The DIARS toolbox: a spatially explicit approach to monitor alien plant invasions through remote sensing

Carol X. Garzon-Lopez‡,§, Tarek Hattab‖,§, Sandra Skowronek‖, Raf Aerts#, Michael Ewald‖, Hannes Feilhauer‖, Olivier Honnay#, Guillaume Decocq‖, Ruben Van De Kerchove‖, Ben Somers‡, Sebastian Schmidtlein‖, Duccio Rocchini†,‡, Jonathan Lenoir#‡,§,|,§ ¶, #, ¤ «, # § » ˄, ˅,¦,ˀ § © Garzon-Lopez CX et al. This is an open access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Corresponding author: Carol X. Garzon-Lopez (c.x.garzon@gmail.com)
Received: 27 Mar 2018 | Published: 30 Mar 2018

Abstract

The synergies between remote sensing technologies and ecological research have opened new avenues for the study of alien plant invasions worldwide. Such scientific advances have greatly improved our capacity to issue warnings, develop early-response systems and assess the impacts of alien plant invasions on biodiversity and ecosystem functioning. Hitherto, practical applications of remote sensing approaches to support nature conservation actions are lagging far behind scientific advances. Yet, for some of these
technologies, knowledge transfer is difficult due to the complexity of the different data handling procedures and the huge amounts of data it involves per spatial unit.

In this context, the next logical step is to develop clear guidelines for the application of remote sensing data to monitor and assess the impacts of alien plant invasions, that enable scientists, landscape managers and policy makers to fully exploit the tools which are currently available. It is desirable to have such guidelines accompanied by freely available remote sensing data and generated in a free and open source environment that increases the availability and affordability of these new technologies.

Here we present a toolbox that provides an easy-to-use, flexible, transparent and open source set of tools to sample, map, model and assess the impact of alien plant invasions using two high-resolution remote sensing products (hyperspectral and LiDAR images). This online toolbox includes a real case dataset designed to facilitate testing and training in any computer system and processing capacity.

Keywords

Biological invasions, ecosystem impact, hyperspectral images, LiDAR, species detection and mapping, species distribution models

Introduction

Biological invasions by non-native, exotic or “alien” species (hereafter IAS; invasive alien species: http://ec.europa.eu/environment/nature/invasivealien/index_en.htm), often related to other threats such as land use intensification and environmental change (Turbelin et al. 2016), are considered one of the biggest threats for biodiversity (Vitousek et al. 1997). Unfortunately, the number of alien species colonizing and impacting ecosystems is likely to increase (Levine et al. 2003, Seebens et al. 2017) due to the increase in global trade with dominant pathways for IAS being horticulture and nursery trade (Chapman et al. 2017). In particular, invasive alien plants (hereafter IAPs) may have significant impacts on ecosystem functioning. Several studies have analyzed the impact of IAPs on ecosystem structure and dynamics, including impacts on nutrient cycling, hydrology and fire regimes (Ehrenfeld 2010, Weidenhamer and Callaway 2010), thereby highlighting the magnitude of the problem and calling for urgent actions to control and manage the invaded as well as the vulnerable ecosystems.

In this context, the development of thorough management actions requires accurate assessments of IAS occurrences at fine spatial resolutions and across large spatial extents. Field surveys are crucial for this task but due to the exhaustive fieldwork required to monitor changes over time and across large spatial extents, this type of survey is sometimes not feasible. Besides, field surveys might be subject to biases in detection, especially in cases of early colonization and in relation to the level of expertise of the observer (Fitzpatrick et al. 2009).
Noteworthy, field observations of IAS’ distribution are often collected as presence-only, despite the fact that accurate data on species absences is also crucial for monitoring and for the development of models capturing the species distributions, also known as species distribution models (SDMs) (Lobo et al. 2010). This fact is of paramount importance in the case of IAS which usually are not in equilibrium with their environment within the invaded range, that is, species would have already occupied all suitable habitats and is absent from all unsuitable habitats: a key assumption of traditional SDMs. In the case of IAS, a careful distinction between observed absences due to unsuitable environmental conditions, also known as environmental absences, and those due to dispersal limitations to locations that are environmentally suitable, known as contingent or dispersal-limited absences (Hattab et al. 2017), is a key prerequisite to the SDM development. In this context, the availability of spatially explicit data of high resolution and over large extents, in combination with ground surveys, could provide the ingredients for the development of a framework for detection, monitoring and assessment of IAPs that generate products that are relevant for management at multiple spatial and temporal scales (He et al. 2011, Hattab et al. 2017)

Remote sensing provides continuous spatially explicit data at several temporal and spatial resolutions ranging from hundreds of meters at high temporal resolution to few centimeters at lower temporal resolutions and across a steadily increasing spatiotemporal extent. A growing number of studies have demonstrated the applicability of remote sensing technology, and specifically of hyperspectral and Light Detection And Ranging (cf. LiDAR) sensors, to detect and monitor IAPs (Huang et al. 2013). This flourishing literature ranges from the differentiation of native vs. alien species (Somers and Asner 2013) to the detection of single species including trees, sub canopy trees (Barbosa et al. 2016) and even tiny moss species (Skowronek et al. 2016). The development of new platforms and instruments are opening new research avenues that integrate remote sensing as a standard tool to monitor IAPs. And along with these new technologies comes the need for transparent and affordable software tools as well as trained experts to process these large datasets promptly and accurately.

Hyperspectral images are often described as a data cube with a spatial X- and Y-dimension, and a third dimension containing information on the earth surface reflectance across the electromagnetic spectrum. This information is provided for hundreds of spectral bands in the wavelength range spanning from the visible to the mid-infrared part of the spectrum. Such data set can be linked to biochemical and biophysical vegetation properties via empirical models which can be used to:

1. estimate key traits such as specific leaf area or leaf chlorophyll content (Ghiyamat and Shafri 2010, Homolová et al. 2013);
2. assess plant diversity (Nagendra 2001);
3. classify plant or plant functional types, e.g. to separate native and alien species (Asner et al. 2008); or
4. identify single species (Bradley 2013, Skowronek et al. 2016).
LiDAR sensors use the light emitted from a laser pulse to measure the travel time of the pulse from the source to the target and back. Using this procedure, the instrument provides information on the surface and the vegetation 3D structure at high spatial resolution (Lim et al. 2003). The data collected is distributed in point clouds with varying density depending on the sensor specifications, flight height and velocity. The higher the point density the higher the accuracy in the products derived from it. For example, digital elevation models (DEMs) can be derived from low density point clouds (1 point per m²) while vegetation studies require higher point density (from 8 point per m² onwards) (Wu et al. 2016). The LiDAR-derived 3D structure of the terrain and the vegetation can be used as surrogates for physiographic and biophysical properties (e.g. Lenoir et al. 2016), respectively, both being potential determinants of the distribution and spread of IAPs. For this reason, LiDAR images are very complementary to hyperspectral images when the aim is to understand the determinants of IAPs’ distribution within the invaded range and subsequently model the potential spread.

The aim of this paper is to provide an open source toolbox to face the challenge of detecting, monitoring and assessing the impact of IAPs on ecosystem functioning through remote sensing, this work is based on the results and knowledge gained from an interdisciplinary BiodivERsA project (DIARS, http://diars.vgt.vito.be/).

The toolbox (http://diarsproject.github.io/DIARS/HomeDIARS.html) features clear guidelines to process and analyze ground and remote sensing data to map, model and assess the impact on ecosystem functioning of IAPs. A dataset specially designed to allow computation even at small processing power as well as the “iSDM” R package (https://cran.r-project.org/web/packages/iSDM/) to help inform the sampling of IASs as well the mapping and modelling of IASs are also provided together with the toolbox.

**Overview of the toolbox**

The toolbox is designed as an easy-to-use, free and open-source solution for the detection, monitoring and impact assessment of IAPs through remote sensing (Fig. 1). It consists of a series of tutorials that include exemplary datasets of ground surveys, LiDAR point cloud data, hyperspectral images and several R functions to:

- implement an optimal sampling design for ground surveys;
- map and model the realized and potential distributions of IAPs within the invaded range, respectively; and
- assess the impact of IAPs on some ecosystem functions.

The toolbox tutorials are organized into two main sections:

- data preparation and
- applications (Fig. 1).
The first section contains three tutorials on hyperspectral data, LiDAR data processing and the implementation of the method and R functions to generate an optimized systemic sampling design. The second section presents the three main applications of the toolbox: mapping, modeling and impact assessment.

### Data processing

There are three main types of data used in the toolbox: hyperspectral data; LiDAR data; and field data. Each type of data requires a specific processing that is explained in a dedicated tutorial. The first tutorial, hyperspectral data processing guidelines, includes a procedure to filter specific bands with values that might add noise to the analysis (e.g., water vapor) and the steps recommended to have the data ready to use in R (R Core Team 2015).

The LiDAR guidelines include instructions on how to import the data (in LAS format) into R and extract a high-resolution digital terrain model (DTM), as well as various canopy height metrics and statistics. This part of the tutorial is based on an interface between GRASS GIS 7.0 (GRASS Development Team 2015) and the R environment, which offers the possibility to ecologists (who are more familiar with R) to use GRASS GIS commands from the R command line via rgrass7 package (Bivand 2016).

Finally, in the third part of this section, highlights an approach that optimizes the sampling of observed presence-absence data of IAS in the field and the handling of absence data for subsequent analyses. The tools are presented as part of the “ISDM” R package (Hattab et al. 2017), that provides functions to facilitate the setup of ground surveys that optimally capture the environmental variation available within the studied area through a systematic
sampling design within the environmental space and, to assess the probability that observed field absences are contingent and thus should only be used for mapping the realized distribution but not for modeling the potential distribution (Fig. 2, Hattab et al. 2017).

![iSDM Package](image)

**Figure 2.** doi

Overlook of the “iSDM” R package functions presented in the toolbox.

### Applications

The applications part of the toolbox is based on three main pillars: mapping; distribution modeling; and impact assessment. Early detection and monitoring of IAPs is key to track and minimize its negative impacts on natural ecosystems, while ground surveys are of crucial importance to ensure early detection and monitoring of IAS. Yet, the logistical barriers to reach remote areas and organize periodic surveys limit its success. The potential of hyperspectral images has already proven useful to face the challenge of detecting even relatively low cover fractions of a small and inconspicuous moss species (Skowronek et al. 2016). This particularly challenging example is presented in the tutorial (Fig. 3).

Projecting species’ future distributions has become an important tool to manage alien plant invasions (Rocchini et al. 2017). These models (cf. SDMs) produce maps depicting the areas of potential invasion risk, the areas that require close surveillance and those that require mitigation actions (Guisan et al. 2013). In this section, guidelines to run SDMs are provided as well as a set of observations to run SDMs for IAS in a “best practices” framework. This section also combines the outcomes from the LiDAR data processing with those derived from the systematic sampling design proposed in the toolbox.
The last part in the applications’ section deals with the assessment of the impact of plant invasions on the ecosystem functioning. The approach focuses on the impact of IAP species establishment on the nutrient status of the native plant communities (Aerts et al. 2017). In a first step, a combination of LiDAR and hyperspectral data is used to predict measurements of chemical leaf traits into space. In a second step, information from the resulting prediction maps is used to compare canopy nutrient concentrations between invaded and non-invaded sites.

Data resources

The demonstration dataset was generated to provide an easy-to-download and easy-to-use real-world data set including ground surveys, hyperspectral and LiDAR data of actual alien plant invasion cases. One initial challenge when using LiDAR and hyperspectral data is the large size of the files that often require high computer power. For example, the original hyperspectral data file of one of our study site was 27 Gb and the raw point cloud LiDAR file was 413 Gb, which are typical dataset sizes for airborne imagery. To overcome this high demand of computational power, we developed a checkerboard approach (Fig. 4) that reconstructs a hyperspectral/LiDAR image using only information from areas relevant for model calibration and validation, together with information from a set of background areas. This way we derived snippets from the original images that were in a second step rearranged to an image mosaic (Fig. 5). The resulting reconstructed image is similar to a checkerboard where each cell consists of the extracted hyperspectral/LiDAR values for a specific plot. After the reconstruction, plot field data were relocated to match the spatial locations in the reconstructed image and provide a complete but much lighter dataset with real hyperspectral and LiDAR data, and ground data.
The dataset consists of two such reconstructed hyperspectral images with 248 spectral bands and spatial resolutions of 1.8 m x 1.8 m (Fig. 5A) and 3 m x 3 m (Fig. 5B) for the island of Sylt (Germany) and for the forest of Compiègne (France), respectively. A set of LiDAR-derived rasters with the same spatial resolutions and a corresponding table of plot-based field data include repositioned geographical coordinates (in the reconstructed image space), grid categories (i.e. calibration, validation, background) and the percentage cover of the invasive alien moss species *Campylopus introflexus* (Hedw.) Brid., 1819 (Sylt island) and the invasive alien tree species *Prunus serotina* Ehrh., 1788 (forest of Compiègne). Additional field data are provided for the forest of Compiègne on the community weighted mean of leaf phosphorus and nitrogen concentration for plots with varying native and alien invasive species abundance. The original remote sensing data were acquired in two flights and field campaigns within the DIARS project. For Sylt island, an APEX (Airborne Prism Experiment) sensor covering a spectral range between 412 and 2432 nm was used by The Flemish Institute for Technological Research (VITO) to acquire
The DIARS toolbox: a spatially explicit approach to monitor alien plant ...  

the hyperspectral images during July 2014. VITO also preprocessed the hyperspectral images with geometric calibration, correction of spectral smile effects and atmospheric correction (Schaepman et al. 2015). Aerodata France acquired and preprocessed airborne discrete return LiDAR data across the island of Sylt with an average point density of 23 points/m² across the island of Sylt, also in July 2014. For the forest of Compiègne, an APEX sensor covering a spectral range between 380 nm and 2500 nm was used by VITO to acquire the hyperspectral images during July 2014. Aerodata France acquired and preprocessed airborne discrete return LiDAR data across the forest of Compiègne with an average point density of 14 points/m² in February 2014.

The toolbox is transparent and open, allowing for changes and customizations to fit other datasets and sources, and also presents the method to create training data, via the “virtualspecies” R package (Leroy et al. 2015), for SDMs (Hattab et al. 2017) and other applications.

Accessing the toolbox

The toolbox can be accessed at http://diarsproject.github.io/DIARS/HomeDIARS.html. The dataset can be downloaded from the site and all the tutorials have been developed using the following free and open source software (FOSS): R and GRASS GIS (Neteler et al. 2012).

Acknowledgements

This study is part of the project DIARS (Detection of invasive plant species and assessment of their impact on ecosystem properties through remote sensing) funded by the ERA-Net BiodivERsA, with the national funders: ANR; BelSPO; and DFG. The authors would like to thank the Office National des Forêts (ONF) for granting permission for leaf sampling and for providing the airborne LiDAR data.

References

• Bivand R (2016) rgrass7: Interface Between GRASS 7 Geographical Information System and R. R package version 0.1-8. URL: https://CRAN.R-project.org/package=rgrass7


