gini index and were standardized across runs by dividing by the value of the most important variable.

To assess the usefulness of soil carbon for identifying priority sites for the conservation of threatened plant species, we assembled surrogates based on abiotic variables including and excluding soil carbon following the method proposed by Albuquerque and Beier<sup>12</sup>. This approach mimics the planning situation in which species data are available for only a subset  $q\%$ , of the planning plots. It uses species data in each subset to determine the conservation importance of plots based on (i) an optimal conservation scenario and (ii) the use of a surrogate derived from a set of environmental variables characterizing the planning area. Under the optimal conservation scenario, the conservation scores of plots in each subset reflect their importance for finding the smallest set of geographical units that maximizes species representation. Here we used the species richness algorithm originally proposed by Rebelo and Sigfried<sup>18</sup> to this end. Importance scores ranged from 0 (plot not part of the optimal conservation scenario) to 1 (plot with highest richness of unprotected species). Next, we developed random forest models<sup>19</sup> to predict plot conservation importance scores in each subset from the explanatory variables shown in Fig. 2, using the scores from the optimal conservation scenario for model calibration. Model predictions for the entire planning area were then used to assess their efficiency for prioritizing conservation plots, compared to the optimal conservation scenario and the random selection of plots, through calculation of the species accumulation index (SAI)<sup>20</sup> for values of  $q$  ranging from 5 to 60%. The SAI measures surrogacy value by comparing  $S$ , the number of species represented in sites selected using the surrogate (here random forest model predictions), with  $O$ , the largest number of species that can be represented in the same number of sites (optimal conservation scenario), and with  $R$ , the number of species represented in the same number of randomly selected sites. SAI was calculated as  $SAI = (S - R) / (O - S)$  whereby  $S$  represents the number of unique species averaged across 100 random forest models per  $q$ , and  $R$  the number of unique species averaged across 1000 randomly drawn subsets of plots.



**Figure 3.** Boxplots of soil carbon values for plots with different numbers of threatened species in the PEBANPA network. Only plots with zero threatened species had significantly lower soil carbon than plots with higher numbers of threatened species  $p < 0.01$ , Tukey post-hoc test for ANOVA).

SAI values range from  $-\infty$  to 1: negative values indicate a worse than random result, near zero indicates random performance, and positive SAI values (times 100) indicate percent efficiency. For further methodological details please refer to Albuquerque and Beier<sup>12</sup>. All analyses were performed using the *randomForest* package<sup>21</sup> for R<sup>22</sup>, using default settings.

## Results

In the PEBANPA network, 470 plant species from 216 genera and 68 families were recorded in the quadrats surveyed. Across the entire permanent plot network there were 37 species belonging to 20 genera and 13 families that are listed as threatened on the PlanEAR website. There was a strong trend between the number of threatened species found in plots and soil carbon (Fig. 3). The random forest models accurately predicted the presence or absence of threatened plant species in approximately 70% of the plots ( $70.0 \pm 3.8\%$  and  $71.8 \pm 6.0\%$  for training and testing data, respectively). Soil carbon consistently featured as the most important predictor variable across the twenty random forest models and was more than twice as important than the second and third-most important variables, i.e. soil age, and the landform classification of Meybeck *et al.*<sup>23</sup> (Fig. 2 and S1). SAI scores obtained showed that the inclusion of soil carbon significantly increased the efficiency of a surrogate based on abiotic variables to identify sites with the highest conservation priority (Fig. 4; red dots  $7 \pm 1\%$  higher SAI scores than green dots). Remarkably, soil carbon alone was a better surrogate at  $q < 40\%$  than surrogates that included all other environmental variables (Fig. 4; black dots 6–25% higher SAI scores than green dots).

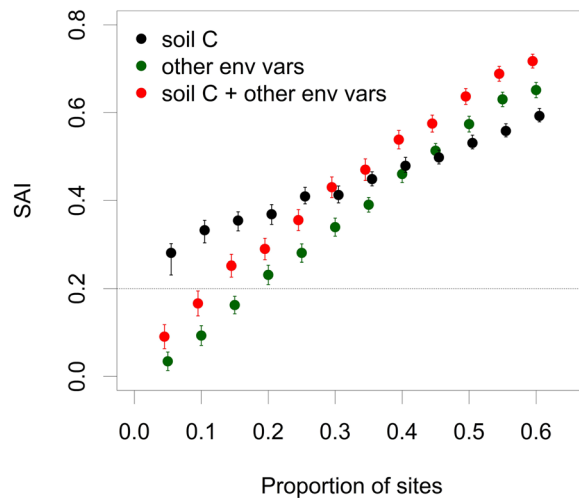
## Discussion

The inclusion of abiotic factors in biodiversity surrogates appears to be a promising approach towards more objective conservation planning in data poor regions<sup>13,14</sup>. Here we show for the first time that including soil carbon significantly improves the performance of a biodiversity surrogate to predict the presence of threatened vascular plant species in Patagonia, adding useful knowledge that will enable further development and refinement of biodiversity surrogates used in conservation planning<sup>13</sup>.

We conclude that over the diverse range of ecosystems found in Southern Patagonia any effort to protect ecosystems with relatively high soil carbon stocks will also benefit threatened plant species. Prioritizing sites with relatively high soil C stocks in the conservation planning process would also help to ensure that the myriad, and poorly documented ecosystem services that flow from soil and soil carbon<sup>24</sup> are also protected<sup>6</sup>.

While the results show a clear correlation between soil carbon and threatened biodiversity, this relationship may be coincidental rather than causal. Low lying parts of the landscape tend to accumulate organic carbon in soil and sediment due to erosional processes, slowed decomposition (anoxia) and reduced burning<sup>25–27</sup>. The relative abundance of water and rich fertile soil with abundant soil carbon make these same areas attractive for agriculture<sup>15</sup>, resulting in extensive land use change. As such, biodiversity that is specialized to these areas may become threatened due to a loss of suitable habitat, over-harvesting of wildlife, eutrophication of water bodies due to high fertilizer use and impacts of invasive species, common in highly modified landscapes, among others. Additionally, because of conflicting land uses these areas tend to be poorly represented in reserve networks<sup>2</sup>. Thus, species that exist in low lying areas rich in soil carbon may face a larger number of more intense extinction threats than species that occur in dry upland habitats where soil carbon stocks are generally lower.

Consistent with previous studies that have used abiotic variables to elaborate biodiversity surrogates<sup>12</sup>, the SAI values for the surrogate based on environmental variables was  $>0.20$  at  $q \geq 20\%$  for the PEBANPA network. This implies that the surrogate is 20% more effective in identifying priority sites for conservation than randomly selected sites when the surrogate model is elaborated using data from 20% of the sites, indicating that the addition of soil carbon to the environmental variables increases the surrogate's power. It is also striking that a surrogate



**Figure 4.** The efficiency of (i) a biodiversity surrogate based on soil carbon alone, (ii) a surrogate elaborated with environmental variables, and (iii) a surrogate elaborated with environmental variables and soil carbon for identifying priority sites for the conservation of threatened species in the PEBANPA network of permanent plots, expressed as the Species Accumulation Index (SAI). The error bars represent the 95% confidence intervals across 100 SAI values, each corresponding to a random forests model developed using the percentage of sites  $q$  indicated on the x-axis. The horizontal line at  $SAI = 0.2$  represents the threshold above which a surrogate is considered good or reliable<sup>12</sup>.

based only on soil carbon had an SAI value  $> 0.25$  for  $q \geq 5\%$ , indicating that soil carbon alone is a powerful predictor for identifying sites of the highest conservation value.

However, it should be noted that while the SAI curve implies that soil carbon initially outcompetes all other variables as a standalone surrogate, it performs less well when more sites are added to the network. This is what we would expect, as it implies that although threatened species may be concentrated on sites with high soil carbon, these sites are relatively homogeneous and a few sites will be able to represent the biodiversity adapted to these areas. To add additional biodiversity to a network design will thus require inclusion of other types of area. Furthermore, relying only on soil carbon could lead to the selection of sites that are adjacent to one another; for example, if all remaining high-soil carbon sites are close together. However, as carbon-rich soils are often associated with water courses, the inclusion of these areas could improve ecological connectivity at a landscape level, increasingly important as the climate changes. Thus spatial optimization methods should always be used, even when surrogates appear to have a high degree of efficacy.

Thanks to recent advances in our ability to map soil carbon using freely available geographic datasets<sup>28,29</sup> obtaining soil carbon estimates is relatively straightforward. While it is obvious that factors other than soil carbon also contribute to species extinction risk, the results presented here demonstrate that soil carbon could be a useful additional variable for the elaboration of biodiversity surrogates that aim to identify networks of sites critical for the conservation of threatened species.

## References

- Margules, C. R. & Pressey, R. L. Systematic conservation planning. *Nature* **405**, 243 (2000).
- Lentini, P., Fischer, J., Gibbons, P., Lindenmayer, D. & Martin, T. Australia's Stock Route Network: 2. Representation of fertile landscapes. *Ecological Management & Restoration* **12**, 148–151, <https://doi.org/10.1111/j.1442-8903.2011.00585.x> (2011).
- Venter, O. *et al.* Bias in protected-area location and its effects on long-term aspirations of biodiversity conventions. *Conservation Biology* **32**, 127–134 (2018).
- Myers, N., Mittermeier, R. A., Mittermeier, C. G., Da Fonseca, G. A. & Kent, J. Biodiversity hotspots for conservation priorities. *Nature* **403**, 853 (2000).
- Olson, D. M. *et al.* Terrestrial Ecoregions of the World: A New Map of Life on Earth: A new global map of terrestrial ecoregions provides an innovative tool for conserving biodiversity. *BioScience* **51**, 933–938 (2001).
- Sheil, D., Ladd, B., Silva, L. C. R., Laffan, S. W. & Heist, M. V. How are soil carbon and tropical biodiversity related? *Environmental Conservation*, 1–11 (2016).
- Theobald, D. & Hobbs, N. T. A framework for evaluating land use planning alternatives: protecting biodiversity on private land. *Conservation Ecology* **6** (2002).
- Manhães, A. P. *et al.* Low-cost strategies for protecting ecosystem services and biodiversity. *Biological Conservation* **217**, 187–194 (2018).
- Pressey, R. L., Cabeza, M., Watts, M. E., Cowling, R. M. & Wilson, K. A. Conservation planning in a changing world. *Trends in Ecology & Evolution* **22**, 583–592 (2007).
- Naidoo, R. *et al.* Global mapping of ecosystem services and conservation priorities. *Proceedings of the National Academy of Sciences* **105**, 9495–9500 (2008).
- Polak, T. *et al.* Efficient expansion of global protected areas requires simultaneous planning for species and ecosystems. *Royal Society open science* **2**, 150107 (2015).
- Albuquerque, F. & Beier, P. Using abiotic variables to predict importance of sites for species representation. *Conservation Biology* **29**, 1390–1400 (2015).

13. Beier, P. & de Albuquerque, F. S. Environmental diversity as a surrogate for species representation. *Conservation Biology* **29**, 1401–1410 (2015).
14. Tukiainen, H., Bailey, J. J., Field, R., Kangas, K. & Hjort, J. Combining geodiversity with climate and topography to account for threatened species richness. *Conservation Biology* **31**, 364–375 (2017).
15. Smith, B. Origins of agriculture in eastern North America. *Science* **246**, 1566–1571 (1989).
16. Peri, P. L. *et al.* Biodiversity and ecological long-term plots in Southern Patagonia to support sustainable land management: The case of PEBANPA network. *Journal for nature conservation* **34**, 51–64 (2016).
17. Peri, P., Ladd, B., Lasagno, R. & Pastur, G. M. The effects of land management (grazing intensity) vs. the effects of topography, soil properties, vegetation type, and climate on soil carbon concentration in Southern Patagonia. *Journal of Arid Environments* **134**, 73–78 (2016).
18. Rebelo, A. & Siegfried, W. Where should nature reserves be located in the Cape Floristic Region, South Africa? Models for the spatial configuration of a reserve network aimed at maximizing the protection of floral diversity. *Conservation Biology* **6**, 243–252 (1992).
19. Breiman, L. Random forests. *Machine learning* **45**, 5–32 (2001).
20. Rodrigues, A. S. & Brooks, T. M. Shortcuts for biodiversity conservation planning: the effectiveness of surrogates. *Annu. Rev. Ecol. Evol. Syst.* **38**, 713–737 (2007).
21. Liaw, A. & Wiener, M. Classification and regression by randomForest. *R news* **2**, 18–22 (2002).
22. Team, R. C. R. A language and environment for statistical computing. (2014).
23. Meybeck, M., Green, P. & Vörösmarty, C. A new typology for mountains and other relief classes: an application to global continental water resources and population distribution. *Mountain Research and Development* **21**, 34–45 (2001).
24. Keesstra, S. D. *et al.* The significance of soils and soil science towards realization of the United Nations Sustainable Development Goals. *Soil* **2**, 111–128 (2016).
25. Stallard, R. F. Terrestrial sedimentation and the carbon cycle: coupling weathering and erosion to carbon burial. *Global Biogeochemical Cycles* **12**, 231–257 (1998).
26. Kane, E. S., Kasischke, E. S., Valentine, D. W., Turetsky, M. R. & McGuire, A. D. Topographic influences on wildfire consumption of soil organic carbon in interior Alaska: Implications for black carbon accumulation. *Journal of Geophysical Research* **112**, G03017, <https://doi.org/10.1029/2007jg000458> (2007).
27. Doetterl, S., Six, J., Van Wesemael, B. & Van Oost, K. Carbon cycling in eroding landscapes: geomorphic controls on soil organic C pool composition and C stabilization. *Global Change Biology* **18**, 2218–2232, <https://doi.org/10.1111/j.1365-2486.2012.02680.x> (2012).
28. Duarte, S. Better estimates of soil carbon from geographical data: a revised global approach. *Mitigation and Adaptation Strategies for Global Change* (2018).
29. Peri, P. L. *et al.* Modelling Soil Carbon Content in South Patagonia and Evaluating Changes According to Climate, Vegetation, Desertification and Grazing. *Sustainability* **10**, 438 (2018).

### Author Contributions

P.L.P., R.G.L. and G.M.P. created the permanent plot network and have worked on the data collection over the last 12 years. B.L. and E.T. analyzed the data. All authors contributed to the redaction of the manuscript.

### Additional Information

**Supplementary information** accompanies this paper at <https://doi.org/10.1038/s41598-019-40741-0>.

**Competing Interests:** The authors declare no competing interests.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019