



**Project Title:** ECOPOTENTIAL: IMPROVING FUTURE ECOSYSTEM BENEFITS THROUGH EARTH OBSERVATIONS

**Project number:** 641762

**Project Acronym:** ECOPOTENTIAL

**Proposal full title:** IMPROVING FUTURE ECOSYSTEM BENEFITS THROUGH EARTH OBSERVATIONS

**Type:** Research and innovation actions

**Work program topics addressed:** SC5-16-2014: "Making Earth Observation and Monitoring Data usable for ecosystem modelling and services"

Deliverable No: 5.5

## Datasets following standard requirements

**Due date of deliverable:** 31<sup>st</sup> May 2018

**Actual submission date:** 8<sup>th</sup> June 2018

**Version:** 1.1

**Main Authors:** Samuel Hoffmann, Johannes Peterseil, Dimitris Poursanidis, Carl Beierkuhnlein



This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 641762



<b>Project ref. number</b>	<b>641762</b>
<b>Project title</b>	<b>ECOPOTENTIAL: IMPROVING FUTURE ECOSYSTEM BENEFITS THROUGH EARTH OBSERVATIONS</b>

<b>Deliverable title</b>	Datasets following standard requirements
<b>Deliverable number</b>	5.5
<b>Deliverable version</b>	1
<b>Contractual date of delivery</b>	31 <sup>st</sup> May 2018
<b>Actual date of delivery</b>	8 <sup>th</sup> June 2018
<b>Document status</b>	Update after review comments
<b>Document version</b>	1.1
<b>Online access</b>	<a href="http://www.ecopotential-project.eu/products/deliverables.html">http://www.ecopotential-project.eu/products/deliverables.html</a>
<b>Diffusion</b>	Public
<b>Nature of deliverable</b>	Report
<b>Workpackage</b>	5
<b>Partner responsible</b>	UBT
<b>Author(s)</b>	Samuel Hoffmann, Johannes Peterseil, Dimitris Poursanidis, Carl Beierkuhnlein
<b>Editor</b>	Guy Ziv
<b>Approved by</b>	Johannes Peterseil
<b>EC Project Officer</b>	Gaëlle Le Bouler

<b>Abstract</b>	The report focuses on the evaluation of representativeness of existing data sets and products and aims to provide recommendation for the harmonisation of in-situ data collections and EO data. For this a critical review on the availability of existing in-situ data on protected areas is conducted and an evaluation of datasets used in the storylines is made. In order to address the spatial, temporal and thematic linking of in-situ data with EO results from a case study on plant communities showcase the relationship between in-situ and remote sensing data. Based on this specific recommendations for the future acquisition of ground information in connection with EO can be derived. Finally, other case studies are introduced to further test data, and account for in-situ data gaps.
<b>Keywords</b>	In-situ, storyline, DEIMS-SDR, remote sensing, field work, harmonisation





## Table of content

1	Introduction .....	9
2	An overview of in-situ data involved in ECOPOTENTIAL.....	10
2.1	Documentation of in-situ data.....	10
2.2	Available in-situ data.....	11
2.3	In-situ data usage.....	17
3	Linking in-situ and remote sensing data.....	18
3.1	Conceptual background .....	18
3.2	Remote sensing signals represent beta diversity among plant community types in a semi-natural system: a case study from a UNESCO Man and Biosphere Reserve.....	19
3.2.1	Introduction .....	19
3.2.2	Methods.....	21
3.2.3	Results.....	25
3.2.4	Discussion .....	29
3.2.5	Conclusion.....	31
4	Identification of in-situ data gaps.....	33
5	Discussion and recommendations.....	44
6	References .....	46
7	Appendix .....	51



## List of figures

- Figure 1 The different levels of metadata as from the DEIMS-SDR Metadata Information Model. .... 10
- Figure 2. The number of available data sets per ECOPOTENTIAL storyline. The data sets were grouped by main categories. The storylines refer to three main ecosystem types: mountain ecosystem (M), (semi-)arid ecosystem (A) and aquatic ecosystem (O). .... 12
- Figure 3 . The number of data sets documented in DEIMS-SDR per ECOPOTENTIAL Protected Area. .... 15
- Figure 4. The number of categorical data sets in DEIMS-SDR per ECOPOTENTIAL Protected Area. .... 15
- Figure 5. The number of data sets per ECOPOTENTIAL protected area as described by scientific publications related to the project. .... 16
- Figure 6. The categorical data sets per ECOPOTENTIAL protected area as described by scientific publications related to the project in May 2018. .... 16
- Figure 7. Problem of precise positioning, errors and linking remote sensing and in-situ data. .... 18
- Figure 8. With the increasing availability of high resolution (fine grain) remote sensing data there is a growing need for precise localization of in-situ data in order to link these two levels of spatial information. GPS errors and other sources of limits in spatial precision require novel approaches such as calculated surfaces of similarity between plots / pixels. .... 19
- Figure 9. Location of sampling plots on La Palma, Canary Islands. The entire island is a UNESCO Man and Biosphere Reserve. The contour lines indicate elevational steps of 200 m. The plots that include *Pinus canariensis* were classified as “Pine forest (PF)” (green). Plots below the pine forest without *Pinus canariensis* were classified as “Succulent scrub (SC)” (yellow), and plots above the pine forest without *Pinus canariensis* as “Subalpine scrub (SA)” (blue). .... 21
- Figure 10. Flow chart describing the remote sensing and in-situ data as well as the statistical analyses to evaluate the relationship between both. For details see Methods section. .... 24
- Figure 11. The location of plots in the two-dimensional ordination space calculated via Non-metric Multidimensional Scaling (NMDS) and Principal Component Analysis (PCA). Species names and abbreviations are listed above the introduction of the main text. a) The PC-rotated NMDS space is representing beta diversity calculated by the Hellinger distance between plots, considering the abundances of perennial plant species. The NMDS-stress value of 0.06 depicts a good fit. b) The PC-rotated PCA space is calculated by the remote sensing (RS) variables derived from the Sentinel-2A image taken on 14 January 2017. A proportion of 63% of total variance is explained by PC1 (39%) and PC2 (24%). The vectors of explanatory variables (brown arrows) and PCA-input variables (black arrows) were fitted after generating the ordination space. .... 26
- Figure 12. Time series analysis of Sentinel-2A (S2A) images and sensitivity analysis concerning grain size. In a) the correlation results between the S2A variables of 13 images with 20 m grain size (see Table S1, Appendix) and the beta diversity are shown. “Ns” highlights non-significant ( $p \geq 0.05$ ) correlation results. The S2A image from 14 January 2017 indicates the strongest correlation in view of the three statistical tests (Multivariate Analysis of Variance, Mantel test, Variation Partitioning). Therefore, this S2A image was used for the sensitivity analysis in b). Here, a grain size of 20 m yields highest correlation between beta diversity and S2A variables, whereas 60 m leads to highest correlation between beta diversity and LiDAR variables. Two out of three



statistical tests reveal strongest correlation between beta diversity and all RS variables for 20 m spatial resolution..... 28

Figure 13. Classification of plots according to remote sensing variables. K-means classification techniques were applied to determine three community classes that are illustrated within the species-based two-dimensional space of Non-metric Multidimensional Scaling (NMDS): K-means clustering involves a) all RS variables b) only Sentinel-2A variables or c) only variables derived from Light Detection and Ranging. Multivariate Analysis of Variance (MANOVA) was used to quantify the explanatory power of classification results explaining the beta diversity, i.e. the Euclidean distances between plots in the two-dimensional NMDS space. For details see Methods section. 29

## List of tables

Table 1 List of publications used for the analysis of data used.....	14
Table 2. Relevant in-situ research by the Biogeography Department of the University of Bayreuth in ECOPOTENTIAL.....	33
Table 3. Overview on in-situ datasets and data products included in DEIMS-SDR and scientific publications related to ECOPOTENTIAL protected areas. ....	51
Table 4. Available in-situ data sets and products described by the ECOPOTENTIAL storylines for arid/semi-arid ecosystems (A), coastal/marine ecosystems (C) and mountain ecosystems (M).....	54



<empty page>



## Executive summary

The use of in-situ data is an important element in workflows developed for the validation and calibration of both modelling approaches as well as the development of derived earth observation data products. Collecting data at the appropriate scale with respect to thematic, spatial and temporal resolution is critical. Datasets collected for different purposes lack this appropriate link. Identifying shortcomings of existing in-situ data sources as well as defining requirements for data collection in terms of usage readiness are therefore important tasks.

The current work is focusing on the evaluation of representativeness of existing data sets and products. A critical review and check of existing information and data provided on ECOPOTENTIAL sites, storylines and categorical data sets is conducted. A case study on plant communities illustrates the relationship between in-situ and remote sensing data to deliver specific frameworks for future efficient acquisition of ground information in connection with EO. Finally, we introduce other case studies conducted by the Biogeography Department of the University of Bayreuth. These case studies include further test data, and account for in-situ data gaps.



<empty page>





## 1 Introduction

Using earth observation techniques in order to improve ecosystem benefits involves the implementation of in-situ data for calibration/validation activities (see TERN 2018). In contrast to earth observation data, in-situ data collection is performed at the actual object of interest (e.g. ecosystem, site, individual tree) or material (soil, rock) and is thus also referred to as ground data. Nevertheless, a wide range of data collection methods can be applied ranging from human based observations (e.g. vegetation surveys), sensor based in-situ observations (e.g. soil moisture) to stationary or air-craft based remote sensing techniques (e.g. LiDAR). In-situ data is then linked to remote sensing data to calibrate and validate model algorithms and image classification for Land Use Land Cover (LULC) analysis, for instance. In-situ data are an essential part of spatial data analysis. Uncertainties and spatial inaccuracies of the underlying input data will therefore strongly influence the results and quality of the output of each analysis. However, pre-existing in-situ data often lack the appropriate thematic, spatial and/or temporal resolution – as often collected for other purposes.

The quality of in-situ data influences the overall model performance. Quality evaluation schemes are consequently indispensable. Alonso et al. (2017, D5.3)<sup>1</sup> already elaborated on the development and implementation of user-friendly quality evaluation schemes for an efficient assessment of the quality of pre-existing data sets based on metadata evaluation following the principles of ISO 19157 and ISO19158. This deliverable profits from pre-existing tools of other projects (BIO\_SOS, EU BON) and from open source software, in which fitness-for-use evaluation tools were developed and tested with a strong input from data users. Deliverable 5.3 facilitates the quality-driven selection of appropriate data for spatial models and the identification of data (quality) gaps. These data (quality) gaps can then lead to the planning of new and targeted data collections.

The current work is focusing on the evaluation of representativeness of existing data sets and products and providing recommendations for EO targeted in-situ data collection. This includes a critical review of the existing information and data provided by protected areas and used in the analysis focusing on the storylines. A case study on plant communities illustrates the relationship between in-situ and remote sensing data to deliver specific frameworks for future efficient acquisition of ground information in connection with EO.

Where necessary, test data will be recorded, and missing data will be accomplished. We finally identified in-situ data gaps that are accounted for by field work of the Biogeography Department of the University of Bayreuth.

---

<sup>1</sup> See <http://www.ecopotential-project.eu/images/ecopotential/documents/D5.3.pdf>

## 2 An overview of in-situ data involved in ECOPOTENTIAL

### 2.1 Documentation of in-situ data

In the context of the current work “in-situ data” are defined as data that is directly collected within or sampled at a specific location using defined data collection methods. This could range from human based observation (e.g. biodiversity surveys), sensor based observations (e.g. meteorological stations) or remote sensing techniques (e.g. LiDAR). In-situ data are seen in contrast to Earth Observation data which are collected from outside the ecosystem using satellite or airborne tools (such as Drones) implemented at different levels. Within the project context data collection is focused on the ECOPOTENTIAL protected areas. In this respect biodiversity surveys and meteorological stations are considered as in-situ, whereas satellite based land cover classifications are not.

In order to ensure the discoverability and reusability of data metadata are needed. Metadata are used to document available in-situ data. Thereby, the context of the data but also limitations of data usage can be described. Depending on the domain a number of different metadata standards exist. For spatial and observation data resulting from the environmental domain ISO19115/19139 is fostered by the INSPIRE directive<sup>2</sup> and can be seen as standard. Within ECOPOTENTIAL DEIMS Site and Dataset Registry (DEIMS-SDR, <https://deims.org/>) was used to maintain and query metadata for data generated in the project context. DEIMS-SDR addresses not only the level of the dataset but also addresses the organisational context of the observation, like the research site or observation facilities. An overview of the different levels of metadata is provided in Figure 1.

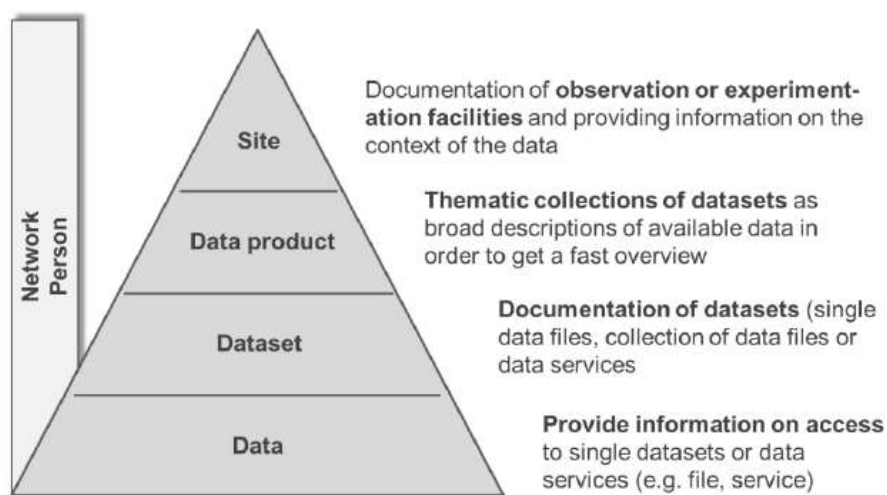


Figure 1 The different levels of metadata as from the DEIMS-SDR Metadata Information Model.

<sup>2</sup> See <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32008R1205&from=EN>



In DEIMS-SDR the data are structured by different levels of metadata, i.e. levels of information. At the highest level (i.e. smallest information content) the site describes the study area, from which the data is coming from. Concerning ECOPOTENTIAL data, these sites are protected areas and corresponding observational and experimental facilities such as LTER stations inside protected areas. The data product comprises a thematic collection of datasets. The dataset offers more information on single or multiple data files and/ or services. The lowest metadata level (i.e. largest information content) delivers access to data files and services. The detailed description of the metadata model is provided in the documentation section of DEIMS-SDR (<https://deims.org/models/>).

The data of ECOPOTENTIAL protected areas described in DEIMS-SDR are stored by data storage formats and data services. Data storage formats involve relational databases, structured files and spreadsheets, spatial data files, spatial databases, scientific publications, unstructured files, XML databases and proprietary file format. Data services comprise data portals, web catalogue service (CSW), web coverage service (WCS), web feature service (WFS), web map service (WMS) among other data services. Thereby, a minority of five ECOPOTENTIAL protected areas work with data portals and only one to three protected areas implemented data services. Catalogue services are barely used. The majority of 19 protected areas store their data offline and provide the data on request via email or phone. Eight protected areas even provide online data services on request. Moreover, the ECOPOTENTIAL protected areas manage the data differently. Fifteen protected areas implemented a central management location, whereas two protected areas organize their data at different locations in the same institute. Five protected areas distributed their data between institutions. Another five protected areas did not describe any data management information.

## 2.2 Available in-situ data

The ECOPOTENTIAL storylines are narratives that were developed by the ECOPOTENTIAL members. They serve as guidelines for the scientific, the management and policy output. The storylines are distinguished by three ecosystem types; mountain ecosystems (M), aquatic ecosystems (O), and (semi-)arid ecosystems (A) (see D5.1). The storylines particularly focus on the required remote sensing and in-situ data to build ecosystem models that fulfil demands of stakeholders and decision-makers. The storylines are expected to develop during the project. However, we screened the current versions for in-situ data and extracted the information on in-situ data sets (see Appendix). The available in-situ data sets were grouped per storyline and main categories according to their context (see Figure 2).

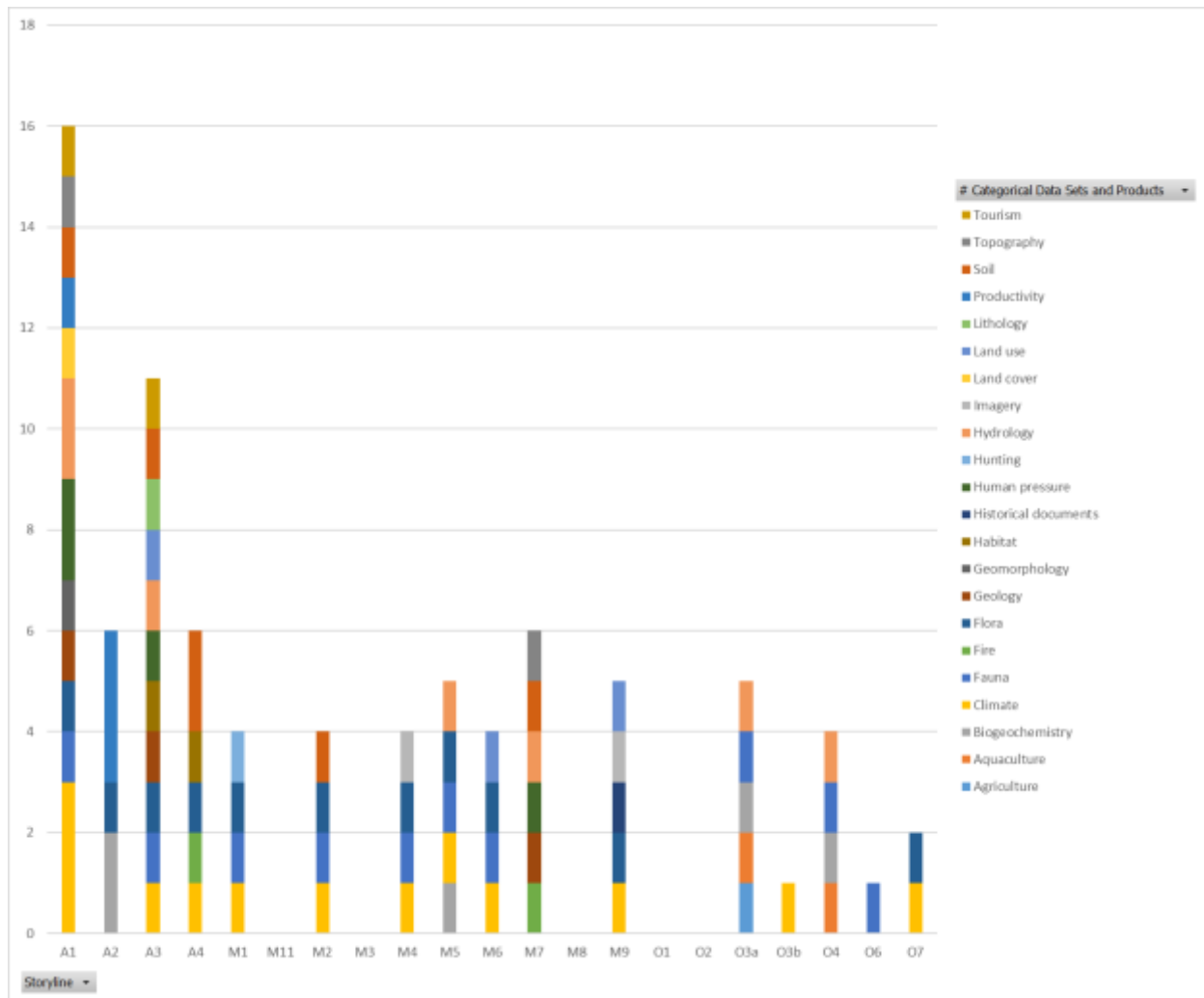


Figure 2. The number of available data sets per ECO POTENTIAL storyline. The data sets were grouped by main categories. The storylines refer to three main ecosystem types: mountain ecosystem (M), (semi-)arid ecosystem (A) and aquatic ecosystem (O).

Data available for ECO POTENTIAL protected areas can be discovered via the web-based data catalogue DEIMS-SDR and the ECO POTENTIAL Virtual Laboratory or Research Environment, respectively (VRE). Deliverable D5.1<sup>3</sup> section 3.3.2 provides a first overview of existing in-situ data sets and data products provided by ECO POTENTIAL protected areas. For this overview, data sets included in the storylines and documented within DEIMS-SDR have been considered.

Here this analysis acts as an up-to-date complement. It has to be noted that the in-situ data description in the storylines does not always give clear information whether the data are actually available or just required, but not existing yet. Concerning DEIMS-SDR we screened all ECO POTENTIAL protected areas and searched for provided in-situ data. In order to allow for a better characterisation a grouping according to (a) the ecosystem characteristics addressed (abiotic and biotic), (b) the ecosystem domain observed (marine and terrestrial), and (c) the type of data generation conducted (in-situ, remote earth observation, models).

<sup>3</sup> See <http://www.ecopotential-project.eu/images/ecopotential/documents/D5.1.pdf>



In addition the datasets were grouped according to data categories. Each dataset was classified according to these three categories.

We additionally grouped the data by the following overarching categories: flora fauna, climatology, geology, imagery, habitats, hydrology, topography, land cover, soil chemistry, agriculture, land use, fire, chemistry, ecosystem services. When possible, we provide information on the variable, the data format, the spatial resolution, the temporal resolution, the data usage, the source, the responsible contact and the work package addressed (see Appendix). A reference list is given below that includes the ECOPOTENTIAL publications that involve in-situ data. In Table 3 in the Appendix the source data for the analysis is provided. For each of the classifications an analysis is provided showing the number of datasets for each category.

Figure 3 and Figure 4 are focusing on datasets and data products documented within DEIMS-SDR. The data were collected using queries from the discovery portal of DEIMS-SDR. Figure 3 focuses on domain and type of data provided and shows the number of categorised datasets and data products per protected area contributing to a selected storyline. Figure 4 provides an overview on the thematic grouping of data. The figure also shows the number of datasets and data products per protected area contributing to a selected storyline.

Figure 5 and Figure 6 are focusing on datasets described in scientific publications related to ECOPOTENTIAL. The list of datasets was analysed and categorised according to the above mentioned groups. Figure 5 shows the number of datasets assigned to the domain and type of observation. Figure 6 provides an overview on the thematic categories of data used in the publications. Considering only these publications, none of the protected areas included presents more than one data product of a particular category (Fig. 6). A review on all ECOPOTENTIAL publications revealed the following studies to include in-situ data of any kind. In total a number of 10 publications were taken into account for the analysis.



Table 1 List of publications used for the analysis of data used

- Bustamante J., Aragonés D., Afán I., Luque C.J., Pérez-Vázquez A., Castellanos E.M., & Díaz-Delgado R.* (2016). Hyperspectral Sensors as a Management Tool to Prevent the Invasion of the Exotic Cordgrass *Spartina densiflora* in the Doñana Wetlands. *Remote Sensing*, 8(12), 1001. doi:10.3390/rs8121001
- Choudhury B.U., Ferraris S., Ashton R.W., Powlson D.S., & Whalley W.R.* (2018). The effect of microbial activity on soil water diffusivity. *European Journal of Soil Science*, 2018, doi:10.1111/ejss.12535
- Díaz-Delgado R., Aragonés D., Afán I., & Bustamante J.* (2016). Long-Term Monitoring of the Flooding Regime and Hydroperiod of Doñana Marshes with Landsat Time Series (1974–2014). *Remote Sensing*, 8(9), 775, doi:10.3390/rs8090775
- Filipponi, F., Valentini, E., & Taramelli, A.* (2017). Sea Surface Temperature changes analysis, an Essential Climate Variable for Ecosystem Services provisioning. In *Analysis of Multitemporal Remote Sensing Images (MultiTemp)*, 2017 9th International Workshop on the (pp. 1-8). IEEE.
- Hoffmann, S., Schmitt, T., Chiarucci, A., Irl, S., Rocchini, D., Vetaas, O.R., Tanase, M.A., Mermoz, S., Bouvet, A., & Beierkuhnlein, C.* (2018). Remote sensing of beta diversity: evidence from plant communities in a semi-natural system, *Applied Vegetation Science*, doi:10.1111/avsc.12403
- Hummel, C., Provenzale, A., van der Meer, J., Wijnhoven, S., Nolte, A., Poursanidis, D., Janss, G., Jurek, M., Andresen, M., Poulin, B., & Kobler, J.* (2017). Ecosystem services in European protected areas: Ambiguity in the views of scientists and managers?. *PloS one*, 12(11), e0187143.
- Raffelli, G., Previati, M., Canone, D., Gisolo, D., Bevilacqua, I., Capello, G., Biddoccu, M., Cavallo, E., Deiana, R., Cassiani, G., & Ferraris, S.* (2017). Local-and Plot-Scale Measurements of Soil Moisture: Time and Spatially Resolved Field Techniques in Plain, Hill and Mountain Sites. *Water*, 9(9), 706, doi:10.3390/w9090706
- Tomaselli, V., Adamo, M., Veronico, G., Sciandrello, S., Tarantino, C., Dimopoulos, P., Medagli, P., Nagendra, H. & Blonda, P.* (2017). Definition and application of expert knowledge on vegetation pattern, phenology, and seasonality for habitat mapping, as exemplified in a Mediterranean coastal site. *Plant Biosystems-An International Journal Dealing with all Aspects of Plant Biology*, 151(5), 887-899.
- Turco, M., Bedia, J., Di Liberto, F., Fiorucci, P., von Hardenberg, J., Koutsias, N., Llasat, M.-C., Xystrakis, F., & Provenzale, A.* (2016). Decreasing Fires in Mediterranean Europe. *PLoS ONE* 11(3), e0150663, doi:10.1371/journal.pone.0150663
- Valentini, E., Filipponi, F., Nguyen Xuan, A., Passarelli, F.M., & Taramelli A.* (2016). Earth Observation for Maritime Spatial Planning: Measuring, Observing and Modeling Marine Environment to Assess Potential Aquaculture Sites. *Sustainability*, 8(6), 519, doi:10.3390/su8060519

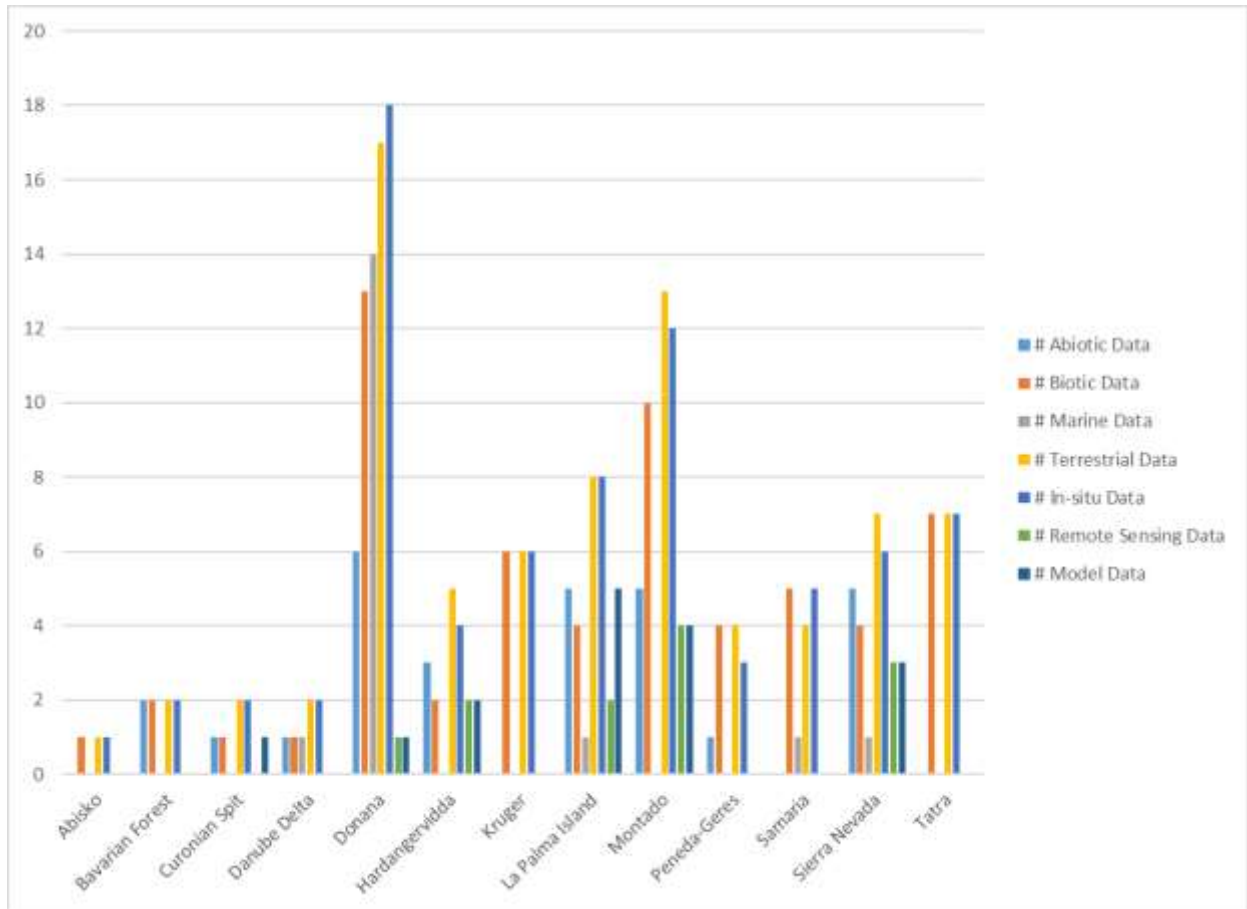


Figure 3 . The number of data sets documented in DEIMS-SDR per ECO-POTENTIAL Protected Area.

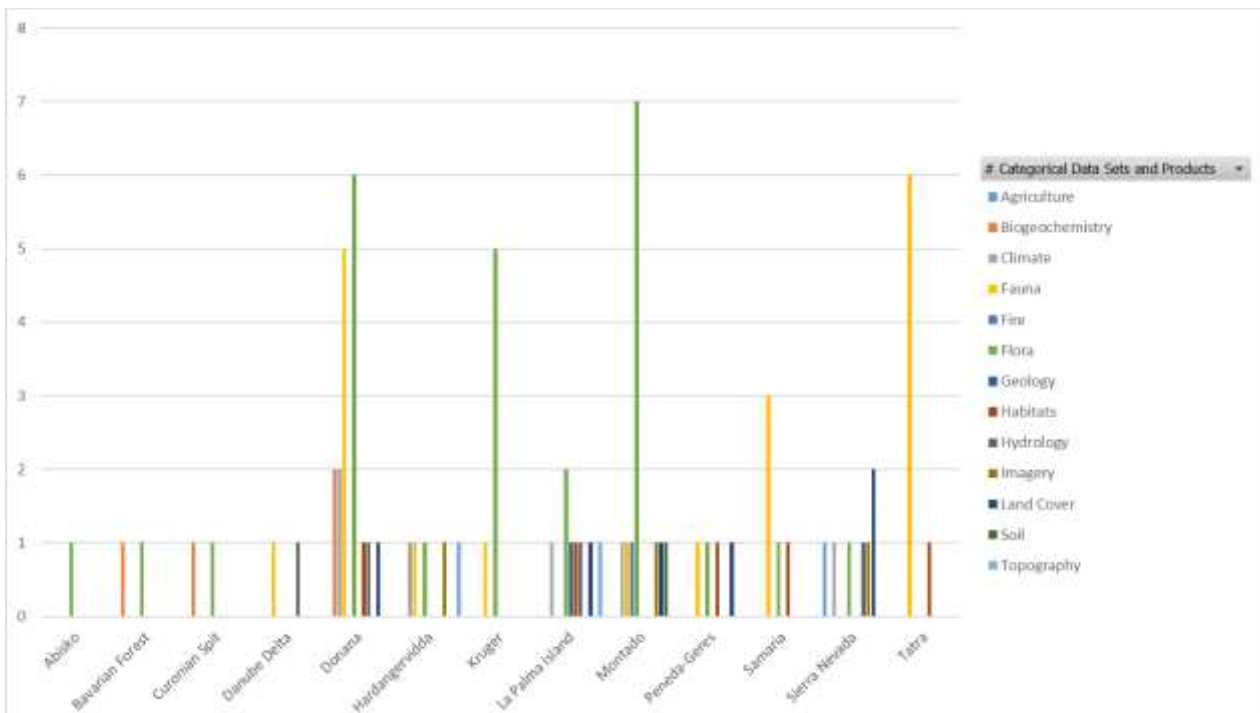


Figure 4. The number of categorical data sets in DEIMS-SDR per ECO-POTENTIAL Protected Area.

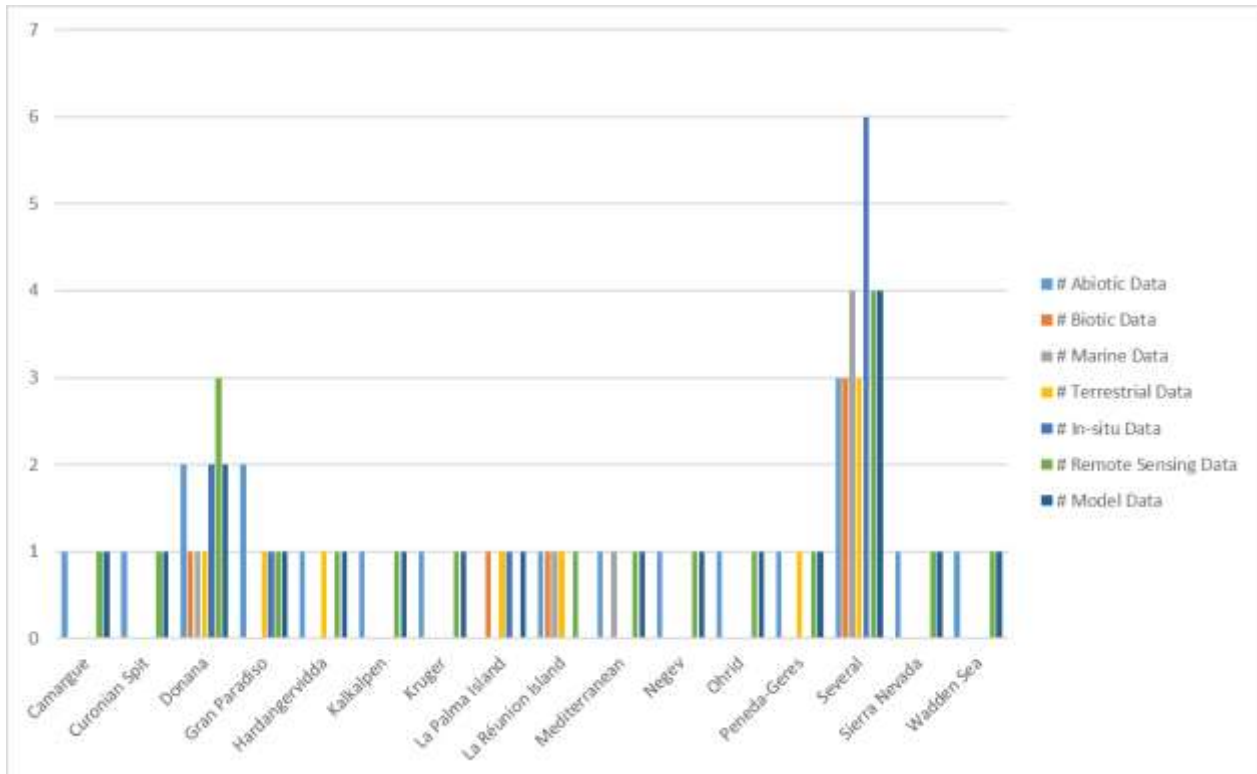


Figure 5. The number of data sets per ECO POTENTIAL protected area as described by scientific publications related to the project.

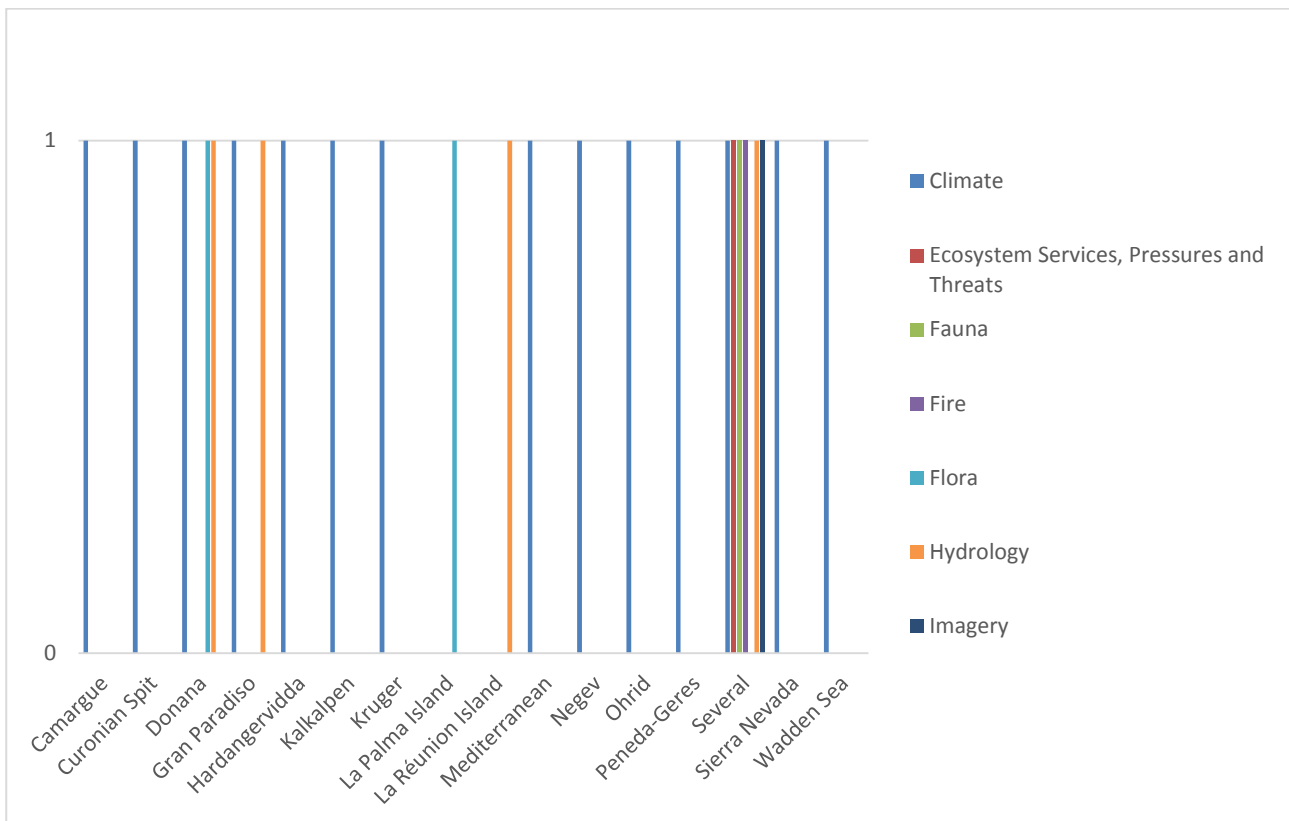


Figure 6. The categorical data sets per ECO POTENTIAL protected area as described by scientific publications related to the project in May 2018.





## 2.3 In-situ data usage

In the context of Earth Observation, in-situ data are used for ground-truth EO data products on the one side (e.g. WP3 and WP4) and to perform model calibration (WP) on the other. In-situ data are generally used for model development and validation (WP6). An overview of existing models considered within ECO POTENTIAL WP6 is provided in Deliverable D5.1 Table 3.1.

### 3 Linking in-situ and remote sensing data

#### 3.1 Conceptual background

Dependent on the time scale, species group, national regulations, and the type of assessments, previous data collection has been carried out with differing motivation, accuracy, grain size, spatial density, precision in localization etc. The connection of ground information with EO has to acknowledge these biases and the problems that emerge in the face of increasing spatial resolution (see Figure 7 and Figure 8). In former remote sensing approaches, where spatial resolution (pixels) was 100 or 250 m or even less (larger pixels), such problems were just ignored. Now with the high resolution of Sentinel products and other sensors, new algorithms for spatial adjustment of different levels of information need to be developed.

Geo-statistical approaches will be applied to standardize data quality (relation to Task 5.5) and to deliver specific frameworks for future efficient acquisition of ground information in connection with EO. This Task is closely linked to WP3, WP4. Where necessary, test data will be recorded and missing data will be accomplished.

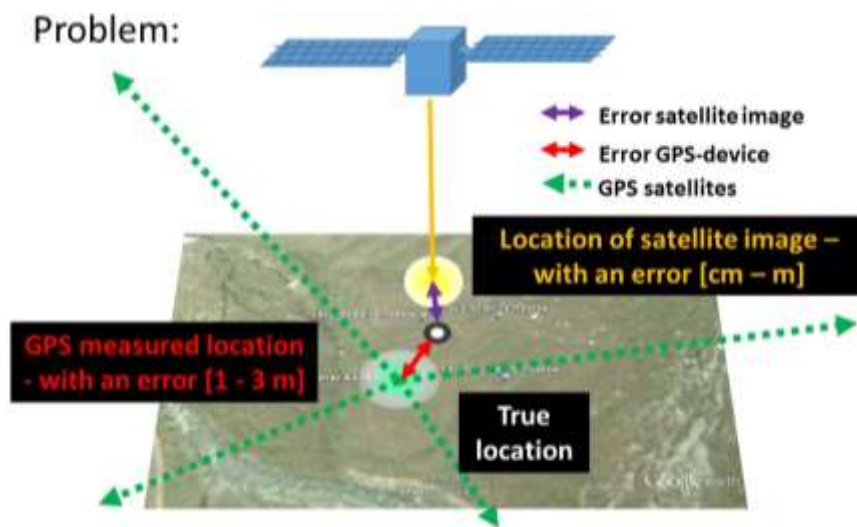


Figure 7. Problem of precise positioning, errors and linking remote sensing and in-situ data.

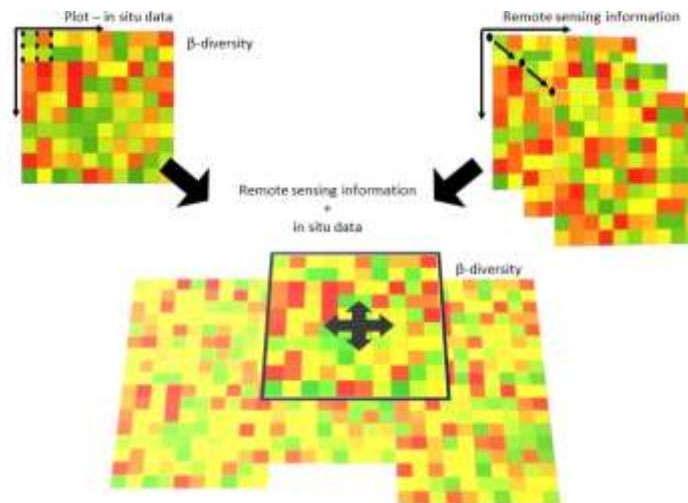


Figure 8. With the increasing availability of high resolution (fine grain) remote sensing data there is a growing need for precise localization of in-situ data in order to link these two levels of spatial information. GPS errors and other sources of limits in spatial precision require novel approaches such as calculated surfaces of similarity between plots / pixels.

In the following we present a case to on the degree to which remote sensing signals reflect beta diversity, i.e., the heterogeneity in species composition. Beta diversity is a crucial component of species diversity, and thus biodiversity, since beta diversity is a measure of dissimilarity between species assemblages that ultimately determines species diversity in general. Understanding beta diversity means understanding species distribution in space and time. Biodiversity conservation profits from remote sensing technique, when species distributions can be successfully represented by remote sensing signals; an approach that is less time consuming and costly than field sampling and monitoring over long time and large geographical extent.

## 3.2 Remote sensing signals represent beta diversity among plant community types in a semi-natural system: a case study from a UNESCO Man and Biosphere Reserve

### 3.2.1 Introduction

The spatial and temporal patterns in species composition are at the heart of community ecology ever since Clements (1916). And yet, the community definition is still largely debated (Palmer and White 1994, Chiarucci 2007, Ricklefs 2008). The controversy revolves around the coherence and integrity of ecological entities through different scales of space and time (Jax 2006). Community models are generally deterministic or stochastic, and community distinction is either discrete or continuous. In this study, community types are considered as operational units, i.e. as the set of plants coexisting in a given unit of space and time, which can show regularities and assembly rules at any spatial and temporal scale (Chiarucci 2007). Processes responsible for observed patterns of species coexistence, usually referred to as “assembly rules”, are manifold, interrelated and contingent, which led Lawton (1999) to call community ecology “a mess”. For reasons of clarification, Vellend (2010) proposes the following four overarching, categorical processes shaping species communities: selection, drift, speciation and dispersal.



The existence of discrete community units implies the delimitation of community types. As natural boundary sharpness varies (Wilson and Agnew 1992, Auerbach and Shmida 1993), community limits are associated with transition zones also known as ecotones (Livingston 1903). In early vegetation science, an ecotone was referred to as the delimitation of community physiognomy (Clements 1905). Contrary, the definition of Lloyd and colleagues (2000) involves beta diversity, by perceiving an ecotone as a “zone where directional change in vegetation (i.e. qualitative and quantitative species composition) is more rapid than on the other side of the zone.” Although ecotones are a standard entity in landscape ecology (Wiens et al. 1992), Hufkens et al. (2009) point out that they do not have standardized spatial and temporal units.

The complexity of community definition implies that comprehensive field sampling and monitoring for community determination is time consuming and costly. Since plant communities may be distinguished by the plants’ physiognomy, remote sensing (RS) can be a powerful tool to identify plant community types over large extent, in short time and at low costs (Rocchini et al. 2016). RS provides data that reveals biodiversity patterns from local to global extent as well as temporally resolved. RS is used to detect changes in community composition, with changes in spectral diversity as a measure of beta diversity (Rocchini et al. 2005). This application of RS rests on the spectral variation hypothesis (SVH) explaining the relationship between environmental heterogeneity, species diversity and spectral information (Palmer et al. 2002). Environmental heterogeneity increases habitat heterogeneity and, thus, biodiversity (i.e. habitat-heterogeneity hypotheses; Simpson 1949). Environmental heterogeneity is expected to increase with spectral heterogeneity. Therefore, spectral variation is associated with alpha and beta diversity of communities (Palmer et al. 2002, Rocchini et al 2004). However, the SVH does not apply to all ecosystems (Schmidtlein and Fassnacht 2017).

In this study, we question the agreement of plant community patterns emerging from species composition with the RS patterns of the same communities. For this test, a continuous elevational gradient on the subtropical-Mediterranean island of La Palma (Canary Islands) was selected as study case. In the study region, literature describes three vegetation types along elevation; succulent shrub, pine forest and subalpine (summit) shrub (del Arco Aguilar et al. 2010). This may also result from dominating plant physiognomies (i.e. scrubs or trees) that is apparent at first sight. But does such strict classification really conform to compositional dissimilarities considering abundances of all occurring perennial vascular plant species? And does such strict classification conform to community patterns that were detected by remote sensing? In comparison to other studies on the SVH, we also use structural RS parameters from light detection and ranging (LiDAR) next to multispectral RS signals from a time series of Sentinel-2A images. By applying and comparing ordination and classification techniques to identify plant community types, we test whether remotely sensed community patterns reflect in-situ observations. To our knowledge, this combination of data sources and techniques has not been used before, to reveal and compare community patterns. Since RS products can hardly account for understory species diversity invisible for many RS sensors, we suppose that RS-based community patterns do not accurately conform to patterns arising from ground surveys. The different methodological approaches are expected to demonstrate similar patterns due to similar underlying algorithms.

## 3.2.2 Methods

### 3.2.2.1 Vegetation Data

We applied a stratified random sampling along the elevation gradient, constrained by aspect (east to west), slope ( $<20^\circ$ ), avoidance of anthropogenic land use, and accessibility and selected as many sampling sites as possible. Due to ridges and steep slopes some sampling sites appear linearly arranged (Fig. 9). In each sampling site, a 10 m x 10 m plot was used to record plant community data. We recorded abundances of all vascular plant species within the plot, by using their coverage within three standardized vegetation strata (tree, shrub and herb layer).

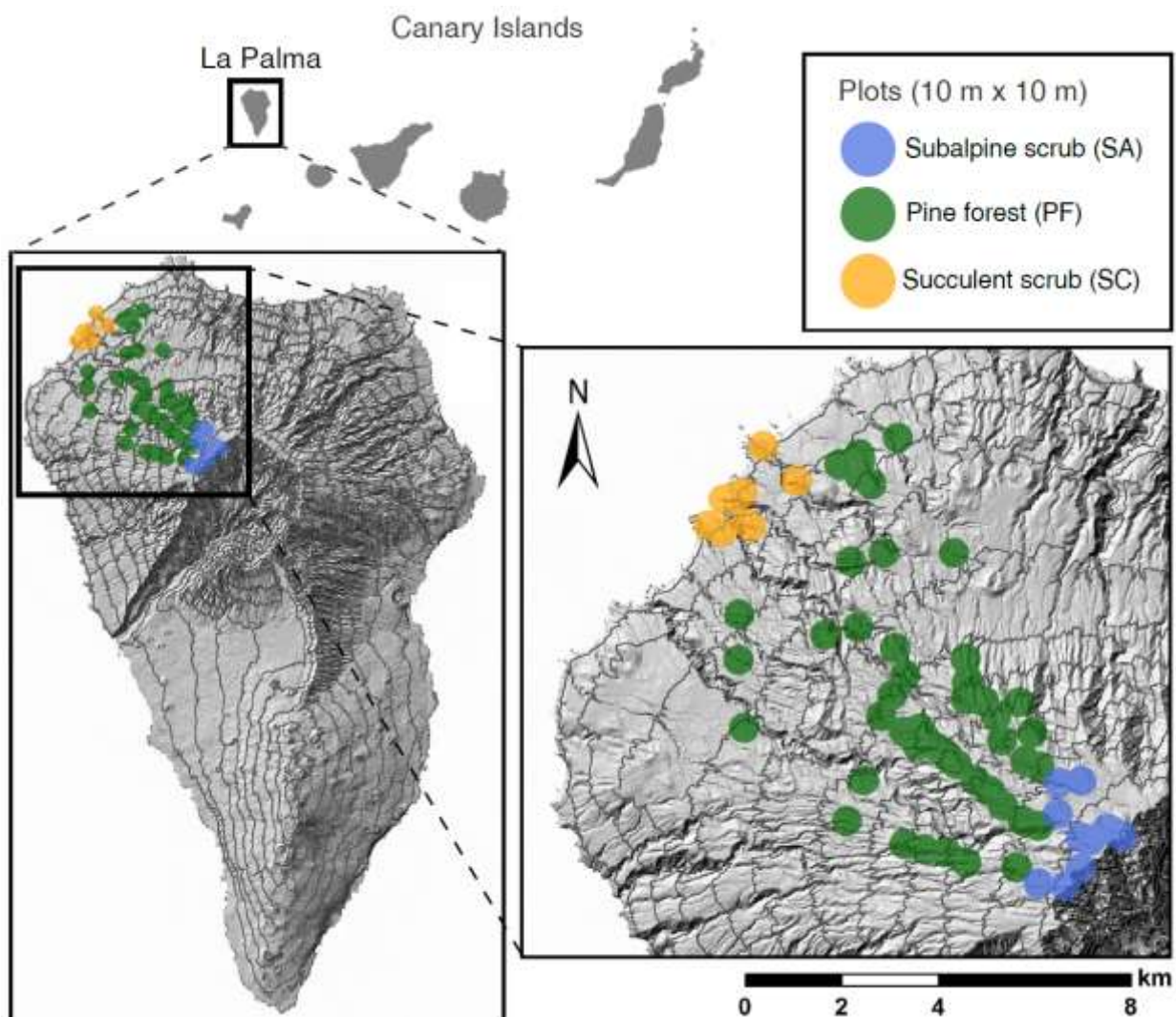


Figure 9. Location of sampling plots on La Palma, Canary Islands. The entire island is a UNESCO Man and Biosphere Reserve. The contour lines indicate elevational steps of 200 m. The plots that include *Pinus canariensis* were classified as “Pine forest (PF)” (green). Plots below the pine forest without *Pinus canariensis* were classified as “Succulent scrub (SC)” (yellow), and plots above the pine forest without *Pinus canariensis* as “Subalpine scrub (SA)” (blue).



Since the presence of annual plants is driven by season, and can be stochastically driven by short term weather events, we consider only the perennial plant species. As the vegetation of the island is dominated by perennial plants, this can be justified. Annuals emerge only for short time periods. During our study, they covered rarely more than 1 % of the surface. Bare soil differs in importance between > 90% in desert-like ecosystems in the south to less than 5 % in humid places. In addition, the variation of their presence during the seasons makes it very difficult to conduct reliable comparison with remote sensing data that are recorded at a different time, even if the time difference is relatively small.

Relative abundance per species and plot was calculated as the species' cover abundance divided by the sum of cover abundances of all species in all strata. As not only trees were involved, correction for basal area was not applied. By this definition of relative abundances, it is possible to resolve changes in species composition between plots, because land cover types other than vegetation (i.e. bare soil, rock, litter) are neglected. If other than vegetation cover types were considered we would notice a reduction in absolute species abundances, even if the relative species composition remains constant. However, such cover classes influence the composition of RS signals. Thus, we did estimate the coverage of bare soil, rock and litter that is not covered by any other strata (i.e. detectable for RS sensors) and account for that in RS-specific analyses. Uncertainty remains about the short-term performance of annual plant species after and during periods of favourable conditions. These issues of annual vegetation cover are relevant for all Mediterranean and desert ecosystems.

### 3.2.2.2 Remote Sensing Data

We considered several RS products that are appropriate to distinguish plant communities (Xie et al. 2008, Pettorelli et al. 2014). We selected 13 Sentinel-2A images (S2A; Copernicus Sentinel Data 2017), covering the time period from February 2016 to February 2017 (see Table S1, Appendix). We can consequently account for multispectral differences that may occur during the seasons. We chose S2A, since this sensor provides images of high radiometric (twelve bands), temporal (five days revisit time) and spatial resolution (10 to 60 m) that are publicly available and free of charge (for details see [https://sentinel.esa.int/documents/247904/685211/Sentinel-2\\_User\\_Handbook](https://sentinel.esa.int/documents/247904/685211/Sentinel-2_User_Handbook)). The downloaded images were given as a geometrically and radiometrically corrected Top-of-Atmosphere (TOA) Level-1C product. We applied atmospheric, terrain and bidirectional reflectance distribution (BRDF with cosine of local solar zenith angle) correction using the Sen2Cor plugin (see <http://step.esa.int/main/third-party-plugins-2/sen2cor/>) within the Sentinel-2 toolbox of the Sentinel Application Platform (SNAP). We thus reveal Bottom-Of-Atmosphere (BOA) Level-2A products. These products include a scene classification that was used to mask pixels classified as “medium cloud probability”, “high cloud probability” and “cirrus”. The cloud mask covered a maximum of two plots per image. Band 1 (aerosol, 60 m), Band 9 (water vapour, 60 m) and Band 10 (cirrus, 60 m) were removed by the preprocessing procedure. The remaining bands are Band 2 (blue, 10 m), Band 3 (green, 10 m), Band 4 (red, 10 m), Band 5 (red edge, 20 m), Band 6 (red edge, 20 m), Band 7 (red edge, 20 m), Band 8 (near-infrared [NIR], 10 m), Band 8a (red edge, 20 m), Band 11 (shortwave infrared [SWIR], 20 m) and Band 12 (shortwave infrared [SWIR], 20 m). We also calculated the normalized differentiation vegetation index (NDVI) by  $(\text{Band 8} - \text{Band 4}) / (\text{Band 8} + \text{Band 4})$ . The NDVI is one of the most appropriate proxies for primary productivity that varies between sampled plant communities



during the seasons (Pettorelli 2013), and is useful to explain plant species richness and rarity (Levin et al. 2007).

Metrics derived from airborne light detection and ranging (LiDAR) are able to account for both 2D and 3D vegetation structure, which helps to distinguish vegetation that differs in structural variables such as growth height and canopy cover (Pettorelli et al. 2014). Airborne laser scanning (ALS) point cloud data from April 2009 was downloaded from the Spanish National Geographic Institute (IGN) with a spatial resolution of 0.5 points per 1 m<sup>2</sup> (see Appendix). After pre-processing, the data consist of several indices with a grain size of 20 m. The canopy height model (CHM) returns the average of normalized heights above ground. The tree fraction cover (TFC) is the proportion of first ALS returns over 2 m above ground from the total amount of first ALS return in the raster cell. The vegetation fraction (VF) reflects the number of all returns over 0.5 m height divided by the number of all returns within the cell. The return proportion (RP) indices were calculated as the number of ALS returns in different strata (0.5 m, 2 m, 5 m, 10 m, 15 m, 20 m, 25 m) divided by the total number of ALS returns in the cell. Thus, RP informs about the three-dimensional vegetation structure. The effective leaf area index (LAI) was computed based on the gap probability, but not corrected for woody elements or the clumping effect. For classifications based on RS data all variables were standardized to zero mean and unit variance. Euclidean distances were applied to retrieve distance between these standardized RS variables.

To reduce the bias induced by GPS inaccuracy for the extraction of RS data by plot centroids, we use RS data with a minimum grain size of 20 m. We therefore aggregate RS data by taking the mean. In addition, we evaluated the results sensitivity to coarser grain sizes (40 m and 60 m). Data processing and statistical analyses were conducted using open-source R Statistics (R Core Team 2017, Version 1.0.136) and corresponding default settings, if not mentioned differently.

### **3.2.2.3 Statistical Analyses**

We prepared a flow chart that summarizes our methodological approach analyzing the relationship between in-situ and RS variables (Figure 10). In order to describe the given plant communities and demonstrate the species' realized environmental niches, we modelled the coenoclines of the ten most abundant species. A coenocline is a response curve of the species abundance along a single gradient (Whittaker 1967). Species with overlapping coenoclines form communities. We applied two environmental gradients: mean annual temperature and mean annual precipitation. Coenoclines were generated by fitting generalized additive models (GAM) with Gaussian distribution and link function, and thin plate regression splines as the single penalty smooth class (Wood 2017; R-function `gam()` in package "mgvc"). Because we were facing unequal sample sizes between community types, we conducted non-parametric Kruskal-Wallis Analysis of Variance (ANOVA; R-function `kruskalmc()` in package "stats") to identify differences in species richness. Linear regression models (R-function `lm()` in package "stats") were applied to determine the relationship between species richness and environmental gradients. Model assumptions were verified visually.

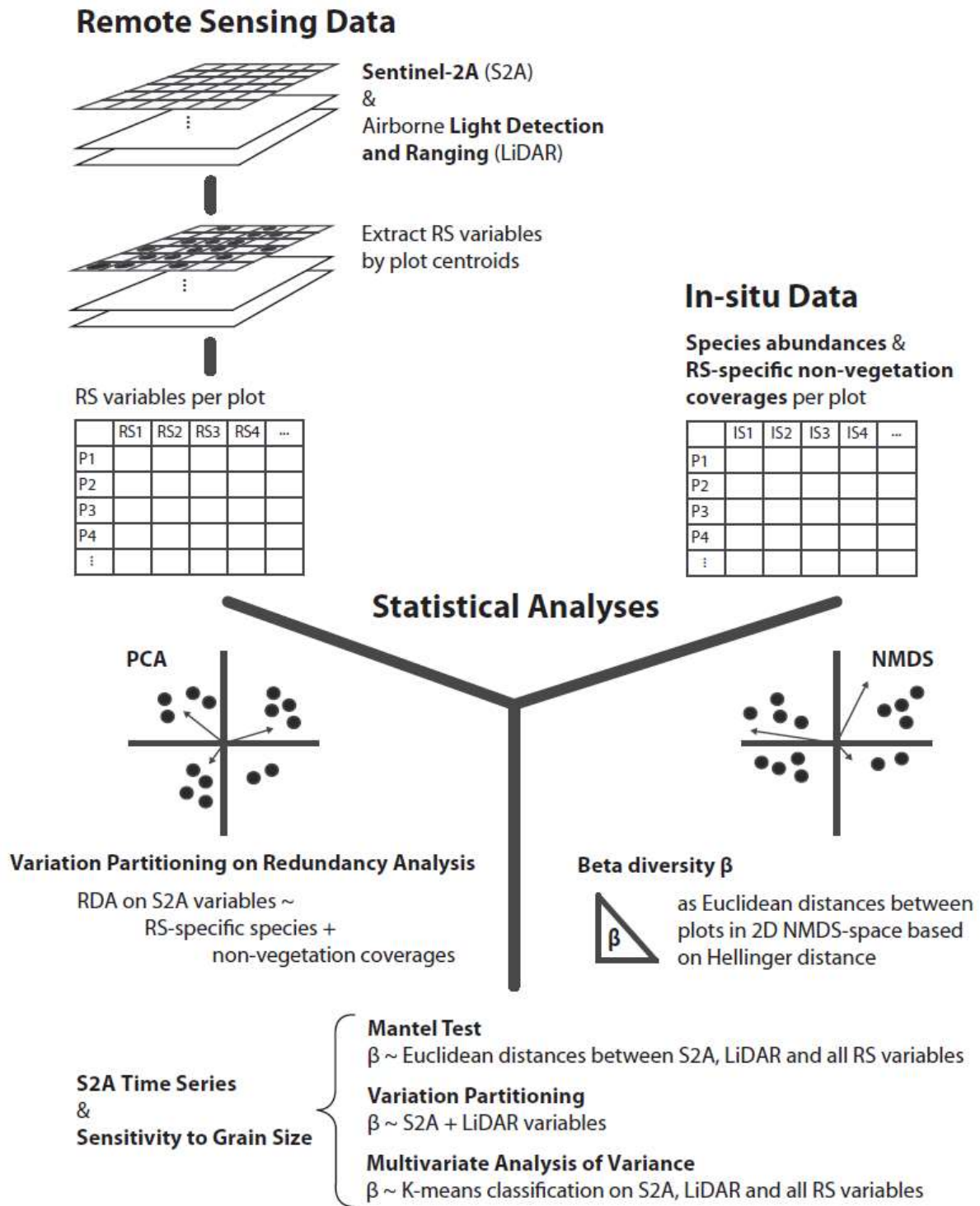


Figure 10. Flow chart describing the remote sensing and in-situ data as well as the statistical analyses to evaluate the relationship between both. For details see Methods section.

Beta diversity can be understood as the dissimilarity between plots regarding their species composition (Whittaker 1967). We applied Nonmetric Multi-Dimensional Scaling (NMDS) to assess beta diversity and distinguish plant communities (Legendre and De Caceres 2013). The NMDS is a distance-based, indirect ordination technique. We avoid direct ordination methods, since we are interested in unconstrained results





that rest on compositional dissimilarity only (McCune and Grace 2002). The NMDS ranks distances between input data (plots). Therefore, NMDS bypasses the linearity assumptions of metric ordination methods. Here, we apply Hellinger distance to calculate the distance matrix among plots regarding their species composition (Legendre and De Caceres 2013). The Hellinger distance downweights the occurrence of rare species. Thus, we control for overrated influence of rare species in dissimilarity calculations. We calculated a two-dimensional ordination space running 100 tries and involving random starting configurations, to find the optimal solution by NMDS, i.e. the lowest stress value (R-function `metaMDS()` in package “vegan”). The NMDS-space was rotated to principal components, i.e. most variation in the data is shown along the first axis, followed by the second. We conducted post-hoc correlation of explanatory variables to the NMDS via surface and vector fitting (R-function `ordisurf()` and `envfit()` in package “vegan”), to interpret the influence of explanatory variables onto the compositional dissimilarity represented by the location of plots in the NMDS-space. We eventually calculate beta diversity as the Euclidean distances between plot locations in the two-dimensional NMDS space.

Subsequently, we utilized the Mantel test to analyse the relationship between beta diversity and RS variables (R-function `mantel()` in package “stats”; Pearson’s correlation coefficient, 1000 permutations). Moreover, variation partitioning was used to reveal the combined and independent effects of S2A and LiDAR variables explaining the beta diversity (R-function `varpart()` in package “vegan”). Variation partitioning is based on a Redundancy Analysis (RDA), linearly modelling the relationship between a set of dependent variables and two sets of explanatory variables. We also employed K-means unsupervised classification algorithm (R-function `kmeans()` in package “stats”; 1000 iterations of random starting configurations) to distinguish three community types considering RS variables only. We aimed at creating three classes, because existing vegetation maps predefine three main community types in the study region: succulent scrub, pine forest and subalpine scrub. K-means algorithm has been used before to test the SVH (Schmidtlein and Fassnacht 2017). We then conducted Multivariate Analysis of Variance (MANOVA) to estimate how K-means classification on RS variables fits to the beta diversity (R-function `adonis()` in package “vegan”).

Furthermore, we applied a Principal Component Analysis (PCA) to the RS variables to illustrate the variation in RS signals, and to depict the RS products that add most variation (R-function `prcomp()` in package “stats”). As for the NMDS, we applied post-hoc correlation of explanatory variables via vector fitting. In addition, variation partitioning onto a Redundancy Analysis (RDA) was used to separate the variation among S2A variables that can be explained by RS-specific coverage of the ten most abundant species and of non-vegetation cover types (i.e. bare soil, rock, pine needles and deadwood).

### 3.2.3 Results

The NMDS based on the species abundances (Stress=0.06) demonstrates no clear distinction between PF and SA (Figure 11.a). At lower altitudes, a considerable gap between PF and SC does become obvious. Consequently, the similarity in species composition between SA and PF is considerably higher than between SC and PF. Within PF we find an emphasized compositional variation in the lower part close to the transition to the SC. Such variation along the second NMDS axis appears in the subalpine zone as well. The S2A variables of the image from 14 January 2017 correlate, on average, strongest with the beta diversity. In addition, RS data with 20 m grain size correlate the most with beta diversity. Therefore, we focus in the



following on RS variables derived from RS data with 20 m spatial resolution and on the S2A image from 14 January 2017; among these RS variables, Band 3, Band 5, Band 6, Band 7, Band 8, Band 8a, NDVI, RP0.5m, RP2m, RP5m, RP10m, LAI, VF correlate significantly ( $p < 0.05$ ) with the NMDS scores. These variables are mostly associated with the second NMDS axis, which does not distinguish the three community types.

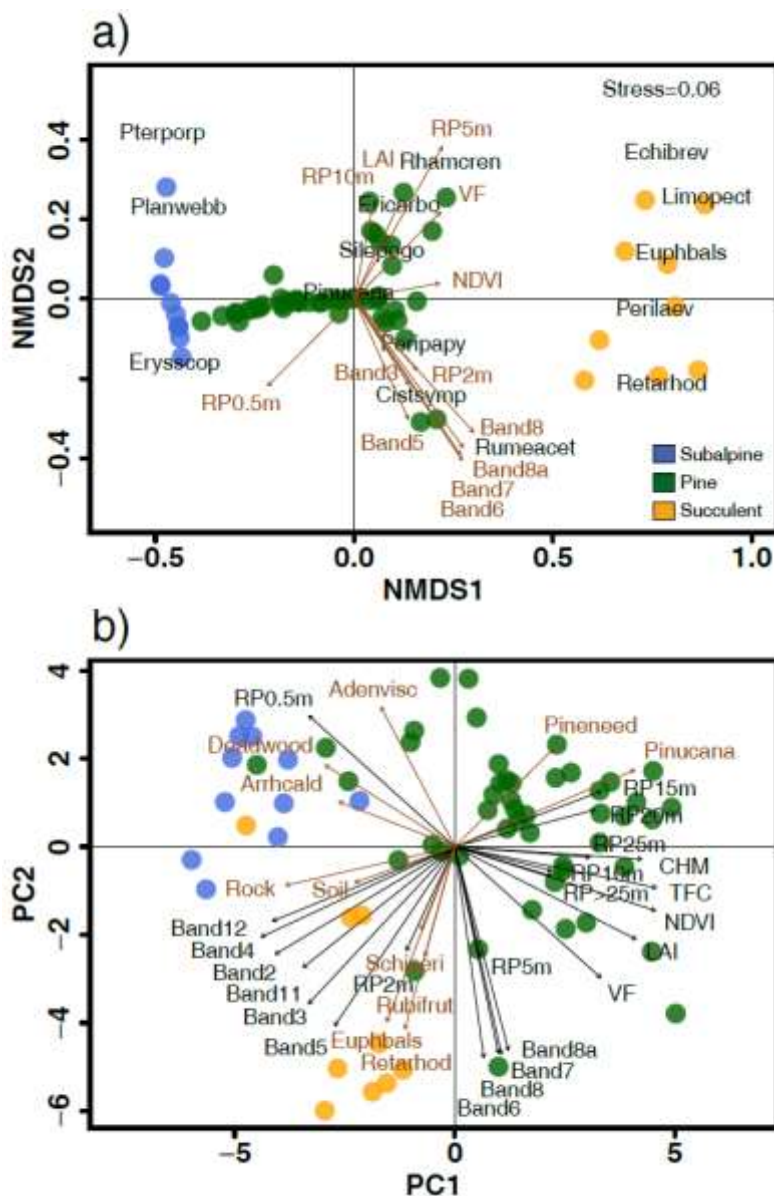


Figure 11. The location of plots in the two-dimensional ordination space calculated via Non-metric Multidimensional Scaling (NMDS) and Principal Component Analysis (PCA). Species names and abbreviations are listed above the introduction of the main text. a) The PC-rotated NMDS space is representing beta diversity calculated by the Hellinger distance between plots, considering the abundances of perennial plant species. The NMDS-stress value of 0.06 depicts a good fit. b) The PC-rotated PCA space is calculated by the remote sensing (RS) variables derived from the Sentinel-2A image taken on 14 January 2017. A proportion of 63% of total variance is explained by PC1 (39%) and PC2 (24%). The vectors of explanatory variables (brown arrows) and PCA-input variables (black arrows) were fitted after generating the ordination space.



The PCA based on RS variables shows that both axes contribute to the differentiation of vegetation types, but the three communities appear poorly separated (Fig. 10b). The distances between SC, PF and SA are not as pronounced as in the species-based NMDS ordination. Among the explanatory variables, only the RS-specific coverage of *Cistus monspeliensis* and *Cistus symphytifolius* are not significantly correlated with the PCA scores. The S2A bands 2 (blue), 3 (green), 4 (red), 11 (SWIR) and 12 (SWIR) are related to SC, whereas gaining reflectance indicated by Band 6 (red edge), 7 (red edge), 8 (NIR) and 8a (NIR) represents PF. The NDVI is closely linked to PF. We find an association between most structural LiDAR variables and PF. The structural variable RP0.5m is related to SA, whereas RP2m reflects SC. The RS-specific coverage of *Pinus canariensis* is strongly correlated with PC1. The RS-specific coverages of succulent and subalpine species are associated with PC2. RS-specific coverage of deadwood is linked to SA, of rock and bare soil to SC as well as SA.

Furthermore, variation partitioning leads to a total  $R^2$  of 0.66 ( $p=0.001$ ) that can be explained by a combination of the RS-specific coverages of species and non-vegetation types (rock, bare soil, deadwood, pine needles). Thereby, RS-specific species' coverages independently account for an  $R^2$  of 0.27 ( $p=0.003$ ), whereas the independent effect of non-vegetation coverages score a non-significant ( $p=0.105$ )  $R^2$  of 0.05. The combined effects of vegetation and non-vegetation coverages result in  $R^2=0.34$ .

Figure 12.a reveals that the S2A variables from 14 January 2017 (20 m grain size) correlate strongest with the beta diversity. We additionally observe a "W"-shape. Consequently, the correlation between S2A variables and beta diversity seems to be stronger during the wet (December-March) and dry season (June-September), compared to other months. Considering all RS variables from 20 m resolution data (Fig. 12b), the Mantel test results in  $r=0.44$  ( $p<0.001$ ). Considering only LiDAR variables yields a Mantel  $r$  of 0.18 ( $p=0.008$ ). Variation partitioning of the beta diversity through a combination S2A and LiDAR variables (20 m grain) leads to a total  $R^2$  of 0.80 ( $p<0.001$ ). The  $R^2$  resulting from independent effects of S2A signals is 0.54 ( $p=0.001$ ). The  $R^2$  of the independent effect of LiDAR signals is 0.01 and not significant ( $p=0.173$ ). The combined effects of S2A and LiDAR variables produce  $R^2=0.25$ . Moreover, the correlation results between S2a variables and beta diversity decrease with increasing grain size (low resolution), while for LiDAR variables the correlation increases (Figure 11.b). Overall, two out of three statistical tests reveal that the correlation between all RS variables and beta diversity is strongest for 20 m grain size.

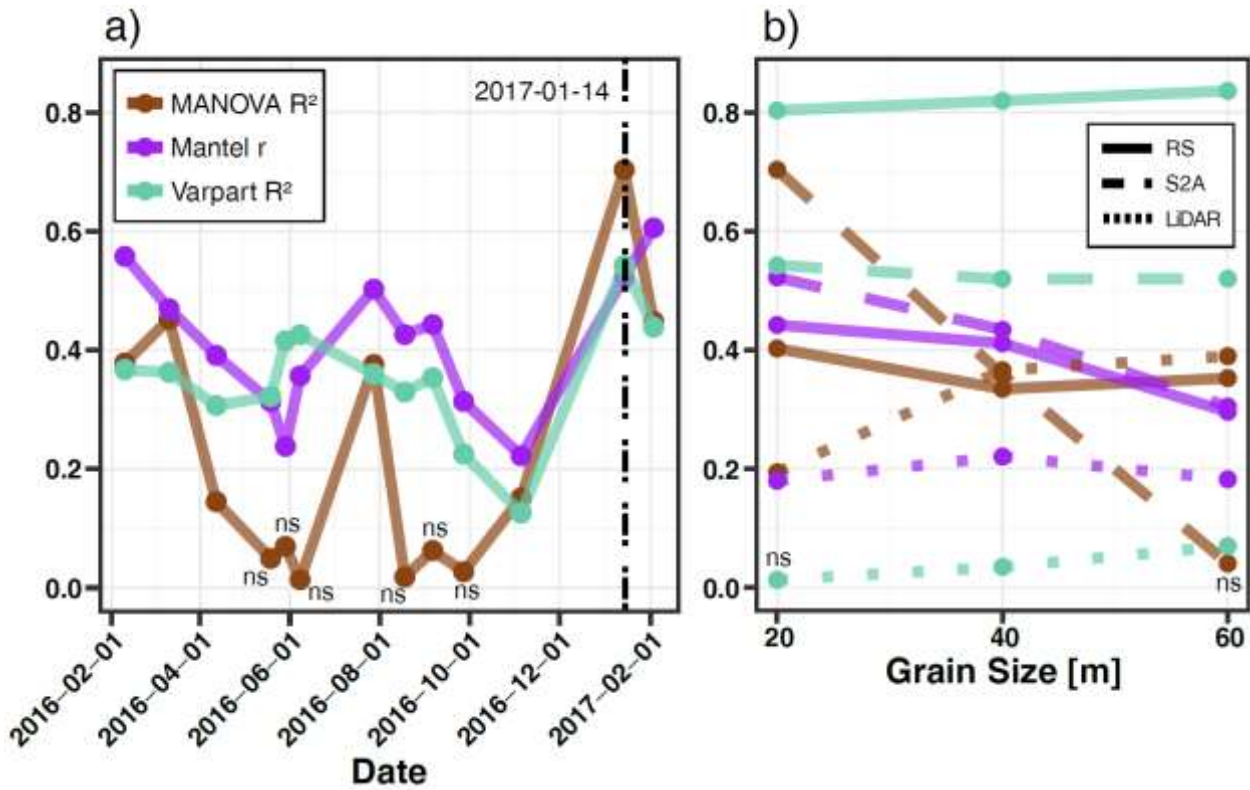


Figure 12. Time series analysis of Sentinel-2A (S2A) images and sensitivity analysis concerning grain size. In a) the correlation results between the S2A variables of 13 images with 20 m grain size (see Table S1, Appendix) and the beta diversity are shown. “Ns” highlights non-significant ( $p \geq 0.05$ ) correlation results. The S2A image from 14 January 2017 indicates the strongest correlation in view of the three statistical tests (Multivariate Analysis of Variance, Mantel test, Variation Partitioning). Therefore, this S2A image was used for the sensitivity analysis in b). Here, a grain size of 20 m yields highest correlation between beta diversity and S2A variables, whereas 60 m leads to highest correlation between beta diversity and LiDAR variables. Two out of three statistical tests reveal strongest correlation between beta diversity and all RS variables for 20 m spatial resolution.

Applying K-means classification algorithm to all RS variables leads to three classes that can moderately explain the dissimilarities in species composition (Figure 13.a; MANOVA:  $p < 0.001$ ,  $R^2 = 0.40$ ). When considering S2A variables only (Figure 13.b), K-means classification outcomes adequately reflect the beta diversity ( $p < 0.001$ ,  $R^2 = 0.70$ ), which is also illustrated by marginally overlapping class-polygons drawn into ordination space. A classification solely based on LiDAR variables yields a much worse fit (Figure 13.c;  $p = 0.001$ ,  $R^2 = 0.19$ ).

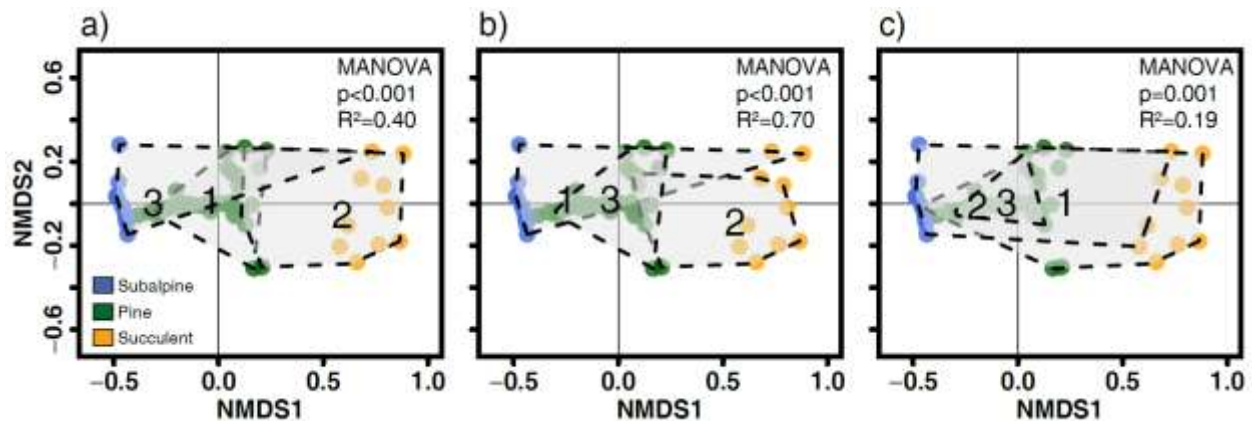


Figure 13. Classification of plots according to remote sensing variables. K-means classification techniques were applied to determine three community classes that are illustrated within the species-based two-dimensional space of Non-metric Multidimensional Scaling (NMDS): K-means clustering involves a) all RS variables b) only Sentinel-2A variables or c) only variables derived from Light Detection and Ranging. Multivariate Analysis of Variance (MANOVA) was used to quantify the explanatory power of classification results explaining the beta diversity, i.e. the Euclidean distances between plots in the two-dimensional NMDS space. For details see Methods section.

### 3.2.4 Discussion

Other than expected, our results demonstrate that a combination of remotely sensed, multispectral and structural variables is able to adequately represent beta diversity as it emerges from in-situ sampling of plant species composition. Thereby, multispectral S2A variables account for much more explanatory power than structural LiDAR variables. Particularly in the wet, but also in the dry season, multispectral S2A signals are strongly correlated to the dissimilarity in species composition. Increasing grain size increases the explanatory power of LiDAR variables, but decreases the power of S2A signals. Moreover, *Pinus canariensis* is a key species in the study region and predominantly determines community discrimination as expressed by RS and in-situ data. At the lower ecotone, the beta diversity between the succulent scrub and the pine forest is high, whereas the diversity of RS signals is low. At the upper ecotone, the difference in species composition between the pine forest and the subalpine scrub is low as is the divergence of RS variables. In general, perennial species richness linearly decreases from low to high elevation, even though precipitation increases. Succulent scrub consists of considerably more species than both other vegetation types.

We show that a combination of multispectral and structural RS variables explains over 80 % of beta diversity in the study system. The S2A variables constitute much more explanatory power than the LiDAR variables. These outcomes are in line with similar studies that consider different scales. He et al. (2009) quantified the relationship between NDVI-distances (derived from MODIS with 250 m resolution) and plant beta diversity (using pairwise Bray-Curtis dissimilarity) within entire US counties. The highest Mantel  $r$  was achieved at the species level ( $r=0.4$ ); see He & Zhang (2009) for a similar approach at the global scale. Hall and colleagues (2012) used multispectral variables derived from QuickBird imagery with a grain size of 2.4 m. They applied variation partitioning on grassland beta diversity (local-to-regional richness ratio), sampled in 0.5 m plots representative for larger sites, which resulted in an  $R^2$  of 0.27 for the independent effect of multispectral RS variables. That is lower than the explanatory power we found, although their study scale is much smaller.



Obviously, the different extents of pixels and plots affect the correlation between RS signals and beta diversity. On the one hand, pixels larger than the plot extent imply a mixture of spectral signals that do not only stem from the plot extent (Nagendra et al. 2010). On the other hand, applying a sampling design with pixels smaller than the plot extent implies either to sample vegetation in larger plots or to use RS data with higher spatial resolution (Rocchini et al. 2010). Vegetation plots larger than 10 m by 10 m are rarely applied in vegetation ecology, because the sampling effort is huge, particularly in open vegetation types (Chytrý & Otýpková 2003). Moreover, as Rocchini (2007) demonstrates, the Mantel  $r$  of 0.69 of the correlation between species diversity sampled in 10 m by 10 m plots and QuickBird data with much smaller spatial resolution (3 m) is not considerably larger than our findings; high-resolution data may contain a considerable amount of noise (Nagendra & Rocchini 2008) and an increase in spectral resolution can also compensate low spatial resolution (Rocchini et al. 2007).

Usually communities that are subject to climate seasonality can be well separated by RS data (Horning et al. 2010). During the wet (December to March) and dry season (June to September), multispectral variables correlate stronger with the dissimilarity in species composition than in other months. These findings may stem from the fact that dominant species of the succulent zone such as *Euphorbia balsamifera* and *Euphorbia lamarckii* are stem succulent and shed their leaves in the dry season, which induces a different multispectral signature of the succulent vegetation. In addition, understory species of the pine forest and subalpine species frequently show discolouration during dry spells, which may spectrally separate vegetation types to high degree. The yellow flowers of the dominating *Adenocarpus viscosus* will also lead to multispectral differentiation of subalpine vegetation in June (Muer et al. 2016). Furthermore, in the wet season, ice-storms can cause discoloration of *Adenocarpus viscosus* (Palomares Martínez et al. 2012).

As we found, multispectral S2A variables explain beta diversity more accurately than structural LiDAR variables. One reason might be that the structural physiognomies of the succulent and subalpine scrub are similar. The growth heights of both vegetation types are about 0.5-2 m on average and vegetation densities are alike. We also observed that both vegetation types are characterized by rocky outcrops and bare soil. However, increasing grain size results in increasing explanatory power of LiDAR variables applying MANOVA, while explanatory power of S2A variables consistently decreases in all statistical tests. We suppose that high variation (noise) in LiDAR variables is the reason for weak correlations with beta diversity at small scales (20 m). The noise is reduced by averaging pixel values, i.e. increasing grain size. Hence, the average structural signatures of entire community types are rather reflected by relatively large grain sizes (60 m), which then lead to more distinct LiDAR-based classes in K-means clustering that correlate stronger with the beta diversity.

High NDVI values correspond to the pine forest, probably indicating high biomass production, where annual precipitation is highest. Most LiDAR-derived structural variables represent the physiognomic forest structure very well (see also Rees 2007, Ørka et al. 2012). The association of structural variables representing different heights above ground (RP variables) with the community types in the PCA ordination agrees with observed vegetation heights in the field. Furthermore, the high reflectance of red light was mostly associated with the succulent and subalpine scrub. This is an indicator for low leaf pigment content and small leaf area, next to of brown rock, soil and litter, (Frampton et al. 2013). In addition, leaf water content is positively related to chlorophyll content (Sims and Gamon 2002). Thus, leaf water content of the succulent and subalpine scrub may be low due to aridity resulting in less chlorophyll and higher reflectance. Indeed, at highest altitudes trade winds prevent the orographic and convective rise of moist air leading to aridity also in the subalpine zone over several months (González Henríquez et al. 1986). Another reason for



high reflectance in the visible spectrum refers to succulent leaf thickness, which prevents light penetration and absorption of lower leaf layers (Sims and Gamon 2002).

A proportion of 66% of variation in S2A signals from 14 January 2017 can be explained by RS-specific coverages of species and non-vegetation cover, but the RS-specific coverages of bare soil, rock and litter barely add to the differentiation of plots based on S2A signals only. Eventually, 34% of variation in S2A variables can neither be explained by the species' coverages, nor by non-vegetation cover types, probably because of differing spatial extents of plots and pixels and GPS-location bias.

GPS-inaccuracy affects the location of RS and in-situ data. For S2A imagery, a GPS-location error of 3, 6 and 18 m is given for 10, 20 and 60 m bands, respectively (Baillarin et al. 2015). Due to the field-sampling conditions (i.e. cloud-free, aspect east to west, slope  $<20^\circ$ , no obstacles), the GPS accuracy of the plot locations could be reduced to a mean of 3.6 m ( $\pm 1.0$  m standard deviation). However, the cardinal direction of the true location shift remains unknown. Thus, a total GPS-error of 6 m for 20m-bands plus the GPS error of the plot locations is possible and likely to cause unexplained variation when correlating RS with in-situ data. However, as the sensitivity analysis shows, the GPS bias seems to be minor, since the lowest grain size of 20 m yields equally high correlation results compared to 40 m and 60 m.

The moderate conformity of RS-based classes with the beta diversity pattern reveals that both sets of variables, S2A and LiDAR, are able to reclassify the pine forest plots, even though unsupervised classifications may be less accurate than supervised techniques (Horning et al. 2010). Interestingly, the S2A variables perform much better here. Both sets seem to contradict each other, because the explanatory power decreases when it comes to defining classes considering a combination of both sets. Therefore, increasing the number of RS variables does not necessarily lead to more variation explained.

Eventually, the here applied RS data were not able to totally resolve the community types and beta diversity in this semi-natural system, which suggest similar RS properties of different species assemblages. Under storey species may be highly abundant and determine beta diversity, but are not detectable for RS sensors. In case heterogeneous, yet distinct plant communities comprise the same spectral signals (Sha et al. 2008), the potential of RS approaches in vegetation science is limited.

### 3.2.5 Conclusion

Our study demonstrated the potential of multiple RS products to represent patterns in plant community composition over large extent, in short time and at low costs. In-situ sampling is thereby indispensable to precisely determine and understand beta diversity and community distinction. The degree of accordance between spectral and beta diversity depends not only on the study system, but also on the methods applied (see also Schmidlein and Fassnacht 2017). Such methods that identify and map discontinuities in beta diversity are necessary for conservation planning and wildlife management (Socolar et al. 2016).

On the one hand, spatial and temporal resolution of RS data may limit the potential of linking field observation with RS data, since interaction between species and environment may occur at scales finer than those RS can deliver. In such cases, other techniques than those applied here may be appropriate (e.g. high spatio-temporal and hyperspectral resolution, space-borne LiDAR), but most high-quality RS data are costly. On the other hand, in-situ data are also often missing. Facing these limitations, project



collaborations are necessary to bring together scientist from ecology and remote sensing to exploit the vast potential of a combination of in-situ data and earth observation for science and conservation practice.





## 4 Identification of in-situ data gaps

Apart from ECO POTENTIAL publications that include in-situ data, several field studies in ECO POTENTIAL protected areas have been conducted by the Biogeography Department of the University of Bayreuth that produced in-situ data. An overview is given in Table 2. The field studies were designed in consideration of the representativeness of existing in-situ data, standardisation of data quality and data gaps. By analysing existing literature and data repositories of any kind, we identified data gaps that have been filled by field studies. The newly collected data are already documented with DEIMS-SDR or will be documented when the scientific work is finished. This led to the completion of the data landscape for ECO POTENTIAL and beyond.

Table 2. Relevant in-situ research by the Biogeography Department of the University of Bayreuth in ECO POTENTIAL.

Title	PA	Authors	Status	Data
Bridging Between Scales: Optimizing Spatial Resolution of Biodiversity Field Data in Alpine Grasslands for Remote Sensing using a Moving Windows Approach	Gran Paradiso	Laura Bethke Carl Beierkuhnlein	Master Thesis completed,  Publication In prep.	Plant Species in three Valleys and three habitats therein  <a href="https://deims.org/dataset/b549ff14-f40f-4749-8e2f-f16f6e523753">https://deims.org/dataset/b549ff14-f40f-4749-8e2f-f16f6e523753</a>
Quantifying beta diversity pattern in alpine grassland in the Gran Paradiso National Park	Gran Paradiso	Jonas Benner Carl Beierkuhnlein	Master Thesis completed,  Publication In prep.	Plant Species (as above)  <a href="https://deims.org/dataset/b549ff14-f40f-4749-8e2f-f16f6e523753">https://deims.org/dataset/b549ff14-f40f-4749-8e2f-f16f6e523753</a>
Mediterranean high mountain flora: Patterns of endemism, species richness and leaf colours along an elevational gradient in the Spanish Sierra Nevada	Sierra Nevada	Pia Eibes David Kienle Carl Beierkuhnlein	Master Thesis completed,  Publication In prep.	Plant Species (along elevational gradients)  <a href="https://deims.org/dataset/9ea0d382-df85-4b0e-b866-263f59ec91cd">https://deims.org/dataset/9ea0d382-df85-4b0e-b866-263f59ec91cd</a>
A closer look from space: The potential of Sentinel-2 data for beta-diversity measurements on small scale in a tundra alpine ecosystem	Hardangervidda	Edvinas Rommel Duccio Rocchini Carl Beierkuhnlein	Master Thesis completed,  Publication In prep.	Plant Species (in a nested design)  <a href="https://deims.org/activity/21d441a9-9230-4d83-bcf8-d1eaae6b817a">https://deims.org/activity/21d441a9-9230-4d83-bcf8-d1eaae6b817a</a> <a href="https://deims.org/dataset/72ed33da-de95-42f6-9023-6448fe81d46a">https://deims.org/dataset/72ed33da-de95-42f6-9023-6448fe81d46a</a>



Title	PA	Authors	Status	Data
Plotsize-decay of beta diversity in an alpine tundra ecosystem	Hardangervidda	Viola Hipler Samuel Hoffmann Carl Beierkuhnlein	Bachelor Thesis completed,  Publication In prep.	Plant Species (as above)  <a href="https://deims.org/activity/21d441a9-9230-4d83-bcf8-d1eaae6b817a">https://deims.org/activity/21d441a9-9230-4d83-bcf8-d1eaae6b817a</a> <a href="https://deims.org/dataset/72ed33da-de95-42f6-9023-6448fe81d46a">https://deims.org/dataset/72ed33da-de95-42f6-9023-6448fe81d46a</a>
Leaf coloration along an elevational gradient on the Island of La Réunion - Exploring the possibilities of field sampling and remote sensing	Réunion	Esther Baumann Carl Beierkuhnlein Dominique Strasberg Erwann Lagabrielle	Master Thesis completed,  Publication In prep.	Shrub species and leaf color  <a href="https://deims.org/dataset/d2ad523b-8105-4e20-b9b6-16136bf66b4a">https://deims.org/dataset/d2ad523b-8105-4e20-b9b6-16136bf66b4a</a>
Development of a marine protected area with increasing shark attacks	Réunion	Jan-Christopher Fischer Carl Beierkuhnlein Erwann Lagabrielle	Master Thesis in progress,  Publication In prep.	Hydrological parameters  In prep.
Finding the optimal spatial resolution for classification of habitat types in protected areas	Bayerischer Wald	Alexander Obermaier Marco Heurich Carl Beierkuhnlein	Master Thesis completed,  Publication In prep.	Habitat types  <a href="https://deims.org/site/letter_eu_de_015">https://deims.org/site/letter_eu_de_015</a>
Correlating diversity patterns and hyper-spectral earth observation	Bavarian Forest	Franziska Hauch Marco Heurich Carl Beierkuhnlein	Bachelor Thesis completed,  Publication In prep.	Plant Species (in a systematic distribution)  <a href="https://deims.org/site/letter_eu_de_015">https://deims.org/site/letter_eu_de_015</a>
Bark beetle outbreaks and inertia to natural forest regeneration	Bayerischer Wald	Philipp Kohler David Kienle Marco Heurich Carl Beierkuhnlein	Master Thesis in progress,  Publication In prep.	Plant Cover, Soil Data  <a href="https://deims.org/dataset/f1b8d1fd-5287-401c-a6b3-27b1906a1a56">https://deims.org/dataset/f1b8d1fd-5287-401c-a6b3-27b1906a1a56</a>
Hydrochemical parameters and Vegetation of helocrenic springs in the Bavarian Forest National Park	Bayerischer Wald	Jamyra Gehler David Kienle Marco Heurich Carl Beierkuhnlein	Master Thesis completed,  Publication In prep.	Higher plants, mosses, water samples  <a href="https://deims.org/dataset/f1b8d1fd-5287-401c-a6b3-27b1906a1a56">https://deims.org/dataset/f1b8d1fd-5287-401c-a6b3-27b1906a1a56</a>



Title	PA	Authors	Status	Data
Spatial patterns of spruce regeneration after bark beetle outbreak in dependence of forest margins	Bayerischer Wald	Stephanie Propp David Kienle Marco Heurich Carl Beierkuhnlein	Bachelor Thesis completed,  Publication In prep.	Tree individuals  <a href="https://deims.org/dataset/f1b8d1fd-5287-401c-a6b3-27b1906a1a56">https://deims.org/dataset/f1b8d1fd-5287-401c-a6b3-27b1906a1a56</a>
Climate change projections and consequences for nature conservation on an oceanic island	La Palma	Severin Irl Carl Beierkuhnlein	In prep.	Endemic plants  <a href="https://deims.org/site/ecopot/es/001">https://deims.org/site/ecopot/es/001</a>
Pennisetum setaceum as a problematic invasive grass species	La Palma	Anna Walentowitz Barbara Zennaro Severin Irl Carl Beierkuhnlein	Data analysis progress  Publication In prep.	Invasive plant species  <a href="https://deims.org/activity/0e1274a4-04e3-4020-b722-7001dc42dc50">https://deims.org/activity/0e1274a4-04e3-4020-b722-7001dc42dc50</a>
Assessing the forest response along treelines to an Epirrita autumnata outbreak in Abisko, using a combination of fieldwork and remote sensing	Abisko	Frank Weiser Carl Beierkuhnlein	Master Thesis completed,  Publication In prep.	Tree individuals  <a href="https://deims.org/dataset/952e2f72-762d-4696-90af-73c9b1de65ff">https://deims.org/dataset/952e2f72-762d-4696-90af-73c9b1de65ff</a>
Identifying drivers of the treeline ecotone in the Alps – a remote sensing and GIS approach	various	Bernadette Menzinger David Kienle Carl Beierkuhnlein	Master Thesis completed,  Publication In prep.	FTP-server
Geographical analysis of habitat richness in European protected areas for conservation of biodiversity.	various	Bravedo Mwaanga Samuel Hoffmann Carl Beierkuhnlein	Master Thesis completed,  Publication In prep.	Habitat types from EEA Data Base  <a href="https://www.eea.europa.eu/data-and-maps/data/article-17-database-habitats-directive-92-43-eeec-1">https://www.eea.europa.eu/data-and-maps/data/article-17-database-habitats-directive-92-43-eeec-1</a>
Remote sensing of beta diversity: evidence from plant communities in a semi-natural system	La Palma	Hoffmann, S., Schmitt, T., Chiarucci, A., Irl, S., Rocchini, D., Vetaas, O., Tanase, M., Mermoz, S., Bouvet, A., & Beierkuhnlein, C.	Journal of Applied Vegetation Science, <a href="https://doi.org/10.1111/avsc.12403">https://doi.org/10.1111/avsc.12403</a>	Plant Species  <a href="https://deims.org/activity/0e1274a4-04e3-4020-b722-7001dc42dc50">https://deims.org/activity/0e1274a4-04e3-4020-b722-7001dc42dc50</a>



Title	PA	Authors	Status	Data
An island view of endemic rarity – environmental drivers and consequences for nature conservation, Diversity and Distributions, 1-11 (2017) doi:10.1111/ddi.12605 Irl, S: Plant diversity on high elevation islands – drivers of species richness and endemism, Frontiers of Biogeography, 8, e29717 (2016) doi:10.21425/F58329717	La Palma	Irl, S; Schweiger, A; Medina, F M; Fernández-Palacios, JM; Harter, D; Jentsch, A; Provenzale, A; Steinbauer, MJ; Beierkuhnlein, C	Diversity and Distributions, 23 (10), 1132-1142 (2017) doi:10.1111/ddi.12605	Plant Species <a href="https://deims.org/site/ecopot_es_001">https://deims.org/site/ecopot_es_001</a>
Plant diversity on high elevation islands – drivers of species richness and endemism	La Palma	Irl, S.	Frontiers of Biogeography, 8, e29717 (2016) doi:10.21425/F58329717	Plant Species (as above) <a href="https://deims.org/site/ecopot_es_001">https://deims.org/site/ecopot_es_001</a>
Climate Change Threats to the Protected Areas of Madagascar	Madagascar	Moenjanahary Ange Raharivoloniaina; Beierkuhnlein, C	Master Thesis completed,  Publication In prep.	On DVD; to be documented with publication.
Uniqueness of Protected Areas for Conservation Strategies in the European Union	various	Hoffmann, S.; Beierkuhnlein, C.; Field, R.; Provenzale, A.; Chiarucci, A.	Scientific Reports 8 (6445), doi:10.1038/s41598-018-24390-3	Reported Plant Species (from EEA Data Base)  <a href="https://deims.org/site/ecopot_es_001">doi:10.1038/s41598-018-24390-3</a>
Inseln als globale Versuchsanordnung und natürliche Laboratorien der Vegetationsökologie	various islands	Beierkuhnlein, C.	Berichte der Reinhold-Tüxen-Gesellschaft, 29 (2017)	Endemic Plant Species  <a href="https://deims.org/site/ecopot_es_001">https://deims.org/site/ecopot_es_001</a>

Information for the unpublished studies listed above is given in the following.

**Leaf coloration along an elevational gradient on the Island of La Réunion - exploring the possibilities of field sampling and remote sensing (Esther Baumann, Carl Beierkuhnlein, Dominique Strasberg, Erwann Lagabriele)**

Colour was and still is fascinating mankind for centuries. In ecology, a broad variety of studies are present, concerning the distinct differences in coloration of different organisms, animals and plants alike. However, the matter of leaf coloration is somewhat neglected: while studies concerning the colour of flowers are found more often, leaf coloration per se triggers not much interest. Mainly studies focus on the chemical reasons that are manifested in colour changes or the agricultural value of leaf colours expressing, for example, the maturity of the crop or indicate malnutrition.



But leaf coloration itself is a morphological and physical property of a plant, which can react to biotic and abiotic drivers. There are studies exploring the possibility of leaf coloration acting as camouflage or repellent against herbivores. And there are as well studies, hinting at different leaf colours as a product of differences in environmental conditions. One prominent gradient to test the plants reaction on changing conditions is the elevation gradient. Incorporation changes of many different abiotic conditions; it can be seen as a proxy for overall harsher environmental condition.

However, the topic of changing leaf coloration as a reaction of an altered physical environment is not well researched. This study therefore aimed at detecting patterns of leaf coloration changes along an elevation gradient and subsequently gaining further knowledge about adaptive strategies of plants.

The study was conducted in the subalpine shrubland of La Réunion, an oceanic island of volcanic origin located in the Indian Ocean. Colours were measured according to the Munsell Colour System. Additionally, Sentinel 2 images were utilized to relate the sampled field data with a dataset derived via remote sensing. Plots were located along a gradient between roughly 2000 and 2900 m.a.s.l. and leaf coloration as well as growth height, species specific cover and leaf area were recorded and analysed.

Results evidently show an overall increasing leaf colour brightness with altitude. However, the patterns seemed to be heavily influenced by few dominant species rather than a change in species composition. The combination of the field data and the satellite images in the form of calculated Vegetation Indices turned out to be problematic: despite several significant relationships found, their reliability is questionable. The direct comparison of the field data and the remote sensing data, translated in a common colour space, showed very clearly, that both colour assessment systems are not picturing the same reality, most likely due to their fundamental difference in how they measure colour.

### **Quantifying beta diversity pattern in alpine grassland in the Gran Paradiso National Park (Jonas Benner, Carl Beierkuhnlein)**

Rapid declining of biodiversity makes an observation of the biodiversity compellingly necessary. The problem with only collecting in-situ data is that it not possible to grant a satisfactory observation. A large scale and long-term observation method should be implemented. ECO-POTENTIAL tries to combine remote sensing and in-situ data to find a successful method for observation. In the Gran Paradiso National Park, in-situ data was collected and used to study three methodological questions for quantifying  $\beta$ -diversity patterns. The focus was on the calculation of  $\beta$ -diversity patterns based on three different amounts of neighbours. The  $\beta$ -diversity patterns were calculated based on two, four and eight neighbours. For the  $\beta$ -diversity calculation with two neighbours, significant differences in heterogeneity were found between the analysed vegetation types. There were, also significant differences between the  $\beta$ -diversity patterns, when using two, four or eight neighbours for calculating the  $\beta$ -diversity. Summarizing, when analysing  $\beta$ -diversity patterns, the amount of neighbours should be adapted to the vegetation type and to the homogeneity of a landscape. Further, calculating with  $\beta$ -diversity the heterogeneity of a landscape could be developed to a useful tool, for long-term observation of changing landscape.

### **Bridging Between Scales: Optimizing Spatial Resolution of Biodiversity Field Data in Alpine Grasslands for Remote Sensing using a Moving Windows Approach (Laura Bethke, Carl Beierkuhnlein)**



The current loss of biodiversity threatens important ecosystem services and hence, human well-being. To understand the mechanisms behind this threat and to conserve biodiversity, monitoring of changes of species communities is crucial. However, in-situ monitoring is cost-intensive, which is why remote sensing plays a growing role in ecological research. Both, field data and remote sensing are scale dependent. The aim of this study is the examination of the behaviour of different alpha diversity measures to increasing scale in field data. Furthermore, Shannon's information entropy was used to find suitable resolutions in remote sensing for studying the biodiversity in grasslands.

Vegetation surveys were conducted in three valleys with (sub-) alpine grasslands of three different types (gravel plains, alpine meadows, wetlands) in the Gran Paradiso National Park in August 2015. The surveys were carried out on sampling sites of 20mx20m divided into 2mx2m subplots. A moving windows approach was used to simulate greater patch sizes and calculate diversity indices (Species Richness, Shannon Index, Pielou's Evenness) for larger scales. Based on this, the information entropy was calculated.

Species richness and the Shannon index increased nonlinearly with increasing sampling size whereas the evenness declined. Thresholds in the behaviour of the indices to increasing sampling size indicated that it does not matter for the choice of the patch size if species richness is to be used alone or if species abundance is included.

Entropy as a measure of information content of the diversity declined with increasing grain size and was shown to be a robust method. It is concluded that using smaller grains of vegetation surveys, thus a higher resolution, is less prone to error. However, thresholds in the behaviour of information entropy indicated that using resolutions of 12mx12m to 14mx14m in remote sensing are sufficient to answer ecological research questions in homogenous grasslands.

### **Mediterranean high mountain flora: Patterns of endemism, species richness and leaf colours along an elevational gradient in the Spanish Sierra Nevada (Pia Eibes, David Kienle, Carl Beierkuhnlein)**

Mountains are important hotspots of biodiversity and often contain rare, endemic or threatened species. Therefore, elevational gradients pose ideal conditions to study the distribution and characteristics of species richness, endemism and adaptive plant traits. The following study aims to test if commonly reported patterns of decreasing species richness and increasing percentage of endemism also apply for the flora of Mediterranean high mountains. According to their biogeographic origins, endemic plant species of the study area can be divided into different endemic classes, respectively. This enables a detailed analysis of immigration paths and main drivers of endemic species with differently narrow distribution areas. A further analysis on how leaf colours of these species change with elevation allows an insight into adaptive strategies of high mountain plant species. Finally, I investigated to what extent vegetation patterns and colour information recorded during fieldwork correspond to vegetation indices derived from remote sensing methods. The study was conducted in the Sierra Nevada National Park in southern Spain, which contains outstanding numbers of endemic plant species. The study transect was placed between 2,000 and 3,470 m a.s.l. on a southern slope towards the Muhlacén, the highest summit in the Iberian Peninsula. Within twenty sampling sites along approximately 1,500 meters of elevation, all present vascular plant species have been recorded. Furthermore, total vegetation cover, and proportions of soil particles were estimated. Leaf colours of the main species have been measured using the Munsell colour space for plant tissues. Subsequently, all endemic species have been categorized into four different endemic classes in accordance with their biogeographic origins. Different vegetation indices and RGB colours were calculated



from a Sentinel satellite image, which was taken within the sampling period. Linear models were performed to test how vegetation patterns and leaf colours are related to elevation and to the patterns derived from remote sensing methods, respectively. While species richness monotonically decreased with elevation, general percentage of endemic species significantly increased. Different endemic classes showed variable patterns. The lightness of leaf colours decreased, while blue and green proportions increased with elevation. Some vegetation indices correlated with the observed vegetation patterns and colour information was partly reflected. When analysing total numbers, findings of species richness and percentage of endemism correspond to commonly reported patterns with a general decrease for species richness and a proportional increase of endemism at high-elevation sites. When focusing on individual endemic classes, the relationship with elevation differs strongly. This highlights the importance of detailed analyses of endemic plant species to better understand main drivers of endemism. The fact that leaf colours change with elevation indicates that high mountain plants have characteristic colour adaptations. Remote sensing products might constitute an additional tool to predict floral biodiversity of mountain ranges, but reach their limits at sparsely vegetated sites around summits.

### **Geographical analysis of habitat richness in European protected areas for conservation of biodiversity (Bravedo Mwaanga, Samuel Hoffmann, Carl Beierkuhnlein)**

Habitats symbolize a most direct link over spatial and geographical scales for conservation of biodiversity with reasons spanning from environmental, to economic and political. Yet, loss and fragmentation of habitats still remain the biggest threat to biodiversity. Linking protected areas (PA), ensuring their performance and the adequate protection is exclusively important in lieu of increasing constrained efforts at country level for PAs' maintenance of biodiversity, hence requiring a transnational perspective. I analysed habitat richness in protected areas strictly National Parks and Man and Biosphere reserves, across Europe with regard to the influence of spatial and geographical variables. Protected areas, combined with latitudinal and topographical orientation metrics of PAs were used for this study. Raster based grid analyses of habitats in PAs were used to generate habitat richness estimates using chain probability followed by correlation analysis to understand the contributions of the geographical variables. Further, relationship between habitat richness and area, latitudinal and altitudinal gradients is modelled and their individual relative importance was generated using variance partitioning.

Across Europe, habitat diversity increased with increasing size of a protected area ( $R^2 = 0.2541$ ;  $P = 2.2e-16$ ), altitudinal range ( $R^2 = 0.328$ ,  $P = 2.2e-16$ ), and increased topographic complexity ( $R^2 = 0.05464$ ,  $P = 3.888e-07$ ) but decreased with an increase in latitude ( $R^2 = 0.01219257$ ;  $P = 0.0207$ ). Altitudinal range had the strongest influence on habitat richness with a linear model but when log-transformed the size of a PA had a higher influence ( $R^2 = 0.3964840$ ;  $P = 2e-16$ ).

Under increased habitat loss and fragmentation and expected climate change, understanding the performance of PAs with regard to spatial factors is crucial for adequate planning, management and protection strategies in biodiversity conservation. Further, the study provides a guide and update on requirements of current protected areas and criteria for future designation of protected areas.

### **Finding the optimal spatial resolution for classification of habitat types in protected areas (Alexander Obermeier, Marco Heurich, Carl Beierkuhnlein)**



Selection of the right scales is important for any scientific study which deals with a spatial component. Especially in remote sensing and landscape ecology, spatial scales have a crucial effect on the results of a study. But how do scientists choose the right spatial resolution for their research? A few decades ago, the semi-variogram gained popularity to identify the optimal spatial resolution (OSR) for accurately mapping land cover and habitat types. This approach is based on auto-correlation of spectral pixel values which returns the distance at which pixels are no longer correlated. Being a measure of spatial heterogeneity, the output of the semi-variogram is affected by changes of habitat type composition and distribution.

The study intends to reveal the effect of management measures on OSR in the Bavarian Forest National Park. Performing an ANOVA analysis reveals differences or similarities of OSR between managed and unmanaged areas. It is the first study ever to investigate how OSR of spectral bands of RapidEye and NDVI are related to certain landscape heterogeneity features. For this purpose, landscape metrics are used as a proxy for certain features of landscape heterogeneity. Regression analysis serves to detect the strength and significance of the relationship of OSR and landscape metrics (a proxy for specific characteristics of landscape heterogeneity). Selected landscape metrics which are calculated in FRAGSTATS are Largest Patch Index (LPI), Area-Weighted Mean Shape Index (AWMSI), Area-Weighted Mean Fractal Dimension Index (AWMPFD), Number of Patches (NP), Contagion (CONTAG), Landscape Shape Index (LSI), Patch Richness (PR) and Shannon's Diversity Index (SHDI).

ANOVA analysis revealed similarity of OSR-results in the visible band spectrum of RapidEye, whereas the near-infrared band exhibited highly significant differences of OSR between managed and unmanaged areas of the national park. NDVI is dependent on the use of a predictor variable in the semi-variogram and only showed significant differences depending on management measure with altitude as predictor variable. Results suggest that differences between habitat types cannot be distinguished in the visible spectrum. But the semi-variogram of the near-infrared spectrum detects differences between forested habitat types. Hence, the different OSR output between managed and unmanaged areas, because unmanaged areas exhibit a higher diversity of forested habitat type patches. Differences in OSR of NDVI can be explained by both climatic and human influences. Linear regression analysis revealed diverging results depending on landscape metric and RapidEye band from which OSR was calculated. Results showed a heterogeneous behaviour depending on location in the national park. Weak, but significant relationships for the southern part could be found. Strongest and most consistent correlations (adjusted  $r^2 \sim 0.3$ ) were found for AWMPFD, PR and SHDI. Results of AWMPFD indicated that increasing complexity and irregularity of patch shapes caused finer OSR. PR and SHDI measure diversity of habitat types and need a coarser OSR when habitat types and their distribution among patch types becomes more equitable.

Results suggested that heterogeneity measures like PARA\_MN, NP or LSI are not related to OSR. Furthermore, expected relationships of selected landscape metrics and OSR were much smaller in the northern part compared to the southern part. The difference is most likely caused by a mismatch between mapped habitat types for calculation of landscape metrics and real detectable spectral differences between habitat types in the semi-variogram.

### **Assessing the forest response along treelines to an *Epirrita autumnata* outbreak in Abisko, using a combination of fieldwork and remote sensing (Frank Weiser, Carl Beierkuhnlein)**

Outbreaks of the moth species *Epirrita autumnata* (i.e. herbivorous caterpillars of this species) have increased in frequency all over Northern Scandinavia due to climate change. The outbreaks are able to





defoliate and severely damage the *Betula pubescens* forests of the area. While the birch forests profit from climate change and are able to both move upward and laterally increase their extent, the pressure by more frequent outbreak damage might hinder this increase or even push back tree lines. Most research regarding the outbreaks is based on very time consuming field methods, remote sensing techniques are severely underused. Therefore, one aim of this study was to assess the potential of RapidEye data due to its 5x5m spatial resolution and its red edge band, which is very sensitive to changes of the vegetation, for detecting tree line shifts due to *E. autumnata* damage.

The expectations were not completely fulfilled. Supervised classifications proved difficult due to low canopy cover of some forests and the gradual shift from forest to bushland that is characteristic for the tundra / taiga ecotone. However, a method based on the standard deviation of the Normalized Difference Red Edge Index (NDRE) of an image stack showed better results at detecting forest borders but still methodological challenges remain. Field work data collected in order to understand the influence of forest structure on the susceptibility to outbreak damage found elevation of the tree line to be the most important predictor for the 2012/2013 outbreak damage. The data showed that low elevation tree lines suffered more damage during the outbreak. This suggests that not all tree lines are influenced to the same extent by *Geometridae* feeding impact. Furthermore, this implies that low elevation forest shifts in order to adapt to climate warming could be hindered more than high elevation shifts.

#### **Identifying drivers of the treeline ecotone in the Alps – a remote sensing and GIS approach (Bernadette Menzinger, David Kienle, Carl Beierkuhnlein)**

The tree line ecotone in the European Alps is the transition zone above the subalpine forest and beneath the alpine zone. As the growing season in the alpine region is limited due to the decreasing temperature with altitude, it determines the establishment of trees. Various authors state, that changes in the transition zone are induced by topography and climatic variables at micro and macro scale. The topographic features determine the distribution of snow and the exposure of plants to wind and sun.

Through an analysis in GIS, I aim to capture differences in altitude of the upper tree limit in the Alps based on temperature and the actual limit derived from remote sensing data. Those differences will be analysed regarding topographic and climatic variables. Especially snow cover and depth are parameters, which are highly dependent of topography and wind, and delimit the establishment of trees. The altitude of the tree line ecotone will be derived in two ways, firstly after the temperature isotherms and secondly from a remote sensing product. In quantifying the differences of altitude of the two products, I will also include parameters like slope, aspect, wind and precipitation. Furthermore, an analysis of the trend of snow cover duration and depth will be included.

The study areas are three National Parks in the Alps, which are involved in the project of ECO-POTENTIAL. The Gran Paradiso National Park in Italy and the Swiss National Park in Switzerland, are two different protected areas considering, temperature, precipitation and topography.

#### **Hydrochemical parameters and Vegetation of helocrenic springs in the Bavarian Forest National Park (Jamyra Gehler, David Kienle, Marco Heurich, Carl Beierkuhnlein)**



Springs are unique ecosystems, which are barely represented in literature. They harbour a lot of different species and can be clearly distinguished to their environment due to their typical shape. The target of this study is helocrenic springs, so springs with a low amount of water sipping out. The water of these ecosystems is mainly transported in late Pleistocene solifluction layers parallel or at least close to the soil surface. In consequence, small catchments are characteristic for these mountain ecosystems including strong precipitation-dependency compared to other types of springs and short distances of water transport. Abiotic factors are +/- constant which delivers an adequate environment for stenoecious and frost as well as heat sensitive species. In this study 65 helocrenic springs in the Bavarian Forest were observed. Abiotic factors as well as species richness and coverage were recorded. It could be found, that pH-value decreases with height. In addition, alpha diversity of herbaceous species exhibits a reverse u-shaped relationship to light influx and does not correlate with the electric conductivity (compound load) of the spring water. Comparing alpha diversity of mosses and herbaceous plants showed no diverging trend along the whole gradient of pH-values. In conclusion, these small catchments were found to be excellent indicator systems for the ecological conditions of forest ecosystems. In the future, such in-situ information can be linked with the state and development of mountain forests reflected in remote sensing data.

**A closer look from space: The potential of Sentinel-2 data for beta-diversity measurements on small scale in a tundra alpine ecosystem (Edvinas Rommel, Duccio Rocchini, Carl Beierkuhnlein)**

Spectral heterogeneity of remote sensing imagery can be used efficiently for biodiversity assessment and presents a valuable source of ecological information, especially when free available data is used. This approach, known as the spectral variation hypothesis (SVH) is based on environmental heterogeneity, which is both related to spectral variability and to biodiversity. Little attention has been paid so far to the question whether the SVH can be applied on small scale and on free available data.

The present study aims to fill this gap by firstly investigating the positive influence of topographic complexity on beta-diversity on community level scale in Hardangervidda National Park. Secondly the performance of spectral heterogeneity, derived from Sentinel-2 data, for beta-diversity prediction is quantified. A vegetation survey was conducted in eight 100 x 100 m sites, each containing 20 systematic randomly placed plots of 2 x 2 m. Four sites were located in high topographic complexity areas (HTC) and four in low topographic complexity (LTC). Jaccard and Bray-Curtis dissimilarity were calculated for each site and correlated with eight different spectral heterogeneity measurements all based on the mean Euclidean distance to the spectral centroid using different spatial resolution and band numbers.

High topographic complexity was related to high beta-diversity. The same pattern could be observed in spectral heterogeneity. Spectral heterogeneity performed better with frequency based dissimilarity than with abundance based dissimilarity, probably because the latter was biased due to vegetation patchiness. The best performing heterogeneity measurement was NDVI heterogeneity with a spatial resolution of 10 x 10 m. Information loss by increased spatial resolution could be partially compensated by increased spectral resolution. The results of this study show that free available remote sensing data can provide reasonable estimates of beta-diversity patterns on small scale. Moreover, as theoretically demonstrated, spectral heterogeneity could even be used to improve efficiency of field surveys as well as mitigating location mismatch between remote sensing and in-situ data.

**Plotsize-decay of beta diversity in an alpine tundra ecosystem (Viola Hippler, Samuel Hoffmann, Carl Beierkuhnlein)**

As a key-concept, beta diversity depicts spatial complexity in order to characterise ecosystems. Spatial structures depend on resolution and so does beta diversity. When working on coarse resolution, small scale heterogeneity may be overlooked. This study aims to answer the questions whether larger plots result in reduced beta diversity and whether this methodical error is intensified by small distances between plots. At the same time the possibly positive influence of weighting shared species is explored. Vegetation data was collected in Hardangervidda National Park in two sites with high and two with low topographic complexity, each of 100 x 100 m. Respectively, 20 plots of 2, 4 and 10 m side length were chosen following systematic random sampling design.

Pairwise Jaccard and simple matching index of different plot sizes are compared applying rank sum tests. Distance dependency is examined using variograms, maximum values and data subsets of different size, the later also with variations of Jaccard Index in order to study influence of weighting shared species. Indices are judged on their ability to express differences in species distribution and relief. Beta diversity decreases with increasing plot size. Jaccard is more sensitive to plot size than simple matching but is also more sensitive to relief. Differences between topographic complexity classes are depicted independent from plot size. Data subsets and under-weighting of shared species reveal dependency of distance. Beta diversity, despite the use of different plot sizes, can be compared in respect to its relative patterns, e.g. when relating field data to satellite imagery. Distance decay studies with small extent can weight shared species, for example in form of Soerensen Index, to reduce methodical errors due to small distance.



## 5 Summary and recommendations

The current work provides an overview of pre-existing data sets and products for ECOPOTENTIAL sites and storylines. By analysing existing information we identified data gaps and assessed the potential of data analyses across PAs' referring to standardized data categories (see chapter 2.2). We furthermore demonstrate how in-situ and remote sensing data can be linked to derive spatial information about plant community composition and habitat types within a UNESCO Man and Biosphere Reserve; such approaches are required for nature conservation management in general. In addition, field studies carried out within the project are targeted to fill these in-situ data gaps in selected protected areas as. All in-situ data collected during these investigations are shared and documented using DEIMS-SDR and are open for further use.

The integration of in-situ data into spatial models based on remote sensing data is necessary to improve ecosystem benefits from earth observation. However, other than the collection of most remote sensing data by the user, in-situ data collection is often time-consuming and costly. As a consequence, one has to be aware of pre-existing in-situ data, of in-situ data gaps and the requirements in terms of quality and fitness for use of in-situ data for spatial models that rest on remote sensing data. Based on the results of the conducted studies a number of general recommendations for the targeted collection of in-situ data can be formulated:

- **[Spatial Accuracy]** Keep the GPS-location errors of the measurements in the field using your field device as small as possible! In addition this information needs to be kept for later analysis. The spatial accuracy of GPS measurements can be improved by avoiding topographic and physical barriers between the GPS device and the satellites. In addition, spatial accuracy is also depending on the topography as on the northern hemisphere, sites on southern aspects are usually exposed to more satellites than on northern aspects. This is vice-versa for the southern hemisphere.
- **[Spatial Resolution]** Before planning and conducting the field sampling one needs to be aware of the spatial resolution (e.g. grain size) of the EO data used. If necessary the spatial resolution of the field sampling (e.g. plot size and distance between plots) needs to be match accordingly. Sensitivity analyses to data resolution may be, however, necessary to validate results. Is the spatial and temporal scale of your field and RS data able to account for the ecological issue in focus?
- **[Temporal accuracy and resolution]** Ensure that in-situ and EO data cover the same or similar time periods relevant for the study. This not only addresses the date and time of sampling but also the update frequency of the datasets. As appropriate RS data might not be available for the date of in-situ data acquisition, proxies e.g. for a similar period (e.g. phenological state) need to be taken into account. In addition, the spatial and temporal dynamics of your study system to match with the EO data needs to be considered.
- **[Thematic accuracy and resolution]** Define the entities used in the field sampling in order to match the categories used in EO data. The semantic aggregation of the recorded entities (e.g. vegetation types, land cover types, or land use categories) need to be matched to information which is intended to be integrated and used from EO data. E.g. land cover systems like LCCS or Eagle try to assess this aspect linking to general habitat categories (GHC) as defined in the EBONE project.



- **[Vertical accuracy]** Depending on the type of EO data used, avoid the sampling of shaded areas in the field. As optical sensors (e.g. Sentinel 2 10-20m & 3 300m) only scan the surface of the ecosystems other types like radar (e.g. Sentinel 1 10m) can penetrate the vegetation. This needs to be taken into account when recording in-situ characteristics of ecosystems in the field.
- **[Documentation]** Provide a full description of the methods applied as well as the reference lists used in your metadata. The provision of metadata need to follow commonly agreed standards in order to ensure the discoverability and re-usability of the data. Information on the above mentioned aspects need to be recorded not only for EO data but moreover for any in-situ observations taken in the field. A proper description of the methods applied and the semantics used is needed. If possible this information should be provided online in web accessible services, e.g. vocabulary service. This is among the most important aspects of the collected data. Associate the in-situ data with documentations.

Based on these recommendations a harmonised provision of data can be fostered. Still in addition syntactic and semantic interoperability (see Magagna et al. 2018<sup>4</sup>) needs to be considered. Deliverable 5.5 generally assesses the availability and usefulness of in-situ data for ecosystem models that incorporate spatial data from Earth observation. Moreover, Deliverable 5.5 makes more in-situ data available that will be used in future.

---

<sup>4</sup> See ECOPOTENTIAL deliverable D5.6 Harmonised delivery of data



## 6 References

- Auerbach, M., & Shmida, A. 1993. Vegetation Change Along an Altitudinal Gradient on Mt-Hermon, Israel - No Evidence for Discrete Communities. *Journal of Ecology* 81: 25–33.
- Baillarin, S.J., Meygret, A., Dechoz, C., Petrucci, B., Lacherade, S., Tremas, T., Isola, C., Martimort, P., & Spoto, F. 2012. Sentinel-2 level 1 products and image processing performances. Geoscience and Remote Sensing Symposium (IGARSS), 2012 IEEE International (pp. 7003-7006). IEEE.
- Beierkuhnlein, C. 2017. Inseln als globale Versuchsanordnung und natürliche Laboratorien der Vegetationsökologie. *Berichte der Reinhold-Tüxen-Gesellschaft*, 29.
- Bustamante J., Aragonés D., Afán I., Luque C.J., Pérez-Vázquez A., Castellanos E.M., & Díaz-Delgado R. 2016. Hyperspectral Sensors as a Management Tool to Prevent the Invasion of the Exotic Cordgrass *Spartina densiflora* in the Doñana Wetlands. *Remote Sensing*, 8(12), 1001.doi:10.3390/rs8121001
- Chiarucci, A. 2007. To sample or not to sample? That is the question. . . for the vegetation scientist. *Folia Geobotanica* 42: 209–216.
- Choudhury B.U., Ferraris S., Ashton R.W., Powlson D.S., & Whalley W.R. (2018). The effect of microbial activity on soil water diffusivity. *European Journal of Soil Science*, 2018, doi:10.1111/ejss.12535
- Chytrý, M., & Otýpková, Z. 2003. Plot sizes used for phytosociological sampling of European vegetation. *Journal of Vegetation Science* 14: 563-570.
- Clements, F.E. 1905. *Research Methods in Ecology*. Science 22: 368.
- Clements, F.E. 1916. *Plant succession: an analysis of the development of vegetation*. Carnegie Institute, Publication 242, Washington, DC, US.
- Copernicus Sentinel Data. 2017. European Space Agency.  
<https://scihub.copernicus.eu/dhus/#/home>, accessed on 6/20/2017.
- Del Arco Aguilar, M.J., González-González, R., Garzón-Machado, V., & Pizarro-Hernández, B. 2010. Actual and potential natural vegetation on the Canary Islands and its conservation status. *Biodiversity and Conservation* 19: 3089–3140.
- Díaz-Delgado R., Aragonés D., Afán I., & Bustamante J. 2016. Long-Term Monitoring of the Flooding Regime and Hydroperiod of Doñana Marshes with Landsat Time Series (1974–2014). *Remote Sensing*, 8(9), 775, doi:10.3390/rs8090775
- Filippini, F., Valentini, E., & Taramelli, A. 2017. Sea Surface Temperature changes analysis, an Essential Climate Variable for Ecosystem Services provisioning. In *Analysis of Multitemporal Remote Sensing Images (MultiTemp)*, 2017 9th International Workshop on the (pp. 1-8). IEEE.
- Frampton, W.J., Dash, J., Watmough, G., & Milton, E.J. 2013. Evaluating the capabilities of Sentinel-2 for quantitative estimation of biophysical variables in vegetation. *ISPRS journal of photogrammetry and remote sensing*, 82, 83-92.



- González Henríquez, M.N., Rodrigo Pérez, J.D. & Suárez Rodríguez, C. 1986. Flora y Vegetación del Archipelago Canario. Edirca, Las Palmas de Gran Canaria, ES.
- Hall, K., Reitalu, T., Sykes, M.T., & Prentice, H.C. 2012. Spectral heterogeneity of QuickBird satellite data is related to fine-scale plant species spatial turnover in semi-natural grasslands. *Applied Vegetation Science*, 15: 145–157.
- He, K., & Zhang, J. 2009. Testing the correlation between beta diversity and differences in productivity among global ecoregions, biomes, and biogeographical realms. *Ecological Informatics*, 4: 93–98.
- He, K.S., Zhang, J., & Zhang, Q. 2009. Linking variability in species composition and MODIS NDVI based on beta diversity measurements. *Acta Oecologica*, 35: 14–2.
- Hoffmann, S., Beierkuhnlein, C., Field, R., Provenzale, A., Chiarucci, A. (2018). Uniqueness of Protected Areas for Conservation Strategies in the European Union. *Scientific Reports*, doi:10.1038/s41598-018-24390-3
- Hoffmann, S., Schmitt, T., Chiarucci, A., Irl, S., Rocchini, D., Vetaas, O.R., Tanase, M.A., Mermoz, S., Bouvet, A., & Beierkuhnlein, C. 2018. Remote sensing of beta diversity: evidence from plant communities in a semi-natural system, *Applied Vegetation Science*, doi:10.1111/avsc.12403
- Horning, N., Robinson, J.A., Sterling, E.J., Turner, W., & Spector, S. 2010. *Remote Sensing for Ecology and Conservation. A Handbook of Techniques*.
- Hufkens, K., Scheunders, P., & Ceulemans, R. 2009. Ecotones in vegetation ecology: Methodologies and definitions revisited. *Ecological Research* 24: 977–986.
- Hummel, C., Provenzale, A., van der Meer, J., Wijnhoven, S., Nolte, A., Poursanidis, D., Janss, G., Jurek, M., Andresen, M., Poulin, B., & Kobler, J. 2017. Ecosystem services in European protected areas: Ambiguity in the views of scientists and managers?. *PloS one*, 12(11), e0187143.
- Irl, S. 2016. Plant diversity on high elevation islands – drivers of species richness and endemism. *Frontiers of Biogeography*, 8, e29717, doi:10.21425/F58329717
- Irl, S.D.H., Harter, D.E.V., Steinbauer, M.J., Gallego Puyol, D., Fernández-Palacios, J.M., Jentsch, A., & Beierkuhnlein, C. 2015. Climate vs. topography - spatial patterns of plant species diversity and endemism on a high-elevation island. *Journal of Ecology* 103: 1621–1633.
- Irl, S., Schweiger, A., Medina, F.M., Fernández-Palacios, J.M., Harter, D., Jentsch, A., Provenzale, A., Steinbauer, M.J., Beierkuhnlein, C. 2017: An island view of endemic rarity – environmental drivers and consequences for nature conservation. *Diversity and Distributions*, 1-11, doi:10.1111/ddi.12605
- Irl, S.D.H., Steinbauer, M.J., Babel, W., Beierkuhnlein, C., Blume-Werry, G., Messinger, J., Palomares Martínez, Á., Strohmeier, S., & Jentsch, A. 2012. An 11-yr exclosure experiment in a high-elevation island ecosystem: introduced herbivore impact on shrub species richness, seedling recruitment and population dynamics. *Journal of Vegetation Science* 23: 1114–1125.
- Jax, K. 2006. Ecological units: definitions and application. *The Quarterly Review of Biology* 81: 237–258.
- Lawton, J.H. 1999. Are there general laws in ecology? *Oikos* 84: 177–192.
- Legendre, P. 2014. Interpreting the replacement and richness difference components of beta diversity. *Global Ecology and Biogeography* 23: 1324–1334.



- Legendre, P., & De Cáceres, M. 2013. Beta diversity as the variance of community data: Dissimilarity coefficients and partitioning. *Ecology Letters* 16: 951–963.
- Levin, N., Shmida, A., Levanoni, O., Tamari, H., & Kark, S. 2007. Predicting mountain plant richness and rarity from space using satellite-derived vegetation indices. *Diversity and Distributions*, 13: 692–703.
- Livingston, B.E. 1903. The Distribution of the Upland Plant Societies of Kent County, Michigan. *Botanical Gazette* 35: 36–55.
- Lloyd, K., McQueen, A., & Lee, B. 2000. Evidence on ecotone concepts from switch, environmental and anthropogenic ecotones. *Journal of Vegetation Science* 11: 903–910.
- McCune, B. & Grace, J.B. 2002. *Analysis of Ecological Communities*. MjM Software Design, Gleneden Beach, OR, US.
- Muer, T., Sauerbier, H., & Calixto, C. 2016. *Die Farn- und Blütenpflanzen der Kanarischen Inseln*. Margraf Publishers, Weikersheim, Germany.
- Nagendra, H., & Rocchini, D. 2008. High resolution satellite imagery for tropical biodiversity studies: The devil is in the detail. *Biodiversity and Conservation*, 17: 3431–3442.
- Nagendra, H., Rocchini, D., Ghate, R., Sharma, B., & Pareeth, S. 2010. Assessing plant diversity in a dry tropical forest: Comparing the utility of Landsat and IKONOS satellite images. *Remote Sensing*, 2: 478–496.
- Ørka, H.O., Wulder, M.A., Gobakken, T., & Næsset, E. 2012. Subalpine zone delineation using LiDAR and Landsat imagery. *Remote Sensing of Environment* 119: 11–20.
- Otto, R., García-del-Rey, E., Méndez, J., & Fernández-Palacios, J.M. 2012. Effects of thinning on seed rain, regeneration and understory vegetation in a *Pinus canariensis* plantation (Tenerife, Canary Islands). *Forest Ecology and Management* 280: 71–81.
- Palmer, M.W., Earls, P.G., Hoagland, B.W., White, P.S., & Wohlgemuth, T. 2002. Quantitative tools for perfecting species lists. *Environmetrics* 13: 121–137.
- Palmer, M.W., & White, P.S. 1994. On the existence of ecological communities. *Journal of Vegetation Science* 5: 279–282.
- Palomares Martínez, Á., López Graciano, C., Freixes Montes, F., Gómez Gómez, M., Moral del Barrio, M.F., León Pérez, A.M., Lorenzo Pérez, J.H., Balsera Pinar, G., Méndez Carvajal, C., Acevedo Rodríguez, A., Ibáñez de Aldecoa, F.B., Rodríguez Lerin, A., and Fernandez San Martín, J.M. 2011. *Memoria anual del Parque Nacional de La Caldera de Taburiente de 2012*. Ministerio de Medio Ambiente, y Medio Rural, y Marino, Madrid.
- Pettorelli, N. 2013. *The Normalized Difference Vegetation Index*. Oxford University Press, Oxford, UK.
- Pettorelli, N., Laurance, W.F., O'Brien, T.G., Wegmann, M., Nagendra, H., & Turner, W. 2014. Satellite remote sensing for applied ecologists: Opportunities and challenges. *Journal of Applied Ecology* 51: 839–848.
- Raffelli, G., Previati, M., Canone, D., Gisolo, D., Bevilacqua, I., Capello, G., Biddoccu, M., Cavallo, E., Deiana, R., Cassiani, G., & Ferraris, S. 2017. Local-and Plot-Scale Measurements of Soil Moisture: Time and





- Spatially Resolved Field Techniques in Plain, Hill and Mountain Sites. *Water*, 9(9), 706, doi:10.3390/w9090706
- Rees, W.G. 2007. Characterisation of Arctic treelines by LiDAR and multispectral imagery. *Polar Record* 43: 345–352.
- Ricklefs, R. 1987. Community Diversity: Relative Roles of Local and Regional Processes. *Science* 235: 167–171.
- Ricklefs, R.E. 2008. Disintegration of the Ecological Community American Society of Naturalists Sewall Wright Award Winner Address. *The American Naturalist* 172: 741–750.
- Rocchini, D. 2007. Effects of spatial and spectral resolution in estimating ecosystem  $\alpha$ -diversity by satellite imagery. *Remote Sensing of Environment*, 111: 423–434.
- Rocchini, D., Balkenhol, N., Carter, G.A., Foody, G.M., Gillespie, T.W., He, K.S., Kark, S., Levin, N., Lucas, K., Luoto, M., & Nagendra, H. 2010. Remotely sensed spectral heterogeneity as a proxy of species diversity: recent advances and open challenges. *Ecological Informatics*, 5: 318-329.
- Rocchini, D., Boyd, D.S., Féret, J.-B., Foody, G.M., He, K.S., Lausch, A., Nagendra, H., Wegmann, M., & Pettorelli, N. 2016. Satellite remote sensing to monitor species diversity: potential and pitfalls. *Remote Sensing in Ecology and Conservation* 2: 25–36.
- Rocchini, D., Butini, S.A., & Chiarucci, A. 2005. Maximizing plant species inventory efficiency by means of remotely sensed spectral distances. *Global Ecology and Biogeography* 14: 431–437.
- Rocchini D., Chiarucci A., Loiselle S.A. 2004. Testing the spectral variation hypothesis by using satellite multispectral images. *Acta Oecologica* 26: 117–120.
- Schmidtlein, S., & Fassnacht, F. E. 2017. The spectral variability hypothesis does not hold across landscapes. *Remote Sensing of Environment* 192: 114–125.
- Sha, Z., Bai, Y., Xie, Y., Yu, M., & Zhang, L. 2008. Using a hybrid fuzzy classifier (HFC) to map typical grassland vegetation in Xilin River Basin, Inner Mongolia, China. *International Journal of Remote Sensing* 29: 2317–2337.
- Sims, D.A., & Gamon, J.A. 2002. Relationships between leaf pigment content and spectral reflectance across a wide range of species, leaf structures and developmental stages. *Remote Sensing of Environment* 81: 337–354.
- Simpson, E. 1949. Measurement of diversity. *Nature* 163: 688
- Socolar, J. B., Gilroy, J. J., Kunin, W. E., & Edwards, D. P. 2016. How should beta-diversity inform biodiversity conservation?. *Trends in ecology & evolution* 31: 67–80.
- TERN (2018). Effective Field Calibration and Validation Practices. A practical handbook for calibration and validation of satellite and model-derived terrestrial environmental variables for research and management. A TERN Landscape Assessment Initiative. Version 1.3. Available at <http://www.tern.org.au/NEW-CalVal-handbook-for-remote-sensing-bgp4370.html>
- Tomaselli, V., Adamo, M., Veronico, G., Sciandrello, S., Tarantino, C., Dimopoulos, P., Medagli, P., Nagendra, H. & Blonda, P. 2017. Definition and application of expert knowledge on vegetation pattern,



- phenology, and seasonality for habitat mapping, as exemplified in a Mediterranean coastal site. *Plant Biosystems-An International Journal Dealing with all Aspects of Plant Biology*, 151(5), 887-899.
- Turco, M., Bedia, J., Di Liberto, F., Fiorucci, P., von Hardenberg, J., Koutsias, N., Llasat, M.-C., Xystrakis, F., & Provenzale, A. 2016. Decreasing Fires in Mediterranean Europe. *PLoS ONE* 11(3), e0150663, doi:10.1371/journal.pone.0150663
- Valentini, E., Filipponi, F., Nguyen Xuan, A., Passarelli, F.M., & Taramelli A. 2016. Earth Observation for Maritime Spatial Planning: Measuring, Observing and Modeling Marine Environment to Assess Potential Aquaculture Sites. *Sustainability*, 8(6), 519, doi:10.3390/su8060519
- Vellend, M., Verheyen, K., Flinn, K.M., Jacquemyn, H., Kolb, A., Van Calster, H., Peterken, G., Graae, B.J., Bellemare, J., Honnay, O., Brunet, J., Wulf, M., Gerhardt, F., & Hermy, M. 2007. Homogenization of forest plant communities and weakening of species-environment relationships via agricultural land use. *Journal of Ecology* 95: 565–573.
- Vellend, M. 2010. Conceptual Synthesis in Community Ecology. *The Quarterly Review of Biology* 85: 183–206.
- Whittaker, R.H. 1970. *Communities and ecosystems*. MacMillan, London, UK.
- Whittaker, R.H. 1967. Gradient analysis of vegetation. *Biological Reviews of the Cambridge Philosophical Society* 42: 207–264.
- Whittaker, R.H. 1956. Vegetation of the Great Smoky Mountains. *Ecological Monographs* 26: 1–80.
- Wiens, J.A., Crawford, C.S., Gosz, J.R., Crawford, S., & Boundary, J.R. 1985. Boundary dynamics: a conceptual framework for studying landscape ecosystems. *Oikos* 45: 421–427.
- Wilson, J.B., & Agnew, A.D.Q. 1992. Positive-feedback Switches in Plant Communities. *Advances in Ecological Research* 23: 263–336.
- Wood, S.N. 2017. *Generalized additive models: an introduction with R*. Chapman and Hall/CRC press, Boca Raton, FL, US.
- Xie, Y., Sha, Z., & Yu, M. 2008. Remote sensing imagery in vegetation mapping: a review. *Journal of Plant Ecology* 1: 9–23.



## 7 Appendix

Table 3. Overview on in-situ datasets and data products included in DEIMS-SDR and scientific publications related to ECOPOTENTIAL protected areas.

Protected Area	Category	Characteristic		Domain		Data type		
		Abiotic	Biotic	Marine	Terrestrial	In-situ	EO	Model
La Palma Island	Flora		x		x	x		
La Palma Island	Habitats		x		x	x	x	x
La Palma Island	Land Cover	x	x		x	x	x	x
La Palma Island	Flora		x		x	x		x
La Palma Island	Climate	x			x	x		x
La Palma Island	Geology	x			x	x		
La Palma Island	Hydrology	x		x	x	x		
La Palma Island	Topography	x			x	x		x
La Palma Island	Flora		x		x	x		x
Hardangervidda	Flora		x		x	x		
Hardangervidda	Fauna		x		x	x		
Hardangervidda	Topography	x			x	x	x	x
Hardangervidda	Imagery	x			x		x	
Hardangervidda	Climate	x			x	x		x
Hardangervidda	Climate	x			x		x	x
Peneda-Geres	Climate	x			x		x	x
Peneda-Geres	Hydrology	x			x	x	x	x
Peneda-Geres	Land Cover	x	x		x			
Peneda-Geres	Habitats		x		x	x		
Peneda-Geres	Flora		x		x	x		
Peneda-Geres	Flora		x		x		x	x
Peneda-Geres	Fauna		x		x	x		
Kruger	Flora		x		x	x		
Kruger	Fauna		x		x	x		
Kruger	Flora		x		x	x		
Kruger	Flora		x		x	x		
Kruger	Flora		x		x	x		
Kruger	Climate	x					x	x
Kruger	Flora		x		x	x		
Abisko	Flora		x		x	x		
Tatra	Fauna		x		x	x		
Tatra	Fauna		x		x	x		
Tatra	Fauna		x		x	x		
Tatra	Fauna		x		x	x		



Protected Area	Category	Characteristic		Domain		Data type		
		Abiotic	Biotic	Marine	Terrestrial	In-situ	EO	Model
Tatra	Habitats		x		x	x		
Tatra	Fauna		x		x	x		
Tatra	Fauna		x		x	x		
Samaria	Flora		x		x	x		
Samaria	Habitats		x		x	x		
Samaria	Fauna		x		x	x		
Samaria	Fauna		x		x	x		
Samaria	Fauna		x	x		x		
Curonian Spit	Flora		x		x	x		x
Curonian Spit	Chemistry	x			x	x		
Curonian Spit	Climate	x					x	x
Sierra Nevada	Flora		x		x	x		
Sierra Nevada	Agriculture		x		x	x		
Sierra Nevada	Imagery	x			x		x	
Sierra Nevada	Land Cover	x	x		x	x	x	x
Sierra Nevada	Land Cover	x	x		x	x	x	x
Sierra Nevada	Climate	x			x	x		x
Sierra Nevada	Climate	x					x	x
Sierra Nevada	Hydrology	x		x	x	x		
Danube Delta	Fauna		x		x	x		
Danube Delta	Hydrology	x		x	x	x		
Montado	Flora		x		x	x		
Montado	Fauna		x		x	x		
Montado	Climate	x			x	x		x
Montado	Flora		x		x	x		
Montado	Imagery	x			x		x	
Montado	Fire	x			x	x		
Montado	Land Cover	x	x		x	x		
Montado	Soil	x	x		x	x		
Montado	Flora		x		x	x		
Montado	Flora		x		x	x	x	x
Montado	Flora		x		x	x		
Montado	Flora		x		x	x	x	x
Montado	Flora		x		x	x	x	x
Donana	Fauna		x	x	x	x		
Donana	Fauna		x	x	x	x		
Donana	Flora		x	x	x	x		
Donana	Flora		x	x	x	x		
Donana	Fauna		x	x	x	x		
Donana	Fauna		x	x	x	x		
Donana	Fauna		x	x	x	x		
Donana	Fauna		x	x	x	x		
Donana	Flora		x	x	x	x	x	x



Protected Area	Category	Characteristic		Domain		Data type		
		Abiotic	Biotic	Marine	Terrestrial	In-situ	EO	Model
Donana	Flora		x	x	x	x	x	x
Donana	Climate	x			x	x		
Donana	Habitats		x	x	x	x		
Donana	Chemistry	x		x		x		
Donana	Land Cover	x	x	x	x	x		
Donana	Flora		x		x	x		
Donana	Flora		x	x	x	x		
Donana	Climate	x			x	x		
Donana	Climate	x					x	x
Donana	Hydrology	x		x	x	x		
Donana	Chemistry	x		x	x	x		
Camargue	Climate	x					x	x
Wadden Sea	Climate	x					x	x
Gran Paradiso	Climate	x					x	x
Gran Paradiso	Fauna		x		x	x		
Negev	Climate	x					x	x
Ohrid	Climate	x					x	x
Kalkalpen	Climate	x					x	x
Mediterranean	Climate	x		x			x	x
La Réunion Island	Hydrology	x	x	x	x		x	
Swiss National Park	Flora		x		x	x	x	x
Le Cesine (coastal wetland)	Habitats		x		x	x	x	
Gran Paradiso	Hydrology	x			x	x		
Delta Po	Hydrology	x		x		x		
Several	Fire	x	x		x	x		
Donana	Hydrology	x				x	x	
Several	Soil	x	x		x			x
Several	Ecosystem Services, Pressures and Threats	x	x	x	x	x		
Bavarian Forest	Chemistry	x	x		x	x		
Bavarian Forest	Flora	x	x		x	x		
Several	Climate			x		x	x	x
Several	Hydrology			x		x	x	x
Several	Fauna			x		x	x	x
Several	Imagery	x	x		x	x	x	x



Table 4. Available in-situ data sets and products described by the ECO POTENTIAL storylines<sup>5</sup> for arid/semi-arid ecosystems (A), coastal/marine ecosystems (C) and mountain ecosystems (M).

Storyline	Available data
A1	Rainfall
A1	Runoff
A1	Floods
A1	Temperature
A1	Evaporation
A1	Soil moisture
A1	Geodiversity
A1	Geology
A1	Geo-morphology
A1	Topography
A1	Primary production
A1	Pistachio
A1	Isopods
A1	Snails
A1	Porcupine
A1	Beetles
A1	Population of ibex, Dorcas gazelle, vultures sand grouse
A1	Settlements size, population and socioeconomic data
A1	Roads
A1	Land use
A1	Field vehicles and tourists
A2	Herb. Biomass (g/m <sup>2</sup> )
A2	Leaf nitrogen (%)
A2	Tree cover (%)
A2	Tree biomass (tons/ha)
A2	NDVI/LAI
A2	Water and carbon fluxes
A3	Vegetation
A3	Fauna
A3	Hydrology
A3	Habitat
A3	Climate
A3	Tourism
A3	Land use
A3	Soil
A3	Lithology
A3	Geology
A3	Waterways
A3	Infrastructure
A3	Protected area
A4	Precipitation
A4	Soil variables
A4	Vegetation structure, tree metrics, tree health condition, Year of cork extraction, etc.

<sup>5</sup> The list of storylines is provided at <http://www.ecopotential-project.eu/site-studies/storylines.html>



Storyline	Available data
A4	SOM
A4	Birds species monitoring (P/A data)
A4	Habitat structure
A4	Burned areas
M1	Meteorological parameters
M1	Wild ungulates count data and population structure
M1	Hunting data (reindeer, HNP)
M1	Species occurrences (HNP)
M1	Arthropod and bird diversity (GPNP)
M1	Vegetation survey
M1	Green index (GCC - green chromatic coordinate)
M1	Grass samples inside and outside enclosures
M2	Climate
M2	Soil inventory
M2	Forest structure (Tree species, cover, height, dbh, etc.)
M2	Bird species occurrence and diversity
M2	Plant species occurrence and diversity
M2	Beetle species occurrence and diversity
M2	Butterfly species occurrence and diversity
M4	Climate
M4	High resolution photos by drones
M4	Vegetation data
M4	Vertebrates and invertebrates
M5	Climate
M5	Physico-chemical water parameters and nutrients (A).
M5	Phytoplankton and zooplankton abundance and size distribution (A)
M5	Surveys on invertebrate and amphibian presence, fish presence (A)
M5	Hydrological balance (A).
M5	Carbon Fluxes (P)
M5	Macrophytes and phytoplankton dynamics, zooplankton dynamics
M5	Taxonomic data-time series of species richness
M5	Fish abundance, Presence of invasive species
M6	Biodiversity
M6	Land use
M6	Biomass
M6	Climate
M7	Climate
M7	Geology
M7	Soils
M7	DEM
M7	Hydrography
M7	Socio-economy
M7	LC/LU and habitats
M7	Vegetation properties
M7	Wildfires
M7	Invasive species
M7	Biodiversity records
M7	Governance
M9	Climate



Storyline	Available data
M9	Aerial images
M9	Historical documents
M9	Biodiversity
M9	Grazing
M9	Irrigation
O3a	Meteorological conditions
O3a	Surface or groundwater levels in marshes
O3a	Surface water levels and salinity
O3a	Rhône river flow
O3a	Relative sea level
O3a	Total rice area
O3a	Salinity of Vaccarès lagoon
O3a	Wintering ducks
O3a	Nesting colonial shorebirds
O3a	Nesting colonial herons in reeds and trees
O3a	Nesting flamingos (adults, nests & young)
O3a	Wintering birds
O3a	Fish stock
O3a	Rabbit populations
O3a	Herd size of Camargue bulls, horses, sheeps
O3a	Number of Camargue horses
O3b	Precipitation
O3b	Groundwater level
O3b	Water turbidity, Chlorophyll concentration
O3b	Water depth
O3b	Nutrient concentration in water
O3b	Waterbirds
O3b	Colonial waterbirds
O3b	Non-colonial waterbirds
O3b	Wild ungulates
O3b	Domestic livestock
O4	Suspended Solids
O4	water turbidity
O4	water depth
O4	River water level
O4	Bird census
O4	Nutrients
O4	alkalinity
O4	TOC (total organic carbon)
O4	water temperature
O4	pH
O4	conductivity
O4	chlorophyll
O4	Secchi disc depth
O4	Number of tourist
O4	Fish catches reported
O6	Biodiversity
O7	Endemic vegetation
O7	Climate





Storyline	Available data
O7	Biodiversity