Assessing the effects of different land-use/land-cover input datasets on modelling and mapping terrestrial ecosystem services - Case study Terceira Island (Azores, Portugal)

Ina M. Sieber[‡], Malte Hinsch[§], Marta Vergílio[|], Artur Gil[¶,#], Benjamin Burkhard[‡,¤]

‡ Leibniz Universität Hannover, Hannover, Germany
§ Gottfried Wilhelm Leibniz Universität, Hannover, Germany
| FRCT - Regional Fund for Science and Technology (Regional Government of the Azores), Ponta Delgada, Portugal
¶ IVAR - Research Institute for Volcanology and Risk Assessment, University of the Azores, Ponta Delgada, Portugal
# Centre for Ecology, Evolution and Environmental Changes (cE3c) & Azorean Biodiversity Group, Faculty of Sciences and Technology, University of the Azores, Ponta Delgada, Portugal
¤ Leibniz Centre for Agricultural Landscape Research ZALF, Müncheberg, Germany

Corresponding author: Ina M. Sieber (sieber@phygeo.uni-hannover.de)

Academic editor: Joachim Maes

Abstract

Modelling ecosystem services (ES) has become a new standard for the quantification and assessment of various ES. Multiple ES model applications are available that spatially estimate ES supply on the basis of land-use/land-cover (LULC) input data. This paper assesses how different input LULC datasets affect the modelling and mapping of ES supply for a case study on Terceira Island, the Azores (Portugal), namely: (1) the EU-wide CORINE LULC, (2) the Azores Region official LULC map (COS.A 2018) and (3) a remote sensing-based LULC and vegetation map of Terceira Island using Sentinel-2 satellite imagery. The InVEST model suite was applied, modelling altogether six ES (Recreation/ Visitation, Pollination, Carbon Storage, Nutrient Delivery Ratio, Sediment Delivery Ratio and Seasonal Water Yield). Model outcomes of the three LULC datasets were compared in terms of similarity, performance and applicability for the user. For some InVEST modules, such as Pollination and Recreation, the differences in the LULC datasets had limited influence on the model results. For InVEST modules, based on more complex calculations and processes, such as Nutrient Delivery Ratio, the output ES maps showed a skewed distribution of ES supply. Yet, model results showed significant differences for differences in all modules and all LULCs. Understanding how differences arise between the LULC input datasets and the respective effect on model results is imperative when computing model-based ES maps. The choice for selecting appropriate LULC data should depend on: 1) the research or policy/decision-making question guiding the modelling study, 2) the ecosystems to be mapped, but also on 3) the spatial resolution of the mapping and 4) data availability at the local level. Communication and transparency on model input data are

©
.

needed, especially if ES maps are used for supporting land use planning and decision-making.

Keywords

InVEST, ecosystem services modelling, Islands, geospatial data

1. Introduction

Modelling ecosystem services (ES) allows us to predict the spatial distribution of different ES that sustain and support human life. Moreover, modelling ES enables us to assess changes in spatio-temporal distribution and the state of ecosystems and how they affect the flows of ES to people. Therefore, modelling ES has become an essential tool for mapping and assessing ES, which is heavily used, for instance, in the context of the European Union's (EU) Initiative Mapping and Assessing Ecosystems and their Services (MAES*[1]) which supports the implementation of the EU's Biodiversity Strategies 2020 and 2030.

The number of readily available model suites has grown in the last decades, with different options to map the actual or potential supply, demand or use of ES. In 2013, more than 17 decision-support tools for ES quantification and valuation models were identified for the assessment of ES (Bagstad et al. 2013) with a growing tendency (Olosutean 2015). Examples range from simple regression models, mechanistic and stochastic models to complex GIS-based ES models (Martínez-Harms and Balvanera 2012). Amongst the most popular open access models, the Integrated Valuation of Ecosystem Services and Tradeoffs model (InVEST) (Sharp et al. 2018), ESTIMAP (Zulian et al. 2013), Resource Investment Optimization System (RIOS) or ARIES (Villa et al. 2014) are listed, allowing to model ES in either biophysical terms (e.g. Mg of carbon sequestered) or economic terms (e.g. net present value of that sequestered carbon) (Natural Capital Assessment (Sharp et al. 2018)).

However, issues related to ES maps and models remain unaddressed, especially those related to uncertainty assessments (Schulp et al. 2014). Numerous studies have compared performances and focused on calibration and validation of models (Cong et al. 2020). Many performance criteria have been established in support of this process and numerous methods to compare maps have been presented (Costanza 1989, Kuhnert et al. 2005, Hagen-Zanker and Martens 2008). Model ensembles have been tested for "goodness of fit" to identify uncertainties (Willcock et al. 2020). For these approaches, map similarity has been in the centre of attention, based on validation by comparing (reference) data from a certain point in time and alignment of model results with reality - assuming that model results are close to reality and, therefore, imply plausibility of the model itself (Hagen-Zanker and Martens 2008).

The InVEST model has been applied in numerous ES mapping applications and first calibration studies highlight the model sensitivities to specific input parameters (Hamel et al. 2015, Redhead et al. 2016, Sharps et al. 2017, Redhead et al. 2018). Whilst model sensitivity to individual parameters has been well documented, an issue that has often been overlooked is how input land-use/land-cover (LULC) data impact the quality of the model outputs. This study tests the performance of three different LULC geodatasets. These included:

1. EU wide CORINE Land Cover (CLC) data,
2. 2018 land cover map of the Azores Archipelago (COS.A 2018), the Azores Region official LULC map and
3. a remote sensing-based LULC and vegetation map of Terceira Island using high-resolution multispectral Sentinel-2 imagery (Fernández-Palacios 2018).

Therefore, the rationale for this ES assessment is two-fold:

- it presents an overview of modelled ES for Terceira Island, based on the InVEST model suite;
- at the same time, it aims to assess the influence of different LULC input datasets for ES modelling and the respective consequences when using these different input LULC to compute ES maps for policy- and decision-making.

2. Material and methods

In this chapter, the case study area, the selected models and selected input datasets, as well as the model parameterisation, are described.

2.1 Case Study Area

Terceira Island is part of the Azores Archipelago, a European Outermost Region (OR) and an Autonomous Region of Portugal with political and administrative autonomy. The Region is an isolated oceanic archipelago located in the Northern Atlantic, approximately between 37° and 40ºN and 24° and 31ºW. Terceira Island is the third largest island of the archipelago and the third oldest (ca. 3.5 million years) with an area of approximately 400 km^2 (maximum length and width of 21 km and 14 km, respectively). It is part of the Central Group, together with the Islands of Faial, Pico, São Jorge and Graciosa (Fig. 1). Most of the Island surface (72%) has altitudes lower than 400 m and only 1% higher than 800 m, with the highest point of the Island located in Serra de Santa Bárbara (1023 m). Hydrography tends to be irregular, with few prevailing small watercourses with torrential regimes, associated with springs located on the northern hills of the massif Guilherme Moniz – Pico Alto that are prone to dry out during drought periods. Existing lagoons are also relatively small (PGRH-Açores 2015).

Terceira Island preserved some pristine areas at high elevation (Gaspar et al. 2011) and few natural areas at lower elevations have remained, especially in Praia da Vitória

Municipality (Borges et al. 2018). However, only 6% of the original native forest exists an area of 23 km^2 (Triantis et al. 2010). The Island's non-urban landscape includes mainly:

- native forest;
- exotic forests plantations (monoculture plantations of *Cryptomeria japonica* and *Eucalyptus globulus* and pure or mixed forest patches dominated by *Pittosporum undulatum* – an exotic invasive species);
- semi-natural pastures (mid- and high-elevation pastures maintaining some indigenous plants), subject to short-term cattle grazing activity and low inputs of fertilisers; and
- intensively managed pastures subject to intensive cattle grazing and high inputs of fertilisers (Rigal et al. 2017).

Additionally to the balanced representation of Azores-specific ecosystems, Terceria Island provides well-documented land use information (see Chapter 2.1). With three official LULC datasets, Terceira surpasses other islands of the Archipelago and, therefore, presents an ideal case study area for conducting an ES modelling study.

The MAES process is comparably advanced in the Azores Archipelago. Many scientific ES assessments of the archipelago have been published (Gil et al. 2017, Madruga et al. 2016, Mendonça 2012, Moreira et al. 2018), as a recent review of scientific studies showed (Sieber et al. 2018). The outputs of such ES modelling studies have been used to support local land management strategies and plans (Cruz et al. 2011, Geneletti et al. 2020). Similar to a global trend, tiered approaches are gaining popularity in the Azores, especially ES models that are easy to operate (e.g. InVEST (NCP 2020) and ESTIMAP (Zulian et al. 2013)). Due to increasing data availability, the number of studies modelling ES on the different islands of the archipelago are increasing, such as Vergílio et al. (2016), Picanço et al. (2017) and Moreira et al. (2018a).

2.2 The InVEST Ecosystem Services model suite

For this study, a modelling approach, based on biophysical models, has been selected (Tier 3; see Grêt-Regamey et al. 2015), applying the InVEST Models. InVEST is a suite of free, open-access, geographic information system (GIS)-based, stand-alone-software models used to map and assess the goods and services provided by nature. InVEST contains several modules and estimates different ES or ES indicators to quantify and map ES, based on spatial, statistical and physical data. InVEST was developed as an ArcGIS Map toolbox in 2008; however, the current version runs as stand-alone software. Besides InVEST's open access availability, its relatively easy implementation is a key advantage. Disadvantages are that, in some cases, the model uses rather simple assumptions, whereas, in other cases, InVEST functions as a black box model. Hence, not all model processing steps can be followed easily by the user. Furthermore, validation of modelling results often proves difficult (Lautenbach et al. 2010, Ochoa and Urbina-Cardona 2017, Polasky et al. 2011, Sharp et al. 2018) due to a lack of reference data. Furthermore, some InVEST modules (e.g. the pollination module) give index values as results which cannot be

measured directly. Therefore, in many cases, only literature-based plausibility checks can be performed instead of a data-based validation.

InVEST is organised in 21 final ES modules and four supporting tools. For this case study, six ES were selected considering the data availability for Terceira Island: recreation, pollination, carbon storage, erosion control, water quality and flow retention (Table 1).

Even though InVEST is stand-alone software, additional GIS software is required to pre-process data, visualise results and perform any further analysis (e.g. data overlays). In this study, data processing was based on ArcGIS ArcMap 10.5. Model post-processing is described in Section 2.5.

2.3 Input LULC datasets

Modelling ES is often implemented on the basis of geospatial data, such as LULC data. The identification of the ecosystems assessed in this study was dependent on the categories of LULC included in each geodataset used (CLC, COS.A 2018 and Sentinel 2-based LULC map, all available as shapefiles).

CLC presents a standardised classification of land use for the entire European Union and some associated countries. The data are publicly available from Copernicus Land Monitoring Services*[2]. CLC data are available from 1990 onwards for the years 2000, 2006, 2012 and 2018. The CLC entails an inventory of altogether 44 LULC types/classes, of which 21 classes are represented on Terceira Island (Table 2). CLC has a minimum mapping unit of 25 ha for areal phenomena and a minimum width of 100 m for linear phenomena (European Environment Agency 2019). Nevertheless, despite its usefulness at European continental and regional level, CLC has been relatively ineffective to address LULC evolution and change issues in European Small Islands/Outermost Regions. This is due to its very broad geographic scale, very large minimum spatial unit (25 ha Minimum Mapping Unit (MMU) and 115 ha average feature size for Terceira) compared to island size and due to its rather inadequate legend, which prevents an efficient characterisation of typical natural and semi-natural LULC units in these territories (Gil et al. 2012).

The official Azores Region LULC Map for 2018 (COS.A 2018) was developed by the Azorean Regional Government (Cruz et al. 2007, Hernâni et al. 2018) and covers the entire Azores Archipelago. The COS.A 2018 is structured hierarchically in three levels, from five major LULC classes on the first level to 29 LULC classes on the third level (Table 2). The LULC nomenclature largely follows the CLC classification, with slight changes. For Terceira, this LULC contains 25 land-cover types (N3), adding aquatic ecosystem types, such as rivers, riparian zones and inland marshes. Initially, the COS.A 2018 was developed using high/very high resolution satellite imagery data classification and an average feature size of 15 ha. The second phase of development consisted of cartographic generalisation and validation tasks in order to address specific spatial planning needs and requirements.

The Sentinel 2-based LULC map of Terceira Island was developed by the Azorean Biodiversity Group*[3] (Fernández-Palacios 2018). It is based on multispectral satellite

imagery with medium/high spatial resolution, derived mostly from available Sentinel-2 data from 2017 and also includes archived Rapideye and Landsat-8 data to fill gaps and increase classification accuracy. It follows its own classification scheme, based on 14 locally-specific ecosystem types and habitats (Table 2), such as *Erica*, *Pittosporum*, *Acacia* , *Pinus* or *Cryptomeria*-dominated forest/scrubland patches. Individual features are most detailed in this dataset with an average feature size of 0.15 ha; however, the focus is clearly on habitat types. Urban LULC is reduced to two classes: built-up areas and bare soil.

All three LULC geodatasets vary in purpose LULC classes and MMU. These differences make them suitable for a comparison of the effects that input datasets can have on ES model outcomes. A visualisation of the three LULC with their individual land use classes is shown in Fig. 2.

2.4 Model parameterisation

In addition to the LULC data, the InVEST modules require additional input data, for example, geospatial or statistical information. The amount of needed input data strongly depends on the complexity of the modules, more specifically, the model processes. These range from very simplistic modules, such as Visitation, to highly complex, multi-parametrical modules, such as Nutrient Delivery or Seasonal Water Yield.

The majority of input data was obtained from a thorough literature review. As locally-specific information on Terceira Island was often scarce, these data gaps were filled with the best available comparable data. For example, most of the data, such as information on Azores-specific soils, were not available online as geodata. In many available datasets, the Azores were not included or were represented as hardly visible, undistinguishable pixels (e.g. FAO World Soil Database (FAO et al. 2009 or Fischer et al. 2008). Therefore, additional soil data were collected from literature and added to the GIS to compile the needed information (Tsui et al. 2013Strohbach and Haase 2012). Table 3 shows a summary of the biophysical and/or statistical input data used for the modelling with InVEST.

2.5 Analysis

In order to assess the relative importance and effects of each input LULC dataset for the outcomes of the model, it was necessary to statistically compare the individual model results. Such an analysis exceeds the scope of existing pairwise map comparison (Hagen-Zanker and Martens 2008, Willcock et al. 2020). InVEST provides all model result maps as either ArcGIS shapefile or grids. Therefore, a fully automated Python script was written, in which data were converted from any InVEST output format to grid format using cell assignment type "cell centre" at an equivalent continuous scale. The resulting raster grids had the same spatial resolution (100 m x 100 m) and were clipped to the Island's extent. For this, Python 2.7 and the ArcPy, NumPy and Pandas libraries were used. The data were transformed from 293 columns and 182 rows (based on the extension of the Island) into a

one-dimensional array with block size of 300 x 200 and flattened to obtain a single column of data with observation values per cell, resulting in arrays of 90,000 observation values and stored in Pandas dataframes for each InVEST module. The 300 x 200 block size was chosen to display the entire grid whilst keeping the dataset as small as possible. This process was repeated for each of the three LULC datasets per InVEST modules, resulting in altogether 18 maps.

Statistical analyses were run for the new data frames, looking at normal distribution, variances and standard deviations for the six InVEST modules with their three LULC datasets. As a normal distribution was not given for any model outputs, a Kruskal-Wallis H test was performed (Kruskal and Wallis 1952), suitable for non-parametric data. This test allows to analyse whether samples originate from different or similar distributions. A significant Kruskal-Wallis test indicates that at least one sample stochastically dominates another sample. However, it neither identifies where such stochastic dominance occurs nor indicates for how many pairs of groups stochastic dominance exists. To minimise the increased risk of type I errors when conducting multiple statistical tests, a Conover-Iman *post hoc* test was applied (both with significance level of $\alpha = 0.05$) (Conover 1999). Results were visualised in boxplots and histograms to show the distribution of values within each raster dataset. Boxplots show the arithmetic means per model result, as well as the standard deviations. Upper whiskers in the boxplots define the third quartile plus 1.5 x the interquartile range (IQR), while lower whiskers indicate the lower quartile minus 1.5 x IQR.

In a last step, the one-dimensional arrays with the results of the statistical evaluation were reshaped and re-assembled to its original format, so that each value was assigned to the former cell in the raster, visualising the difference between the arrays in maps for variance and standard deviation per InVEST module per cell in ArcMap.

3. Results

The six InVEST model modules were applied for Terceira Island. Each module was run with the three input datasets. This process resulted in 18 different ES maps for CLC, COS.A and Sentinel-2-based LULC maps (Figure 3 and Figure 4). Despite the differences amongst the input LULC maps, some models show similar output maps, for example, recreation and pollination. Other model results, for example, carbon storage and nutrient delivery ratio, show differences in the model outputs, indicating differences in the spatial modelled ES supply derived from the differences amongst the input LULC datasets.

Recreation

The InVEST Recreation module quantifies recreation, based on Photo User Days (PUD) uploaded on the online photo-community Flickr*[4] between 2007 and 2017, where 2017 is the last possible year for modelling. The ES “recreation” is supplied predominantly in urban areas, based on the PUD/ha per year, with hotspots in the major cities of Angra do Heroísmo, Praia da Vitória and around Biscoitos (vineyards and coastal site). An exception is located at the touristic spots of "Furnas do Enxofre" (a very popular volcanic cavity). The

overview of the three model results showed very similar results (Figure 3). CLC obtains large areas of no to very weak service supply (0-0.002; 0.003-0.02PUD ha), containing agricultural lands. Continuous urban fabric scored highest for all three input LULC with a total maximum of 22 PUD (0.2 PUD/ha). The city of Angra do Heroísmo, visible as a dark green area in the south of Terceira Island, is one example of this, followed by the urban area around the city of Praia da Vitória in the east. Photo hotspots were found at touristic spots (e.g. Furnas do Enxofre) and natural reserves ("Reserva Natural da Serra de Santa Bárbara e dos Mistérios Negros, Terra Brava e Criação das Lagoas"). Whilst certain areas clearly marked hotspots for recreation and reach between 4 (CLC), 12 (COS.A) and up to 22 PUDs per feature in Sentinel2, averaging the PUDs per hectare results in leaving, for example, moors and heathlands and forested areas highly under-represented (0.004, 0.002 PUD/ha^{-1} per year). For the CLC input LULC with its large polygons and coarse resolution of the pixels, tourist hotspots appeared large on the map. With smaller feature size, the individual maxima became higher; however, tourist hotspots decreased in size.

Pollination

Pollinator abundance was modelled, based on average pollinator abundance, a dimensionless index considering 0 as no to very low abundance to 1 as maximum abundance. Results showed highest potential for pollinator abundance inland of the Island, with the strongest supply in moors and heathlands and coniferous forests edges (CLC) as seen in Figure 3. Both COS.A and Sentinel2-based LULC maps had a lower average pollinator abundance modelled for the same location around the natural reserves (e.g. "Reserva Florestal Natural Parcial do Biscoito da Ferraria" and "Reserva Florestal Parcial da Serra de S. Barbara e dos Misterios Negros") in the Island's centre, with higher potential. The modelled species abundances in the three model results ranged from an average of 0 in port areas and dump sites to average maxima in grasslands (0.316), moors and heathlands (0.329) and territories mainly occupied by agriculture, with significant areas of natural vegetation (0.21). Forested areas, heath and moorlands scored highest in the CLC dataset, but overall pollinator abundance patterns throughout the three LULC remained similar. Agricultural sites, urban areas and port areas scored lowest with all three LULC datasets, with no to very little pollinator abundance (0 – 0.09).

Carbon storage

The carbon storage module calculates a carbon balance for above-ground, below-ground and soil carbon, including carbon stored in dead material. The model computed high capacities for the ES "carbon storage" (Fig. 3) for all three datasets in peat bogs (up to 658 Mg C per ha), followed by forested areas (ranging from 237 – 328 Mg C per ha). Urban areas, airport and port areas scored lowest with all three LULC maps. Nonetheless, model results showed diversity amongst the three LULC model outputs. Under CLC, values were slightly higher than in the other two LULC datasets. These findings are in line with Vergílio et al. (2016), who have calculated similar average Carbon storage values for Pico Island, based on a different LULC map. An exception resulted from the Sentinel-2-based dataset, as a result of the limited differentiation of agricultural classes in this LULC dataset - differentiations of agricultural activities are restricted to “arable land” and “other crops”. Due

to this difference, the class “Other crops” included vineyards, horticulture and pasture areas. Averaging these three classes lead to an underestimation of vineyards and horticulture areas, but caused an overestimation of pastures (modelled values ranging from 81-180 Mg C ha^{-1}), which seemed unrealistically high compared to values of 30 Mg C ha^{-1} found in literature, for example, by Seó et al. (2017). Potentially, this could be solved by adjusting the input parameters according to the LULC surface, taking into account falsely modelled low values for vineyards and horticultural areas. Despite these differences, all three input LULC maps reflected the low carbon storage capacity of urban areas and present similar spatial distribution. Freshwater-related ES nutrient export, sediment export and water quality (flow retention) showed related results.

Nutrient Delivery Ratio (NDR)

The NDR module spatially depicts the outwash of nitrogen (N) from different LULC types. The ES "nutrient export" shows overall low potential for the Island. As shown in Fig. 4, the model outcomes showed a high degree of dissimilarity between the different input LULC maps. Based on CLC, an annual nutrient export, or N outwash, between 45 and 51 kg N per ha was calculated on intensively cultivated agricultural areas and vineyards on the slopes and hillfoots of the Island, which occurred on 18% of the Island’s surface, according to this dataset. N export was modelled with lower loading rates for all other LULC classes. Least annual N export (average of -4.1 - +0.24 kg N y^{-1} per spatial unit) was modelled for forest and peat bogs in CLC. The NDR, based on COS.A, showed less drastic N outwashes. Results showed strong nutrient outwash on 5% of the Island, where steep slopes coincided with humanly-altered land-use types (-45 and -51 kg N y^{-1}), whilst all other areas showed no to very weak nutrient export capacities (-4.1 - +0.24 kg N y^{-1}). A strong contrast was presented in the Sentinel-2-based model results. With only two agricultural LULC classes, the model results for these areas were less distinct. The large surface cover of “other crops”, a mixture of horticulture, vineyards and pastures, showed annual N export values between -30 and -40 kg N y^{-1}. The mosaic of different shades of green in the map in Figure 4 reflects the many fragmented habitat types and their N export and retention capacities.

Sediment Delivery Ratio (SDR)

This module estimates the export of sediment particles (Fig. 4) per ha or watershed per year. It draws upon complex functions of altitude, namely flow accumulation and the Universal Soil Loss Equation (USLE) (Sharp et al. 2018). The modelled average sediment export on Terceira was relatively low for all three LULC datasets (between 0-15 t ha^{-1} y^{-1}) with few locally specific extremes of up to 55 t ha^{-1} y^{-1} in steep agricultural areas towards the centre of the Island. This could be explained by the comparably small scale of agricultural fields, the abundance of support practices, such as stone rows and the high amount of pastures for dairy and meat industry (Gil et al. 2017). Agricultural fields showed an estimated average sediment export value ranging from 4.48 t ha^{-1} y^{-1} (“Land principally occupied by agriculture” CLC, COS.A) to 7.11 t ha^{-1} y^{-1} (“Non-irrigated arable land”, CLC; COS.A), findings that are in line with field experiments by Fontes et al. (2004). The model results computed individual maxima per pixel up to 500 t ha^{-1} y^{-1} (CLC, COS.A, Sentinel-2)

on steep slopes, which is high, considering that losses more than 55 t ha^{-1} y^{-1} are assumed to have a high erosion risk in comparable studies. However, calculated per feature in ArcGIS, this averaged to 35-51 t ha^{-1} y^{-1}. Soil loss in urban and forested areas was calculated to be higher than 1 t ha^{-1} y^{-1} for all three LULC datasets.

Flow retention (Seasonal Water Yield)

The Seasonal Water Yield module calculates the quick flow water recharge and allows to calculate a flow retention index (1-Qn/P). The ES “flow retention" computed very good retention capacities (0.81 – 1) for the majority of the Island, with coastal areas showing slightly lower retention capacities (0.61-0.8) for all three input LULC datasets. The model results (Fig. 4) using CLC showed values ranging between 0.35 to 1.0 [-] for Terceira, with the lowest flow retention values computed for continuous urban fabric (0.35). Highest mean flow retention values were found in natural grasslands (0.95), moors and heathland or urban sports facilities (0.94). Flow retention was lowest using the COS.A dataset with low capacities for streamlines and lower capacities for urban areas: the focus on aquatic ecosystems and riparian zones in the LULC classes might have contributed to this. The highest flow retention was obtained using the Sentinel-2-based LULC dataset: diversity and small size of LULC patterns enhanced retention capacity. Overall, all datasets indicated good flow retention for Terceira. These findings are in accordance with the outputs of the SDR module, showing low average soil exports.

Statistical Results

A statistical analysis of the different datasets reveals trends for each set of model outputs. The arithmetic mean for all modules and LULC datasets is visualised in boxplots (Fig. 5). For some modules, such as Visitation, SDR or Flow Retention, the differences between the LULC seemed small. However, the individual histograms revealed that the magnitude of ES supply differed with input LULC data. Regarding pollination, for instance, potential pollinator abundance was modelled to be highest in CLC with maximum values of up to 0.4, with a standard deviation between 0.05 and 0.18. In both COS.A and Sentinel-2 datasets, this deviation of modelled potential abundance is much smaller with a larger number of outliers. The strongest differences were obtained for the Nutrient Export module. The arithmetic means differ largely, with annual N outwash of -7.8 kg ha^{-1} y^{-1} using the CLC dataset, -43 kg N ha^{-1} y^{-1} for the COS.A map, and -27.9 kg N ha^{-1} y^{-1} year when using the Sentinel-2-based LULC dataset.

Nevertheless, Kruskal-Wallis-H proves significant for all six modules and all three model runs. With p-values < 2.2e-16 < 0.05 = α, the null hypothesis can be rejected in all cases and we conclude that there are significant differences amongst the three model outputs and that there is a significant effect of input LULC datasets on model results. This difference is most distinct for the Nutrient Delivery Ratio module, with differences in nitrogen outwash rates of up to 23 kg ha^{-1} y^{-1} for the agricultural areas deriving from the high nitrogen loading rates from the CLC dataset. In addition, differences between the LULC maps for the Pollination module were distinct on Terceira Island, as results of the size and type of the different forest patches. For Flow retention, based on the SWY

module, the deviation reached up to 0.46, an effect of the aquatic LULC classes present in the COS.A dataset. For the Visitation module, the differences mainly occurred in a few small spots, showing overall the highest consensus of the three input LULC datasets (Fig. 6). In Fig. 6, the distribution of the dissimilarities for the three output maps per InVEST module are shown, highlighting where the differences, based on input LULC, are spatially located on Terceira Island.

4. Discussion

The applications of the six different InVEST model modules on Terceira Island present an overview of the spatial distribution of ES, based on three different input datasets. The statistical analysis showed that the choice of input LULC data largely affects the model outcomes. The maps of the ES model outputs of CLC, COS.A and Sentinel-2 clearly demonstrated significant differences in terms of the modelled distribution of ES supply. Hence, the decision on the input LULC data largely affected the modelled distribution of ES in a case study region. In the following, the reasons for the differences, as well as some practical guidance on what factors to look for when choosing an appropriate LULC dataset, will be given.

For those InVEST modules that model ES, based on specific indicators, such as recreation (Photo User Days), input LULCs only determined the spatial extent of ES supply as seen in the Visitation module. Differences arise, based on the size of LULC features when visualised as maps. Due to the overall low number of PUDs throughout large parts of the Island, differences between the input LULCs are highly local and small, as shown in Fig. 6.

As the InVEST modules require LULC as the major input for their modelling, modules such as Carbon Storage, Visitation or Pollination, showed little differences between the three output maps - potential reason for this could be the similar type of land use. Whether the classification foresees an area to be covered by "broad leaved forest" (CLC and COS.A) or "Pittosporum" (Sentinel-2) or "transitional woodland-shrub" (CLC), "shrubland, bushland, heathlands" (COS.A) or "Shrub peatland" (Sentinel-2), the corresponding data, for example, Carbon Storage values, remain similar and therewith, differences in ES distribution remain rather small, where ES classes closely resemble each other. For example, the outcomes of the pollinator abundance reflect similar trends throughout all three LULC maps. For other modules drawing upon more complex calculations and combinations of input data, the differences in ES maps are substantial. For example, the results of the NDR model show higher deviations between the individual datasets (Fig. 5). However, the extent of this difference becomes only visible on a map (Fig. 6). With an increasing number of model parameters, the differences between the input LULCs became more distinct - here, the choice of input LULC thus becomes more decisive.

The differences in model results can - as was expected - be explained by the use of different input datasets. Differences in input data influence the model outcomes. The different categorisation of LULC classes and the MMU of each dataset are important for spatial accuracy. The InVEST models assume that the supply of ES changes linearly with

the land-use change (Balvanera et al. 2017). With changes in input LULC data, for example, different land uses, the definition of LULC, its spatial resolution, precision and extent, the modelled distribution of ES also changes (e.g. resolution is coarsest in the CLC, followed by COS.A and Sentinel-2-based LULC map). As stated by Gil et al. (2012), an island-scaled LULC cartography with high spatial resolution and with an adequate focus on biodiversity and natural resources protection and management is needed to better address the LULC and derived ES characterisation, evolution and change issues in the Azores Islands. Nevertheless, for ES maps indicating broader trends over larger areas, a coarse resolution may be sufficient. To map a first spatially-explicit distribution of ES, such as recreation (visitation), carbon storage capacity or water quality, the CLC dataset proved sufficient. This is in line with Hamel et al. (2017) who proposed to execute the SDR model on a watershed scale.

Based on our results, the choice of input LULC datasets depends on different factors. The proper selection not only depends on data availability, but also on: 1) the research or policy/decision-making question guiding the modelling study, 2) the ecosystems to be mapped, but also on 3) the spatial resolution of the mapping and 4) data availability at the local level.

Following the MAES guidelines, the purpose of an ES modelling and mapping exercise is often linked to a guiding decision-making or policy question (Maes et al. 2012). Such policy questions often include knowledge requests, policy support questions, questions on resources and responsibilities, application questions and technical and methodological guidance questions (Maes et al. 2018). These questions often bear implications for the scope, scale and resolution of the modelling exercise.

Another factor of importance for the selection of input LULC data is the ecosystem(s) in the focus of the mapping exercise. Here, the level of detail in the classification scheme adopted for each LULC dataset is important for the model output. For example, the carbon storage maps increase in detail with more specific LULC types. Especially for agricultural areas, the difference becomes visible. Both CLC and COS.A datasets contain >=4 subcategories, whereas the Sentinel-2-based LULC map only uses two out of 13 LULC types for agriculture. This is important, when ES maps are to be used for urban or agricultural areas.

The spatial resolution of the modelling exercise is closely linked to the investigated ecosystem. To model ES that highly depend on the structure and composition of the natural mosaic landscape, with its interactions and processes, this study recommends to select habitat-specific LULC, such as the Sentinel-2-based LULC map. Particularly ES, such as pollination, require spatial information on (small) patch sizes of different ecosystem types. It is recommended to use the highest possible spatial resolution for modelling pollinator abundance, even though the results of all three input LULC datasets show a high degree of similarity. A comparison of these results with the work of Picanço et al. (2017), who mapped spatial distribution of insect pollination in Terceira Island, shows similarities even though both studies apply different methodologies (sampling versus modelling approach).

Depending on the size of the study area, the application of different LULC datasets can be useful. For national ES accounting and large scale ES modelling in the European Union, CLC is recommended, as its large average feature size might be sufficient. For regional and local ES assessments, LULC data, such as COS.A and Sentinel-2-based maps, can be used, especially for smaller areas with detailed landscape structures or feature sizes, including small patches and locally-specific ecosystem types. Examples of this are the EU's Outermost Regions and Overseas Countries and Territories, for which CLC is available, but this is unable to capture the dominant local ecosystem types ranging from tropical rainforest to arctic steppe (Petit and Prudent 2008). In this case, the modeller should strive for maximum class types in, for example, aquatic, forest or urban classes for highest output accuracy (Fig. 7).

Lastly, data availability at local level can impact the choice of LULC. With abundant data on a local level, it is possible to model ES for locally-specific ES types, as the example of Sentinel-2 with its endemic ecosystem types shows. Where such local data are limited, COS.A or CLC can be suitable options to conduct the modelling, as reference data on European level can be used as a proxy. However, this comes at the cost of model robustness and introduces uncertainties to the model outputs. Such effects can be severe for small islands as this affects the modelled distribution of ES and hence, the quality of ES maps.

The approach taken in this paper, running ES models with different available LULCs, attempts to minimise uncertainties, based on LULC input datasets. Each model entails uncertainties. As in all computer-based model approaches, model outputs are strongly dependent on the precision, quality and quantity of model input data. Many of the InVEST models require only few data inputs – for example, the carbon storage or visitation modules - which constitute an advantage of InVEST. This makes it broadly applicable across a variety of social-ecological contexts (Balvanera et al. 2017). Often, simple models bear the risk of being incorrect, but may perform well in matching the general patterns of the data, especially when there is a high variance and uncertainty in the data (Costanza 1989). However, if InVEST models rely on such few data inputs, these data need to be as accurate as possible. Nevertheless, on such a local scale as for small (oceanic) islands (as most EU Outermost Regions and Overseas Countries and Territories), many of the required data are not available. As a result, surrogate or proxy data need to be used, which can have a high uncertainty when InVEST models are applied with coarse and simplified secondary data (Balvanera et al. 2017). In short, the more precise and accurate the input LULC data and the better suited the quality of the data for the individual assessment are, the more precise and reliable the modelled maps of ES distribution will be. As this study shows, running ES models with different available LULC datasets can help to add an additional layer to the ES maps - it can predict where the distribution of ES can be modelled with high certainty and where uncertainty in spatial distribution occurs, based on LULC. This knowledge of the degree of uncertainty in ES maps can enhance the quality of modelled ES maps.

Conclusions

This study shows that the choice of input LULC datasets can have a significant impact on the outcomes of ES maps computed with the InVEST model suite on small islands, such as the EU Outermost Region of Terceira on the Azores Archipelago. Comparing three input LULC datasets and six InVEST Modules, significant differences are found between each input LULC. The choice for a particular LULC can either lead to visually enlarging or visually diminishing areas of ES supply on the ES maps. Furthermore, this choice can affect the magnitude of ES supply through the inclusion or omission of certain LULC types. While studies acknowledge sensitivities of the models to input parameters, our work highlights the implications of selecting proper LULC input data - a novel aspect. The use of different input LULC maps in the modelling process can enhance the accuracy of ES maps. Studies and researchers should not only include information on their input parameters, but also on the input LULC dataset with its different classes and feature sizes in order to ensure transparency of the maps for potential users. This is especially relevant if policy and decision-makers or land-use planners are to base their decisions on ES model results.

Acknowledgements

A big thank you to Phil Maurischat for the statistical discussions, to Fernando Santos-Martín for his critical comments and to Angie Faust for language editing. Additionally, Carolina Parelho deserves to be mentioned for her endless efforts to coordinate the MOVE project. Lastly, we would like to thank the reviewers for their valuable feedback. Without them, this paper would not have been possible.

The research leading to these results has received funding from the European Union, under the programme Pilot Project — Mapping and Assessing the State of Ecosystems and their Services in the Outermost Regions and Overseas Countries and Territories: establishing links and pooling resources, MOVE Project (MOVE- Facilitating MAES to support regional policy in Overseas Europe: mobilizing stakeholders and pooling resources, grant agreement Nº 07.027735/2018/776517/SUB/ENV.D2, www.moveproject.eu).

Conflicts of interest

References

- Auxtero E, Madeira M, Sousa E (2004) Variable charge characteristics of selected Andisols from the Azores, Portugal. CATENA 56 (1-3): 111-125. https://doi.org/10.1016/j.catena.2003.10.006

- Bagstad K, Semmens D, Waage S, Winthrop R (2013) A comparative assessment of decision-support tools for ecosystem services quantification and valuation. Ecosystem Services 5: 27-39. https://doi.org/10.1016/j.ecoser.2013.07.004
- Balvanera P, Quijas S, Karp D, Ash N, Bennett E, Boumans R, Brown C, Chan KA, Chaplin-Kramer R, Halpern B, Honey-Rosés J, Kim C, Cramer W, Martínez-Harms MJ, Mooney H, Mwampamba T, Nel J, Polasky S, Reyers B, Roman J, Turner W, Scholes R, Tallis H, Thonicke K, Villa F, Walpole M, Walz A (2017) Ecosystem Services. https://doi.org/10.1007/978-3-319-27288-7_3
- Borges PAV, Gabriel R, Pimentel C, Brito M, Serrano A, Crespo LC, Assing V, Stüben P, Fattorini S, Soares A, Mendonça E, Nogueira E (2018) Biota from the coastal wetlands of Praia da Vitória (Terceira, Azores, Portugal): Part 1 - Arthropods. Biodiversity Data Journal 6 https://doi.org/10.3897/bdj.6.e27194
- Cong W, Sun X, Guo H, Shan R (2020) Comparison of the SWAT and InVEST models to determine hydrological ecosystem service spatial patterns, priorities and trade-offs in a complex basin. Ecological Indicators 112 https://doi.org/10.1016/j.ecolind.2020.106089
- Conover WJ (1999) Practical nonparametric statistics. 3rd ed. John Wiley, New York, Chichester. URL: http://www.loc.gov/catdir/description/wiley032/98008521.html.
- Costanza R (1989) Model goodness of fit: A multiple resolution procedure. Ecological Modelling 47: 199-215. https://doi.org/10.1016/0304-3800(89)90001-x
- Cruz A, Benedicto J, Gil A (2011) Socio-economic benefits of Natura 2000 in Azores Islands–a case study approach on ecosystem services provided by a special protected area. Journal of Coastal Research1955-1959. URL: Cruz, Benedicto et al. 2011 - Socio-economic benefits of Natura 2000.pdf
- Cruz JV, Pereira R, Moreira A (2007) Carta de Ocupação do Solo da Região Autónoma dos Açores. Edited by Secretatia Regional do Ambiente do Ordenamento do Território e dos Recursoso Dídricos.
- European Environment Agency (2019) CORINE Land Cover (CLC) 2018. European Environmental Agency. URL: https://land.copernicus.eu/pan-european/corine-land-cover/clc2018
- FAO, IIASA, ISRIC, ISS-CAS, JRC (2009) Harmonized World Soil Database (version 1.1). online..
- Fernández-Palacios JA (2018) Land-cover/land-use mapping of Terceira Island (Azores, Portugal) by applying the Random Forest supervised classifier to 2017's Sentinel-2 multispectral remote sensing data. Final Report of the Argo Traineeship in the Azorean Biodiversity Group (Angra do Heroísmo, Portugal), pp. 1-64..
- Fischer G, Nachtergaele F, Prieler S, van Velthuizen HT, Verelst L, Wiberg D (2008) Global Agro-ecological Zones Assessment for Agriculture (GAEZ 2008). IIASA, Laxnburg, Austria and FAO Rome, Italy.
- Fontes JC, Pereira LS, Smith RE (2004) Runoff and erosion in volcanic soils of Azores: simulation with OPUS. CATENA 56 (1-3): 199-212. https://doi.org/10.1016/j.catena.2003.10.011
- Gaspar C, Gaston KJ, Borges PA, et al. (2011) Selection of priority areas for arthropod conservation in the Azores archipelago. Journal of Insect Conservation **15**: 671-684. https://doi.org/10.1007/s10841-010-9365-4
- Geneletti D, Adem Esmail B, Cortinovis C, Arany I, Balzan M, van Beukering P, Bicking S, Borges P, Borisova B, Broekx S, Burkhard B, Gil A, Inghe O, Kopperoinen L, Kruse

M, Liekens I, Lowicki D, Mizgajski A, Mulder S, Nedkov S, Ostergard H, Picanço A, Ruskule A, Santos-Martín F, Sieber I, Svensson J, Vačkářů D, Veidemane K (2020) Ecosystem services mapping and assessment for policy- and decision-making: Lessons learned from a comparative analysis of European case studies. One Ecosystem 5 https://doi.org/10.3897/oneeco.5.e53111

- Gil A, Fonseca C, Lobo A, Calado H (2012) Linking GMES Space Component to the development of land policies in Outermost Regions - the Azores (Portugal) case-study. European Journal of Remote Sensing 45 https://doi.org/10.5721/EuJRS20124524.
- Gil A, Fonseca C, Benedicto-Royuela J, et al. (2017) Land Cover Trade-offs in Small Oceanic Islands: A Temporal Analysis of Pico Island, Azores. Land Degradation & Development 29 (2): 349-360. https://doi.org/10.1002/ldr.2770
- Grêt-Regamey A, Weibel B, Kienast F, Rabe SE, Zulian G (2015) A tiered approach for mapping ecosystem services. Ecosystem Services 13: 16-27. https://doi.org/10.1016/j.ecoser.2014.10.008
- Hagen-Zanker A, Martens P (2008) Map Comparison Methods for Comprehensive Assessment of Geosimulation Models. Computational Science and Its Applications – ICCSA194-209. https://doi.org/10.1007/978-3-540-69839-5_15
- Hamel P, Chaplin-Kramer R, Sim S, Mueller C (2015) A new approach to modeling the sediment retention service (InVEST 3.0): Case study of the Cape Fear catchment, North Carolina, USA. Science of The Total Environment166-177. https://doi.org/10.1016/j.scitotenv.2015.04.027
- Hamel P, Falinski K, Sharp R, Auerbach D, Sánchez-Canales M, Dennedy-Frank PJ (2017) Sediment delivery modeling in practice: Comparing the effects of watershed characteristics and data resolution across hydroclimatic regions. The Science of the Total Environment 580: 1381-1388. https://doi.org/10.1016/j.scitotenv.2016.12.103
- Hengl T, Mendes de Jesus J, Heuvelink GM, Ruiperez Gonzalez M, Kilibarda M, Blagotić A, Shangguan W, Wright M, Geng X, Bauer-Marschallinger B, Guevara MA, Vargas R, MacMillan R, Batjes N, Leenaars JB, Ribeiro E, Wheeler I, Mantel S, Kempen B (2017) SoilGrids250m: Global gridded soil information based on machine learning. PlOS ONE 12 (2). https://doi.org/10.1371/journal.pone.0169748
- Hernâni J, Melânia R, Carlos Lima (2018) Carta de Ocupacao do solo da regiao autonima dos acores. COS.A/2018.
- IPCC (2006) Guidelines for National Greenhouse Gas Inventories. Agriculture, Forestry and Other Land Use. With assistance of Eggleston, H. S., Buendia, L., Miwa, K., Ngara, T., & Tanabe, K. IGES, Japan (Vol. 4).. Institute for Global Environmental Strategies, Hayama, Kanagawa, Japan, IGES, Japan. URL: https://www.ipcc-nggip.iges.or.jp/public/2006gl/vol4.html
- Jones A, Montanarella L, Jones R (2005) Soil atlas of Europe. European Communities Off. for Official Publ. of the European Communities, Luxembourg.
- Kruskal W, Wallis WA (1952) Use of Ranks in One-Criterion Variance Analysis. Journal of the American Statistical Association 47 (260): 583-621. https://doi.org/10.1080/01621459.1952.10483441
- Kuhnert M, Voinov A, Seppelt R (2005) Comparing Raster Map Comparison Algorithms for Spatial Modeling and Analysis. Photogrammetric Engineering & Remote Sensing 71 (8): 975-984. https://doi.org/10.14358/pers.71.8.975
- Lautenbach S, Volk M, Gruber B, Dormann C (2010) Quantifying ecosystem service trade-offs. International Congress on Environmental Modelling and Software. 417.

- Madeira M, Pinheiro J, Madruga J, Monteiro F (2007) Soils of volcanic systems in Portugal. In Ó. Arnalds, H. Óskarsson, F. Bartoli, P. Buurman, G. Stoops, E. García-Rodeja (Eds.): Soils of Volcanic Regions in Europe. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 69–81. https://doi.org/10.1007/978-3-540-48711-1_8
- Madruga L, Wallenstein F, Azevedo JM (2016) Regional ecosystem profile – Macaronesian Region. EU Outermost Regions and Overseas Countries and Territories. BEST, Service contract 07.0307.2013/666363/SER/B2, European Commission. 233 p + 10 Appendices,.
- Maes J, Egoh B, Willemen L, Liquete C, Vihervaara P, Schägner JP, Grizzetti B, Drakou E, Notte AL, Zulian G, Bouraoui F, Luisa Paracchini M, Braat L, Bidoglio G (2012) Mapping ecosystem services for policy support and decision making in the European Union. Ecosystem Services 1 (1): 31-39. https://doi.org/10.1016/j.ecoser.2012.06.004
- Maes J, Liekens I, Brown C (2018) Which questions drive the Mapping and Assessment of Ecosystems and their Services under Action 5 of the EU Biodiversity Strategy? One Ecosystem 3 https://doi.org/10.3897/oneeco.3.e25309
- Marcos V (2017) Serviço de Visualização (WMS) da Hidrografia - Região Autónoma dos Açores (RAA) - INSPIRE. Direção Regional do Ambiente/Secretaria Regional da Energia, Ambiente e Turismo/Governo dos Açores. URL: http://wssig3.azores.gov.pt/geoserver/hy-p/wms?service=wms&version=1.3.0&request=GetCapabilities
- Martínez-Harms MJ, Balvanera P (2012) Methods for mapping ecosystem service supply: a review. International Journal of Biodiversity Science, Ecosystem Services & Management 8: 17-25. https://doi.org/10.1080/21513732.2012.663792
- Mendonça FE (2012) Servicos dos Ecossistemas na Ilha Terceira: estudo preliminar com ênfase no sequestro de carbono e na biodiversidade. [PhD Dissertation]. University of the Azores
- Moreira M, Fonseca C, Vergílio M, Calado H, Gil A (2018) Spatial assessment of habitat conservation status in a Macaronesian island based on the InVEST model: a case study of Pico Island (Azores, Portugal). Land Use Policy 78: 637-649. https://doi.org/10.1016/j.landusepol.2018.07.015
- Ochoa V, Urbina-Cardona N (2017) Tools for spatially modeling ecosystem services: Publication trends, conceptual reflections and future challenges. Ecosystem Services 26 (155): 169. https://doi.org/10.1016/j.ecoser.2017.06.011
- Olosutean H (2015) Methods for Modeling Ecosystem Services: A Review. Management of Sustainable Development 7 (1): 5-12. https://doi.org/10.1515/msd-2015-0014
- Ordóñez JA, de Jong BH, García-Oliva F, Aviña FL, Pérez JV, Guerrero G, Martínez R, Masera O (2008) Carbon content in vegetation, litter, and soil under 10 different land-use and land-cover classes in the Central Highlands of Michoacan, Mexico. Forest Ecology and Management 255 (7): 2074-2084. https://doi.org/10.1016/j.foreco.2007.12.024
- Osmond DL, Neas K (2011) Delineating agriculture in the Neuse River Basin. Final report to NCDENR, Division of Water Quality for USEPA 319 program. . - Raleigh, NC..
- Panagos P, Borrelli P, Meusburger K, Alewell C, Lugato E, Montanarella L (2015a) Estimating the soil erosion cover-management factor at the European scale. Land Use Policy 48: 38-52. https://doi.org/10.1016/j.landusepol.2015.05.021
- Panagos P, Borrelli P, Meusburger K, van der Zanden E, Poesen J, Alewell C (2015b) Modelling the effect of support practices (P-factor) on the reduction of soil erosion by

water at European scale. Environmental Science & Policy 51: 23-34. https://doi.org/10.1016/j.envsci.2015.03.012

- Panagos P, Borrelli P, Meusburger K, Yu B, Klik A, Jae Lim K, Yang J, Ni J, Miao C, Chattopadhyay N, Sadeghi SH, Hazbavi Z, Zabihi M, Larionov G, Krasnov S, Gorobets A, Levi Y, Erpul G, Birkel C, Hoyos N, Naipal V, Oliveira P, Bonilla C, Meddi M, Nel W, Al Dashti H, Boni M, Diodato N, van Oost K, Nearing M, Ballabio C (2017) Global rainfall erosivity assessment based on high-temporal resolution rainfall records. Scientific Reports 7 (1): 4175. https://doi.org/10.1038/s41598-017-04282-8
- Paredes P, Fontes JC, Azevedo EB, Pereira LS (2018) Daily reference crop evapotranspiration in the humid environments of Azores islands using reduced data sets: accuracy of FAO-PM temperature and Hargreaves-Samani methods. Theoretical and Applied Climatology 134 (1-2): 595-611. https://doi.org/10.1007/s00704-017-2295-2
- Petit J, Prudent G (2008) Climate change and biodiversity in the European Union overseas entities. IUCN.
- PGRH-Açores (2015) Plano de gestão da Região Hidrográfica dos Açores (rh9) 2016-2021 – Caracterização e Diagnóstico da Situação de Referência - Vol. 3 Terceira. Região Autónoma dos Açores - Secretaria Regional da Agricultura e Ambiente – Direção Regional do Ambiente. URL: https://servicos-sraa.azores.gov.pt/grastore/DRA/PGRHA_20162021/PGRH-A_2016-2021_RT_Parte2_VOL3-TER.pdf
- Picanço A, Gil A, Rigal F, Borges PV (2017) Pollination services mapping and economic valuation from insect communities: a case study in the Azores (Terceira Island). Nature Conservation 18 (1). https://doi.org/10.3897/natureconservation.18.11523
- Polasky S, Nelson E, Pennington D, Johnson KA (2011) The impact of land-use change on ecosystem services, biodiversity and returns to landowners: a case study in the state of Minnesota. Environmental and Resource Economics 48 (2).
- Reckhow KH, Beaulac MN, Simpson JT (1980) Modeling Phosphorus loading and lake response under uncertainty: A manual and compilation of export coefficients. - [In Washington, DC.].
- Redhead J, May L, Oliver T, Hamel P, Sharp R, Bullock J (2018) National scale evaluation of the InVEST nutrient retention model in the United Kingdom. The Science of the Total Environment 610-611: 666-67. https://doi.org/10.1016/j.scitotenv.2017.08.092
- Redhead JW, Stratford C, Sharps K, Jones L, Ziv G, Clarke D, Oliver TH, Bullock JM (2016) Empirical validation of the InVEST water yield ecosystem service model at a national scale. The Science of the Total Environment 569-570: 1418-142. https://doi.org/10.1016/j.scitotenv.2016.06.227
- Rigal F, Cardoso P, Lobo J, Triantis K, Whittaker R, Amorim I, Borges PV (2017) Functional traits of indigenous and exotic ground-dwelling arthropods show contrasting responses to land-use change in an oceanic island, Terceira, Azores. Diversity and Distributions 24 (1): 36-47. https://doi.org/10.1111/ddi.12655
- Schulp CJE, Burkhard B, Maes J, Van Vliet J, Verburg PH (2014) Uncertainties in ecosystem service maps: a comparison on the European scale. PLOS ONE 9 (10): e109643. https://doi.org/10.1371/journal.pone.0109643
- Seó HS, Machado Filho LP, Brugnara D (2017) Rationally Managed Pastures Stock More Carbon than No-Tillage Fields. Frontiers in Environmental Science 5: 711. https://doi.org/10.3389/fenvs.2017.00087

- Sharp R, Tallis HT, Ricketts T, Guerry AD, Wood SA, Chaplin-Kramer R, Nelson E, Ennaanay D, Wolny S, Olwero N, Vigerstol K, Pennington D, Mendoza G, Aukema J, Foster J, Forrest J, Cameron D, Arkema K, Lonsdorf E, Kennedy C, Verutes G, Kim CK, Guannel G, Papenfus M, Toft J, Marsik M, Bernhardt J, Griffin R, Glowinski K, Chaumont N, Perelman A, Lacayo M, Mandle L, Hamel P, Vogl AL, Rogers L, Bierbower W, Denu D, and Douglass J (2018) InVEST 3.5.0.post337+n70cca4fa258b User's Guide. The Natural Capital Project, Stanford University, University of Minnesota, The Nature Conservancy, and World Wildlife Fund.
- Sharps K, Masante D, Thomas A, Jackson B, Redhead J, May L, Prosser H, Cosby B, Emmett B, Jones L (2017) Comparing strengths and weaknesses of three ecosystem services modelling tools in a diverse UK river catchment. Science of The Total Environment 584-585: 118-130. https://doi.org/10.1016/j.scitotenv.2016.12.160
- Sieber IM, Borges P, Burkhard B (2018) Hotspots of biodiversity and ecosystem services: the Outermost Regions and Overseas Countries and Territories of the European Union. One Ecosystem 3 https://doi.org/10.3897/oneeco.3.e24719
- Smith M (1992) CROPWAT : a computer program for irrigation planning and management. Rome: Food and Agriculture Organization of the United Nations.
- Strohbach M, Haase D (2012) Above-ground carbon storage by urban trees in Leipzig, Germany: Analysis of patterns in a European city. Landscape and Urban Planning 104 (1): 95-104. https://doi.org/10.1016/j.landurbplan.2011.10.001
- Triantis K, Borges PV, Ladle R, Hortal J, Cardoso P, Gaspar C, Dinis F, Mendonça E, Silveira LA, Gabriel R, Melo C, Santos AC, Amorim I, Ribeiro S, Serrano AM, Quartau J, Whittaker R (2010) Extinction debt on oceanic islands. Ecography 33: 285-294. https://doi.org/10.1111/j.1600-0587.2010.06203.x
- Tsui C, Tsai C, Chen Z (2013) Soil organic carbon stocks in relation to elevation gradients in volcanic ash soils of Taiwan. Geoderma 209-210 https://doi.org/10.1016/j.geoderma.2013.06.013
- USDA NRCS (2004) National Engineering Handbook: Part 630—Hydrology. Chapter 9: Hydrologic Soil-Cover Complexes. www.nrcs.usda.gov, Washington, DC, USA. URL: https://www.nrcs.usda.gov/wps/portal/nrcs/detailfull/national/water/manage/hydrology/?cid=stelprdb1043063
- Vergílio M, Fjøsne K, Nistora A, Calado H (2016) Carbon stocks and biodiversity conservation on a small island: Pico (the Azores, Portugal). Land Use Policy 58: 196-207. https://doi.org/10.1016/j.landusepol.2016.07.020
- Villa F, Bagstad KJ, Voigt B, Johnson GW, Portela R, Honzak M, Batker D (2014) A methodology for adaptable and robust ecosystem services assessment. PLOS ONE 9 (3). https://doi.org/10.1371/journal.pone.0091001
- Willcock S, Hooftman DP, Blanchard R, Dawson T, Hickler T, Lindeskog M, Martinez-Lopez J, Reyers B, Watts S, Eigenbrod F, Bullock J (2020) Ensembles of ecosystem service models can improve accuracy and indicate uncertainty. Science of The Total Environment 747 https://doi.org/10.1016/j.scitotenv.2020.141006
- Xavier E, Neto A, Reis A, Azevedo JN (2016) Scoping the potencial for outdoor microalgae production in the Azores. ARQUIPÉLAGO - Life and Marine Sciences 33: 57-70.
- Zulian G, Paracchini ML, Maes J, Liquete C (2013) ESTIMAP: Ecosystem services mapping at European scale. Publications Office of the European Union, Luxembourg. https://doi.org/10.2788/64369

Endnotes

*1 http://biodiversity.europa.eu/

*2 https://land.copernicus.eu/pan-european/corine-land-cover

*3 https://gba.uac.pt/

*4 https://www.flickr.com/

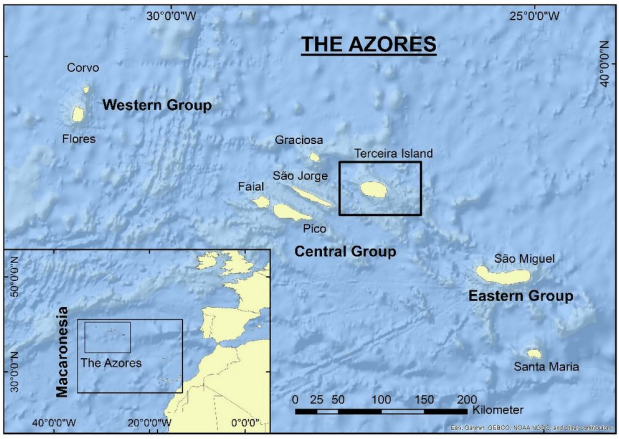


Figure 1.

Location of the Macaronesia biogeographic region, the Azores archipelago and Terceira Island (based on ESRI Basemap).

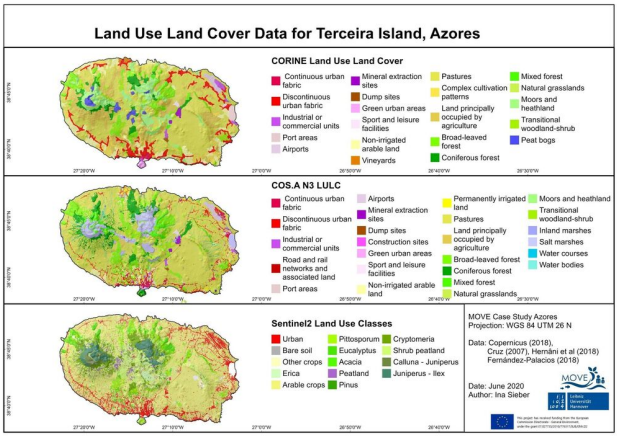


Figure 2.

Overview of the three LULC geodatasets available for Terceira Island.

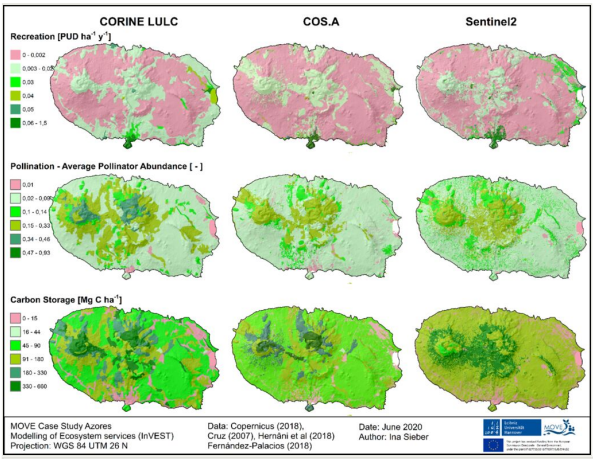


Figure 3.

InVEST Model results for recreation (Visitation), pollinator abundance (Pollinaiton) and carbon storage (Carbon) for the three input LULC datasets CORINE, COS.A and Sentinel-2.

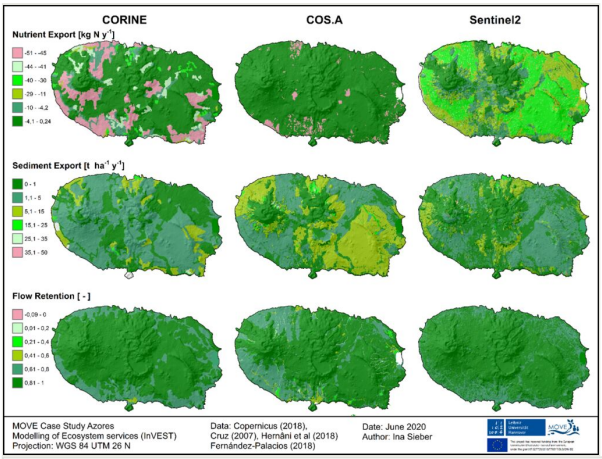


Figure 4.

InVEST Model results for Nutrient export (NDR), Sediment export (SDR) and flow retention (SWY) for the three input LULC datasets CORINE, COS.A and Sentinel-2.

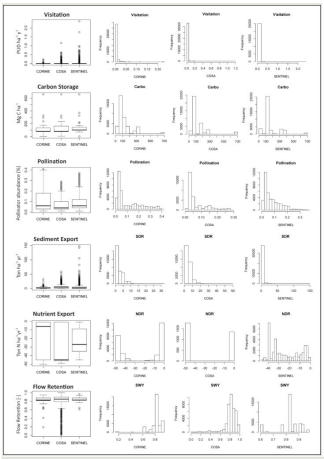


Figure 5.

Statistical analysis of the four analysed InVEST Models with different input LULCs showing boxplots and histograms of the model output data for Terceira Island.

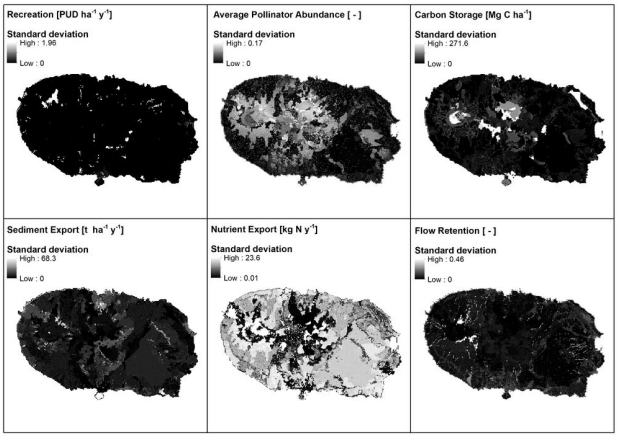


Figure 6.

Spatial distribution of dissimilarities for each InVEST module, based on standard deviations per raster cell for all three input LULC datasets: CORINE, COS.A and Sentinel-2-based map.

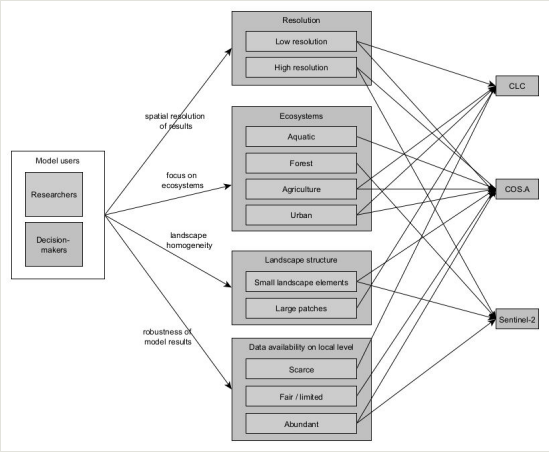


Figure 7.

Decision tree for the selection of input LULC datasets for Terceira Island for the InVEST model suite.

Table 1.

Selected ecosystem services for the modelling on Terceira Island, Azores

Ecosystem service	Ecosystem service type	InVEST module
Recreation	Cultural	Visitation: Recreation and Tourism
Pollination	Regulating and maintenance	Pollinator Abundance: Crop pollination
Carbon Storage	Regulating and maintenance	Carbon Storage and Sequestration
Erosion control	Regulating and maintenance	Sediment Delivery Ratio (SDR)
Water quality	Regulating and maintenance	Nutrient Delivery Ratio (NDR)
Flow retention	Regulating and maintenance	Seasonal Water Yield (SWY)

Table 2.

Overview of the different LULC classes from the three LULC geodatasets, land cover area (ha) and percentage of land cover for Terceira Island.

CORINE Code		Land cover (ha)	%	COS.A Code		Land cover (ha)	%	Sent2 Code		Land cover (ha)	%
111	Continuous urban fabric	101.39	0.25	111	Continuous urban fabric	408.07	1.02	1	Urban	3799.99	9.49
112	Discontinuous urban fabric	2730.38	6.79	112	Discontinuous urban fabric	1318.95	3.29				
121	Industrial or commercial units	314.36	0.78	121	Industrial or commercial units	215.36	0.54				
				122	Road and rail networks and associated land	68.95	0.17				
123	Port areas	289.09	0.72	123	Port areas	60.10	0.15				
124	Airports	319.23	0.79	124	Airports	309.39	0.77				
131	Mineral extraction sites	140.15	0.35	131	Mineral extraction sites	166.76	0.42				
132	Dump sites	40.60	0.10	132	Dump sites	22.80	0.06				
				133	Construction sites	1.02	0.00	2	Bare Soil	493.69	1.23
141	Green urban areas	120.24	0.30	141	Green urban areas	26.53	0.07	1	Urban		
142	Sport and leisure facilities	86.81	0.22	142	Sport and leisure facilities	108.74	0.27				
211	Non-irrigated arable land	3007.67	7.47	211	Non-irrigated arable land	1118.25	2.79	5	Arable crops	13124.52	32.78
				212	Permanently irrigated land	359.05	0.90	5	Arable crops	5547.42	13.86
221	Vineyards	107.26	0.27					3	Other crops	3558.65	8.89
231	Pastures	14862.60	36.93	231	Pastures	23034.29	57.53	3	Other crops	131.49	0.33
242	Complex cultivation patterns	1629.68	4.05								
243	Land principally occupied by agriculture	6516.78	16.19	243	Land principally occupied by agriculture	124.40	0.31				
311	Broad-leaved forest	2348.66	5.84	311	Broad-leaved forest	4123.12	10.30	6	Pittosporum	2607.05	6.51

								7	Eucalyptus	1534.72	3.83
								8	Acacia	337.66	0.84
312	Coniferous forest	1355.01	3.37	312	Coniferous forest	2836.15	7.08	10	Pinus	8.69	0.02
								11	Cryptomeria	2576.86	6.44
313	Mixed forest	115.22	0.29	313	Mixed forest	918.85	2.29	13	Calluna-Juniperus	420.92	1.05
								14	Juniperus-Ilex	1562.04	3.90
CORINE Code		**Land cover (ha)**	**%**	**COS.A Code**		**Land cover (ha)**	**%**	**Sent2 Code**		**Land cover (ha)**	**%**
321	Natural grasslands	954.81	2.37	321	Natural grasslands	45.39	0.11				
322	Moors and heathland	3387.98	8.42	322	Moors and heathland	11.36	0.03	4	Erica	955.43	2.39
324	Transitional woodland-shrub	1223.68	3.04	324	Natural herbaceous vegetation	1117.03	2.79				
				325	Shrubland, bushland, heathlands	1476.33	3.69	12	Shrub peatland	2391.50	5.97
				411	Bare rock	321.15	0.80				
412	Peat bogs	589.42	1.46	421	Flooded zones	1407.63	3.52	9	Peatlands	797.64	1.99
				511	Water courses with vegetation	429.67	1.07				
				512	Water bodies	7.84	0.02				
	Total land-cover	40241,1 ha			Total land-cover	40037.2 ha			Total landcover	40046 ha	

Table 3.

Input data required for the different InVEST modules.

Input data per module	Data type	Data sources
Visitation		
Photo User Days	Image hosting service	Flickr (2017)*[4]
Pollinator abundance: Crop Pollination		
Floral availability and nesting suitability	Literature	Picanço et al. (2017) Zulian et al. (2013)
Carbon		
Carbon storage values	Literature	Vergílio et al. (2016) Strohbach and Haase (2012) Ordóñez et al. (2008) Tsui et al. (2013) IPCC (2006)
Nutrient delivery ratio (NDR)		
Digital Elevation Model (DEM)	Grid (cell size 25 m x 25 m)	Copernicus (2018)
Watershed boundaries	Web Map Service	Marcos (2017)
Nutrient loads	Literature	Osmond and Neas (2011) Reckhow et al. (1980)
Sediment delivery ratio (SDR)		
Digital Elevation Model (DEM)	Grid (cell size 25 m x 25 m)	European Environment Agency (2019)
Soil Data	30 arc-second raster database	JRC (2009) Jones et al. (2005)
Rainfall erosivity	Literature	Panagos et al. (2017)
P Factor	Literature	Panagos et al. (2015b)
C Factor	Literature	Panagos et al. (2015a)
Seasonal water yield (SWY)		
Soil Data	30 arc-second raster database	JRC (2009) Jones et al. (2005) Auxtero et al. (2004)
Annual precipitation	Literature	Xavier et al. (2016) World Weather Online (2018)
Soil hydraulic parameters	Literature	Madeira et al. (2007) Fontes et al. (2004)
Runoff curve numbers	Grey literature	USDA NRCS (2004)
Crop evapotranspiration	CROPWAT Model, Literature	Smith (1992), Paredes et al. (2018)
Soil types	Geotiff	Hengl et al. (2017)
Watershed boundaries	Web Map Service	Marcos (2017)