



Sample-data publishing with the IPT Tool 2.3

*Larissa Smirnova, Patricia Mergen
Royal Museum for Central Africa, Botanic Garden Meise*

ASSUMPTIONS :

- There is a demand from content providers
- Some datasets already published in GBIF would better fit sample-based data concept
- New IPT functionalities allow publishing of a larger variety of datasets
- New type of data should extend range of data providers
- Technical documentation and support are important prerequisites for successful implementation of this new development

SAMPLE-BASED DATA:

Sample-based data is....

“a type of data available from environmental, ecological, and natural resource investigations: one-off studies or monitoring programs. Such data are usually quantitative, calibrated, and follow certain protocols so that changes and trends of populations can be detected.”

EXAMPLES OF SAMPLE-BASED DATA

1. An insect trap set in the forest canopy and picked up one month afterwards
2. A GPS collar attached to an individual constantly sending data to a central database.
3. Satellite imagery downloaded every year to analyze the vegetation composition of an area
4. Car-based bat monitoring where volunteers are recording the ultrasound using time expansion detectors
5. Vegetation sampling (core boring) for screening of subsurface pollution

EU BON, GBIF AND EBVs

Enabling discovery and access to sample-based data

- ✓ GBIF in collaboration with EU BON partners, has adapted the Integrated Publishing Toolkit and Darwin Core Archives to enable the flow of sample-based data (2015)
- ✓ One of the objectives is to support GEO BON's Essential Biodiversity Variables (EBVs)
- ✓ Support detection of changes and trends in populations

EBV CLASSES

EBV Class	Genetic composition	Species populations	Species traits	Community composition	Ecosystem structure	Ecosystem function
EBV example	Allelic diversity	Abundances and distributions	Phenology	Taxonomic diversity	Habitat structure	Nutrient retention

Class	EBV	Definition	How to measure in marine, terrestrial, freshwater (spatial, temporal, taxonomic)
Species populations	Species occurrence	Presence/absence of a given taxon or functional group at a given location	Quantify number/biomass/cover at a sample of selected taxa (or functional gps) at extensive suite of sites (selected from stratified random sample or building on existing networks)
	Population abundance	Quantity of individuals or biomass of a given taxon or functional group at a given location	
	Population structure by age/size class	Quantity of individuals or biomass of a given demographic class of a given taxon or functional group at a given location	

APPLICATIONS OF S-B DATA (1/3)

Goals: Modeling supporting conservation or wildlife management

- prediction of distribution
- identification of important areas (e.g., those areas more suitable for a species or more likely to hold high densities)

Methods:

- camera traps
- gps tracking
- point-count surveys

Measurements:

- abundance
- presence / absence

APPLICATIONS OF S-B DATA (2/3)

Goals: Modeling supporting biodiversity conservation, environmental monitoring, ecosystem health and condition and impact on climate change

- vegetation research (phyto sociology, trend analysis)
- forest degradation
- land use
- carbon monitoring
- ecological dynamics

Methods:

- vegetation plot
- relevé

Measurements:

- species abundance
- cover (e.g. Braun-Blanquet scale)

APPLICATIONS OF S-B DATA (3/3)

Goals: Interactions between biodiversity and environment

- species monitoring
- temporal patterns
- population trends and responses to drivers
- biological indicators of land use change

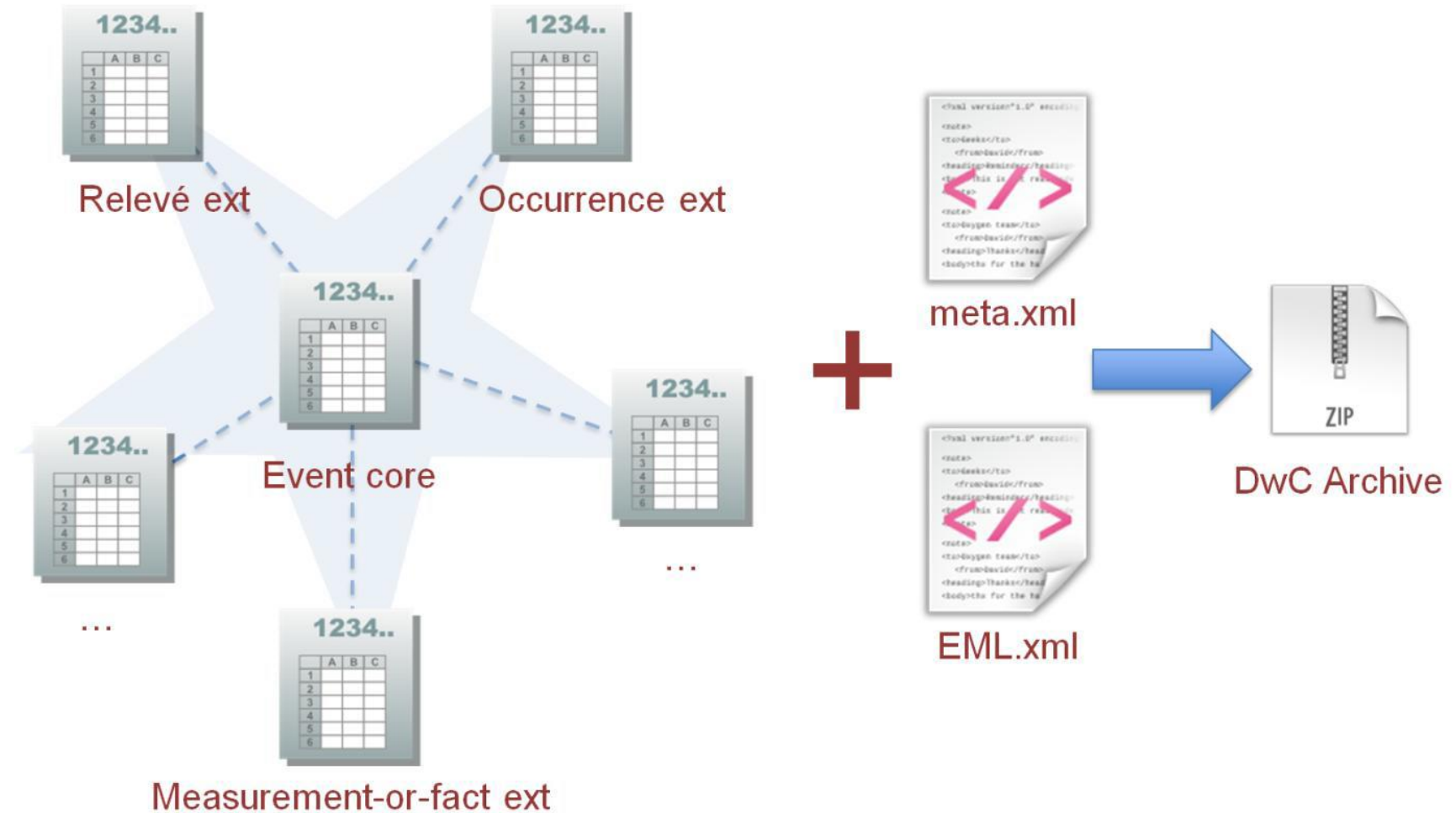
Methods:

- long-term monitoring data (census)
- atlas data (survey)

Measurements:

- number of individuals
- relative abundances

ENABLING SD PUBLISHING: STAR SCHEMA

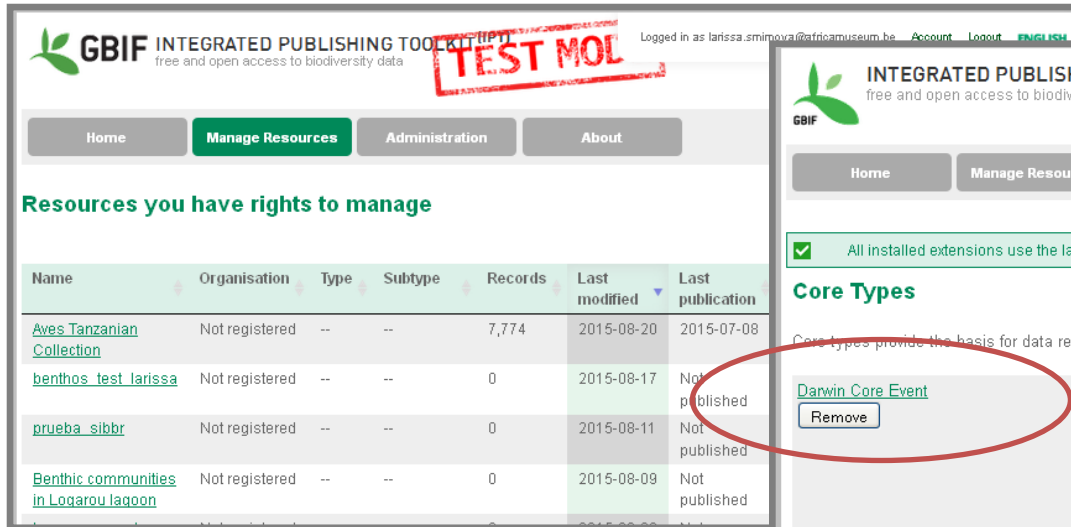


ENABLING SD PUBLISHING : DWC VOCABULARY

7 essential terms for encoding sample data:

1. eventID
2. parentEventID (new)
3. samplingProtocol
4. sampleSize (new)
5. sampleSizeUnit (new)
6. organismQuantity (new)
7. organismQuantityType (new)

ENABLING SD PUBLISHING: IPT v.2.3



GBIF INTEGRATED PUBLISHING TOOLKIT (IPT) v.2.3
free and open access to biodiversity data

Logged in as larissa.sminova@afriamuseum.be Account Logout ENGLISH

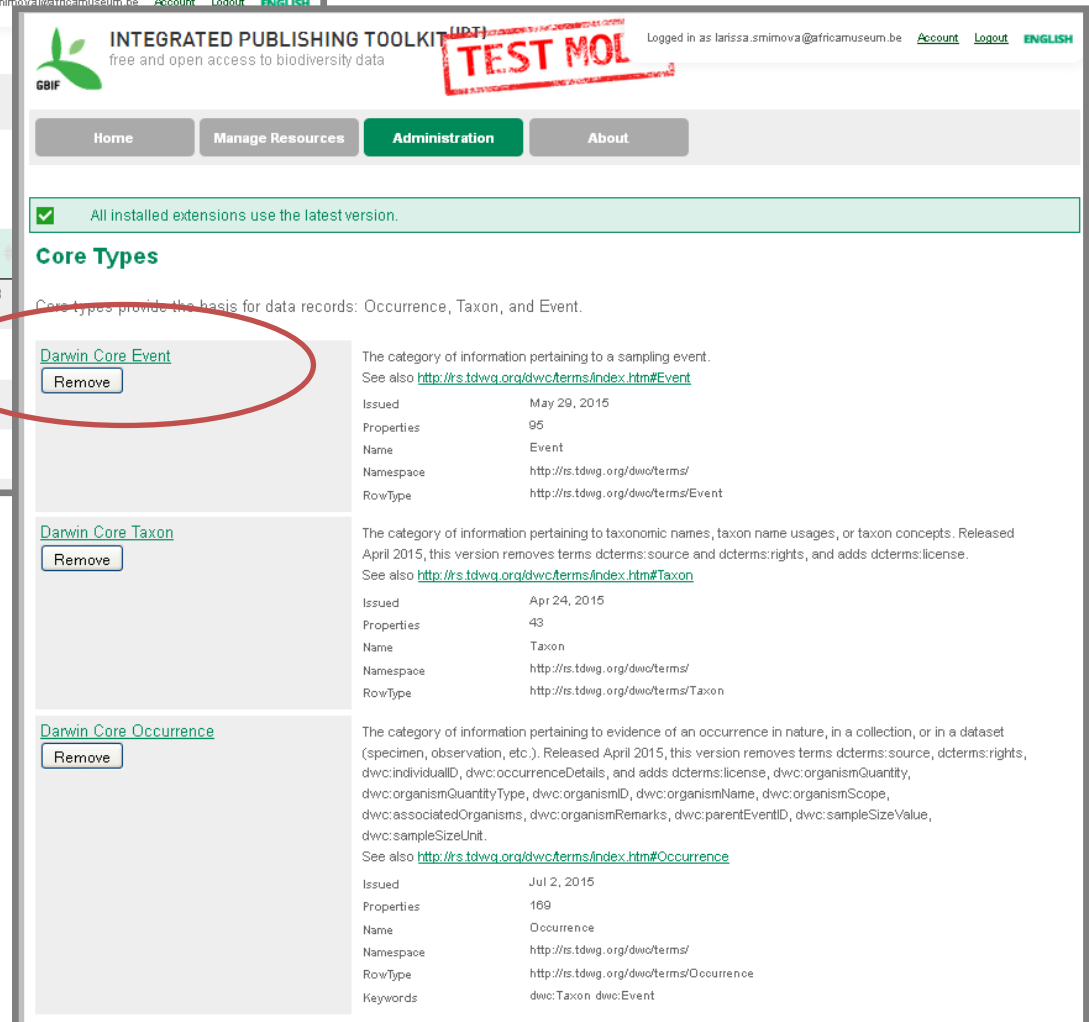
Home Manage Resources Administration About

Resources you have rights to manage

Name	Organisation	Type	Subtype	Records	Last modified	Last publication
Aves Tanzanian Collection	Not registered	--	--	7,774	2015-08-20	2015-07-08
benthos test larissa	Not registered	--	--	0	2015-08-17	Not published
prueba sibbr	Not registered	--	--	0	2015-08-11	Not published
Benthic communities in Logarou lagoon	Not registered	--	--	0	2015-08-09	Not published

NEW FEATURES

- **Event core and extensions**
- Possibility to assign DOIs to resources
- Auto-generation of a citation for a resource including the version number and DOI
- Machine readable licenses



GBIF INTEGRATED PUBLISHING TOOLKIT (IPT) v.2.3
free and open access to biodiversity data

Logged in as larissa.sminova@afriamuseum.be Account Logout ENGLISH

Home Manage Resources Administration About

☒ All installed extensions use the latest version.

Core Types

Core types provide the basis for data records: Occurrence, Taxon, and Event.

[Darwin Core Event](#)

Remove

The category of information pertaining to a sampling event. See also <http://rs.tdwg.org/dwc/terms/index.htm#Event>

Issued May 29, 2015

Properties 95

Name Event

Namespace <http://rs.tdwg.org/dwc/terms/>

RowType <http://rs.tdwg.org/dwc/terms/Event>

[Darwin Core Taxon](#)

Remove

The category of information pertaining to taxonomic names, taxon name usages, or taxon concepts. Released April 2015, this version removes terms dcterms:source and dcterms:rights, and adds dcterms:license. See also <http://rs.tdwg.org/dwc/terms/index.htm#Taxon>

Issued Apr 24, 2015

Properties 43

Name Taxon

Namespace <http://rs.tdwg.org/dwc/terms/>

RowType <http://rs.tdwg.org/dwc/terms/Taxon>

[Darwin Core Occurrence](#)

Remove

The category of information pertaining to evidence of an occurrence in nature, in a collection, or in a dataset (specimen, observation, etc.). Released April 2015, this version removes terms dcterms:source, dcterms:rights, dwc:individualID, dwc:occurrenceDetails, and adds dcterms:license, dwc:organismQuantity, dwc:organismQuantityType, dwc:organismID, dwc:organismName, dwc:organismScope, dwc:associatedOrganisms, dwc:organismRemarks, dwc:parentEventID, dwc:sampleSizeValue, dwc:sampleSizeUnit. See also <http://rs.tdwg.org/dwc/terms/index.htm#Occurrence>

Issued Jul 2, 2015

Properties 169

Name Occurrence

Namespace <http://rs.tdwg.org/dwc/terms/>

RowType <http://rs.tdwg.org/dwc/terms/Occurrence>

Keywords dwc:Taxon dwc:Event

THE IMPORTANCE OF METADATA

Metadata is **ESSENTIAL** for sample-based data:

- ✓ Documenting the project data: description of sampling/monitoring activity.
- ✓ Documenting sampling methods: protocols, data quality and sampling steps.

Dataset documentation enables users to evaluate the **fitness for use** of a dataset

METADATA: PROJECT DATA & SAMPLING METHODS

Project Data

Please enter metadata about the project under which the data in this resource were produced.

Title*

Identifier



A unique identifier for the research project. This can be used to link multiple dataset/EML document instances that are associated in some way with the same project, e.g. a monitoring series. The nature of the association can be described in the project description.

Funding



Study Area Description



Design Description



Section

[Basic Metadata](#)

[Geographic Coverage](#)

Sampling Methods

Please enter metadata about the sampling methods used for the data represented by the resource.

Study Extent*



Sampling Description*



Quality Control



Step Description*



Section

[Basic Metadata](#)

[Geographic Coverage](#)

[Taxonomic Coverage](#)

[Temporal Coverage](#)

[Keywords](#)

[Associated Parties](#)

[Project Data](#)

[Sampling Methods](#)

[Citations](#)

[Collection Data](#)

[External links](#)

[Additional Metadata](#)

METADATA EXAMPLE: STUDY EXTENT

- *“Seasonal (4 seasons) records of zoobenthic communities across confinement gradient in 5 coastal lagoons (Mazoma, Tsopeli, Tsoukalio, Rodia and Logarou)”*
- *“The floristic inventories was conducted at the Sierra Nevada (Andalusia, southeast Spain). Forest cover in Sierra Nevada is dominated by pine plantations (Pinus halepensis Mill., Pinus pinaster Ait., Pinus nigra Arnold. subsp. salzmannii (Dunal) Franco, and Pinus sylvestris L.) that covering approximately 40,000 ha...”*
- *“Transect walks were conducted once in two weeks from the beginning of October to the end of June on the territory of Israel”*

METADATA EXAMPLE: STUDY EXTENT

- “Seasonal (4 seasons) records of zoobenthic communities across confinement gradient in 5 coastal lagoons (Mazoma, Tsopeli, Tsoukalio, Rodia and Logarou)”
- “The floristic inventories was conducted at the Sierra Nevada (Andalusia, southeast Spain). Forest cover in Sierra Nevada is dominated by pine plantations (*Pinus halepensis* Mill., *Pinus pinaster* Ait., *Pinus nigra* Arnold. subsp. *salzmannii* (Dunal) Franco, and *Pinus sylvestris* L.) that covering approximately 40,000 ha...”
- “Transect walks were conducted once in two weeks from the beginning of October to the end of June on the territory of Israel.”

Sampling area/habitat + Sampling frequency

METADATA EXAMPLE: SAMPLING DESCRIPTION

- *“Box corer for 0.03 square meters area sampled”*
- *“Transect lengths range between 300 and 600 m in length, and divide into 50 m sections. Transect usually cover a single habitat type. In each visit, transect-walkers count all butterfly species that can be seen within a range of 5 m range. Special behaviours (egg laying or nectaring), as well as butterfly larvae or eggs, can be registered as well.”*
- *“The dataset includes data from 1,879 sites, collected using standard RLS survey methods, described in detail in an online methods manual. Surveys involve underwater visual census by SCUBA divers along a 50 m transect line, laid along a depth contour on hard substrate. All fish species observed within 5 m of the transect line were recorded on a waterproof datasheet as the diver swam slowly along the line (at approximately 2 m/min)...”*

METADATA EXAMPLE: SAMPLING DESCRIPTION

- “Box corer for 0.03 square meters area sampled”
- “Transect lengths range between 300 and 600 m in length, and divide into 50 m sections. Transect usually cover a single habitat type. In each visit, transect-walkers count all butterfly species that can be seen within a range of 5 m range. Special behaviours (egg laying or nectaring), as well as butterfly larvae or eggs, can be registered as well.”
- “The dataset includes data from 1,879 sites, collected using standard RLS survey methods, described in detail in an online methods manual. Surveys involve underwater visual census by SCUBA divers along a 50 m transect line, laid along a depth contour on hard substrate. All fish species observed within 5 m of the transect line were recorded on a waterproof datasheet as the diver swam slowly along the line (at approximately 2 m/min).”

Description of sampling procedures

METADATA EXAMPLE: QUALITY CONTROL

- *“Record validation and cleaning was incorporated at several steps of the documentation process: the scientific names on labels were checked with a taxonomic thesaurus.... We assigned scientific names in accordance to current taxonomy trends. Geographic coordinates were verified using the “Check Coordinates” function in DIVA-GIS”*
- *“Every reported record is flagged ‘forApproval’. Record status is changed to “Approved” upon and by expert only“*
- *“Each sampling plot was checked to ensure whether the geographical coordinates were correct. The databases of International Plant Names Index (IPNI, 2013) and Catalogue of Life/Species 2000 to verify the taxonomical classification”*

METADATA EXAMPLE: QUALITY CONTROL

- *“Record validation and cleaning was incorporated at several steps of the documentation process: the scientific names on labels were checked with a taxonomic thesaurus.... We assigned scientific names in accordance to current taxonomy trends. Geographic coordinates were verified using the “Check Coordinates” function in DIVA-GIS”*
- *“Every reported record is flagged ‘forApproval’. Record status is changed to “Approved” upon and by expert only“*
- *“Each sampling plot was checked to ensure whether the geographical coordinates were correct. The databases of International Plant Names Index (IPNI, 2013) and Catalogue of Life/Species 2000 to verify the taxonomical classification”*

Quality control actions

METADATA EXAMPLE: STEP DESCRIPTION


- *“Replicates were collected at two stations, one inside and one outside the lagoon, using a box corer (0.03 m² sampling surface). The samples were sieved through a 0.5 mm mesh sieve, stained with Rose Bengal and preserved in 4% formalin. All benthic organisms were sorted in the lab and identified and counted using stereo- and microscopes”*
- *“1. Selection and rout of transect is accompanied by an scientific/principal investigator. 2. Identification of the capabilities of the observer are verified tested along timeline 3. Time start and weather condition are recorded first. 4. During transect walk, the number seen of every species is counted....”*
- *“This dataset was compiled from data collected in a combination of collaborative surveys with scientific colleagues worldwide, targeted RLS field campaigns and ad-hoc local surveys by trained RLS divers at their regular dive sites or when on holidays. Field campaigns involved small groups of divers (usually 4 to 8) undertaking survey dives over a period of four days to two weeks under the direction and supervision of a scientist or experienced survey diver“*

METADATA EXAMPLE: STEP DESCRIPTION





- *“Replicates were collected at two stations, one inside and one outside the lagoon, using a box corer (0.03 m² sampling surface). The samples were sieved through a 0.5 mm mesh sieve, stained with Rose Bengal and preserved in 4% formalin. All benthic organisms were sorted in the lab and identified and counted using stereo- and microscopes”*
- *“1. Selection and rout of transect is accompanied by an scientific/principal investigator. 2. Identification of the capabilities of the observer are verified tested along timeline 3. Time start and weather condition are recorded first. 4. During transect walk, the number seen of every species is counted....”*
- *“This dataset was compiled from data collected in a combination of collaborative surveys with scientific colleagues worldwide, targeted RLS field campaigns and ad-hoc local surveys by trained RLS divers at their regular dive sites or when on holidays. Field campaigns involved small groups of divers (usually 4 to 8) undertaking survey dives over a period of four days to two weeks under the direction and supervision of a scientist or experienced survey diver“*

Methods, procedures, processing steps

METADATA AND DATA PAPERS

PENSOFT [About Pensoft](#) [Books](#) [E-Books](#) [Journals](#) [News & Blog](#) [Contact](#)  Larissa Smit





Journal home About Browse articles

 ZooKeys 475: 119-145 (22 Jan 2015)
doi: 10.3897/zookeys.475.8556   

Study extent description

Over 2,000 locations in estuaries, inland rivers, streams, canals, and enclosed waters in Flanders, Belgium have been sampled, from March to November, since 1992 (Figure 10). In 2001, these locations were consolidated in a monitoring network ("VISmeetnet") of 900 sampling points. Four locations in the Yser estuary and 43 locations in the Scheldt were sampled since 2001. The Yser estuary only covers a small geographical area, the Scheldt estuary is with 350 km² one of the largest estuaries in Europe. It is also one of the few remaining European estuaries. The Scheldt estuary includes the entire gradient from fresh to saltwater tidal areas (Van den Bergh et al. 2004). The sampling locations in the Scheldt estuary are mainly located in the River Scheldt, Rivers Durme, Rupel, Dijle, Zenne and Nete.

The geographic coordinates in both datasets are those of the defined sampling locations (dwc:locationID). However, as these coordinates are not always exact the actual catch, which may be located further up- or downriver, the coordinate uncertainty (dwc:coordinateUncertaintyInMeters) has been set to 250 meter.

Share this article    

Contents Article info Citation Metrics

Figs Tabs Taxa

[Geographic coverage](#)

[Flanders](#)

[Inland waters](#)

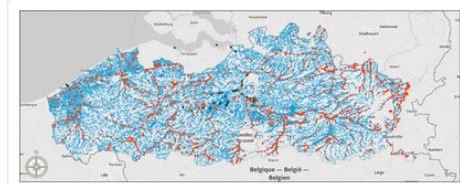







Figure 10.

Map of all sampling locations in VIS. Orange points represent inland waters, green points represent estuarine waters. Image created in CartoDB and Mapbox, basemap by OpenStreetMap.

Sampling description

PENSOFT [About Pensoft](#) [Books](#) [E-Books](#) [Journals](#) [News & Blog](#) [Contact](#)  Larissa Smirnova

Journal home About Browse articles

 PhytoKeys 35: 1-15 (17 Feb 2014)
doi: 10.3897/phytokeys.35.8363   

Methods

Method step description





This inventory was undertaken in 2004 and the database generated contains information relative to forest attributes and occurrence data (see below). This information, originally stored in a Microsoft Access database, has been integrated into the project's information system.

Study extent description: The floristic inventories were conducted at the main forest units of the Sierra Nevada (Andalusia, SE Spain). Forest cover in Sierra Nevada is dominated by pine plantations (*Pinus halepensis* Mill., *Pinus pinaster* Ait., *Pinus nigra* Arnold subsp. *salzmannii* (Dunal) Franco, and *Pinus sylvestris* L.) covering approximately 40, 000 ha. Most of them were planted in the period 1960–1980. The main native forests of Sierra Nevada are dominated by the evergreen holm oak *Quercus ilex* subsp. *ballota* (Desf.) Samp. occupying low and medium mountain areas (8, 800 ha.) and Pyrenean oak *Quercus pyrenaica* Willd ranging from 1, 100 to 2, 000 m a.s.l., covering about 2, 000 ha. Autochthonous pine *Pinus sylvestris* var. *nevadensis* forests can also be found in small patches at high altitudes with a characteristically low tree cover.

Sampling description: SINFONEVADA Forest Inventory was established over an extensive network of 600 long-term permanent plots distributed within the main forest units of the Sierra Nevada: pine plantations, evergreen *Quercus ilex* forests, and deciduous broadleaf forests. The network of plots is a random sample stratified by land cover and altitude, covering a gradient of 974–2439 m a.s.l. (Figure 5).

Each inventory plot has three sampling units: i) a forest inventory plot (20 × 20 m); ii) a 5-m radius subplot for the estimation of the regeneration; iii) and a 10-m radius subplot for species composition and abundance.

Each live tree with a diameter at breast height (dbh) > 7.5 cm was tallied by species and dbh in the forest inventory plot. This information was used to calculate forest attributes (tree basal

Share this article    

Contents Article info Citation Metrics Comments

Figs Tabs Taxa References

[Top](#)

[Abstract](#)

[Resource Citation](#)

[Project details](#)

[Data published through](#)

[Taxonomic coverage](#)

[Taxonomic ranks](#)

[Spatial coverage](#)

[Natural collections description](#)

[Methods](#)

[Datasets](#)

[Dataset description](#)

[Acknowledgements](#)

[References](#)

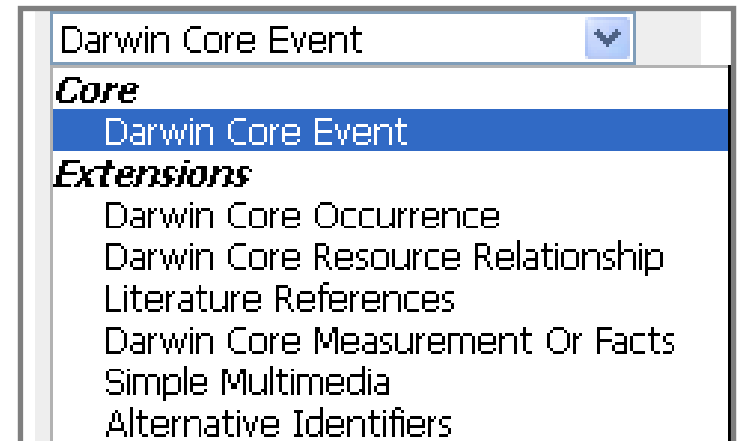
THE NEW EVENT CORE

In IPT 2.3, you can select the new event core:

- When creating new resources
- When mapping terms





































Three classes are available:

- Event
- Location
- GeologicalContext



EVENT CORE – KEY TERMS

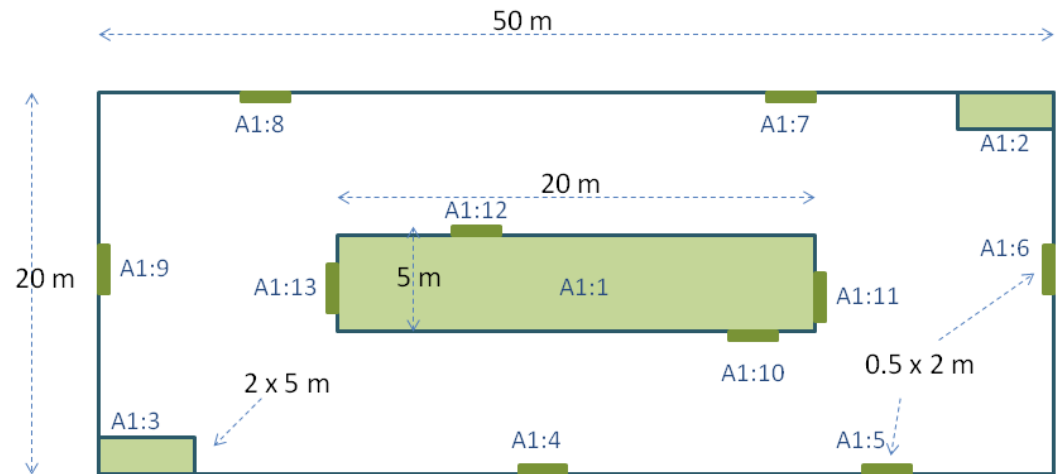
Event

 parentEventID	<input type="text"/>		<input type="text"/>
 samplingProtocol	<input type="text"/>		<input type="text"/>
 sampleSizeValue	<input type="text"/>		<input type="text"/>
 sampleSizeUnit	<input type="text"/>		 
 samplingEffort	<input type="text"/>		<input type="text"/>
 eventDate	<input type="text"/>		<input type="text"/>
 eventTime	<input type="text"/>		<input type="text"/>
 startDayOfYear	<input type="text"/>		<input type="text"/>
 endDayOfYear	<input type="text"/>		<input type="text"/>
 year	<input type="text"/>		<input type="text"/>
 month	<input type="text"/>		<input type="text"/>
 day	<input type="text"/>		<input type="text"/>
 verbatimEventDate	<input type="text"/>		<input type="text"/>
 habitat	<input type="text"/>		<input type="text"/>
 fieldNumber	<input type="text"/>		<input type="text"/>
 fieldNotes	<input type="text"/>		<input type="text"/>
 eventRemarks	<input type="text"/>		<input type="text"/>

NEW DWC TERMS

eventID and parentEventID

- Sampling events can be related to each other (e.g., nested samples) via a common parent identifier
- Information on the nature of the relationship (e.g. part of a monitoring series) can be described in the project section of the accompanying metadata.



A Whittaker plot with sub-plots of varying area

NEW DWC TERMS

sampleSize
























A numeric value for the time duration, length, area or volume involved in the sampling event.

sampleSizeUnit

The unit of measurement used in the sampling event, e.g., minute, hour, day, metre, square metre, cubic metre.

sampleSize	sampleSizeUnit
2	hour
6	m2
1	litre
3	km

OCCURRENCE EXTENSION: KEY TERMS

Occurrence		
 occurrenceID	<input type="text"/>	<input type="text"/>
 catalogNumber	<input type="text"/>	<input type="text"/>
 occurrenceRemarks	<input type="text"/>	<input type="text"/>
 recordNumber	<input type="text"/>	<input type="text"/>
 recordedBy	<input type="text"/>	<input type="text"/>
 individualCount	<input type="text"/>	<input type="text"/>
 organismQuantity	<input type="text"/>	<input type="text"/>
 organismQuantityType	<input type="text"/>	<input type="text"/>  
 sex	<input type="text"/>	<input type="text"/>
 lifeStage	<input type="text"/>	<input type="text"/>
 reproductiveCondition	<input type="text"/>	<input type="text"/>
 behavior	<input type="text"/>	<input type="text"/>
 establishmentMeans	<input type="text"/>	<input type="text"/>
 occurrenceStatus	<input type="text"/>	<input type="text"/>
 preparations	<input type="text"/>	<input type="text"/>
 disposition	<input type="text"/>	<input type="text"/>
 otherCatalogNumbers	<input type="text"/>	<input type="text"/>
 associatedMedia	<input type="text"/>	<input type="text"/>
 associatedReferences	<input type="text"/>	<input type="text"/>
 associatedSequences	<input type="text"/>	<input type="text"/>
 associatedTaxa	<input type="text"/>	<input type="text"/>

DATA: NEW DWC TERMS

organismQuantity

A number or enumeration value for the quantity of organisms. Use with organismQuantityType to indicate the type of entity that is being quantified

organismQuantityType

The type of entity to which the number or enumeration value reported for the quantity of organisms in organismQuantity refers.

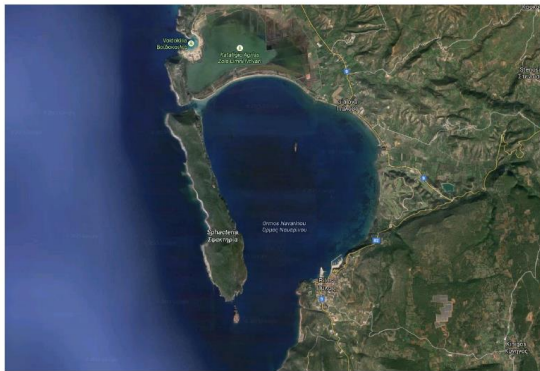
organismQuantity	organismQuantityType
14	Individuals
r	BraunBlanquetScale
31	%Biomass
0.4	%Species

UC1 - AQUATIC INVERTEBRATES



Session 08, Sample-data use case 1: Aquatic invertebrates¹

Environmental impact assessment of oil pollution accidents in Gialova lagoon and Navarino Bay.



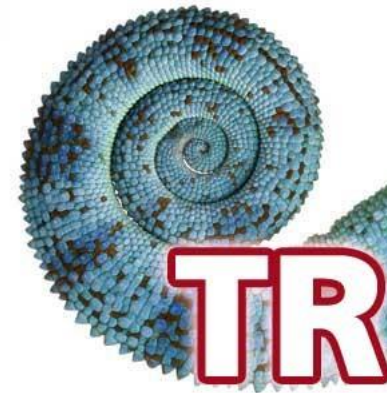
Description

The area of Navarino bay and the Gialova lagoon is located in the west coast of Greece, connected to the Mediterranean sea through the Ionian sea. Due to its strategic location, this area has unfortunately suffered several oil spill accidents in the past (e.g. the *Irenes Serenade* in 1980, the *Happy Leader* and the *Worthy* in 1990, the *Illiad* in 1993).

Several protection figures have been assigned to the area such a Natura 2000 site by the European Union or a Marine Protected area (MPA) for cetaceans by ACCOBAMS, but these efforts have proven insufficient to protect the area. The environmental agency of the regional government of Peloponnese is evaluating other more-effective protection figures. They turned to you as the national reference for biodiversity information to find out which data is openly available.

The initial queries in GBIF.org show insufficient data coverage for this area for the kind of work planned by the policy makers. To help you mitigate this situation, your contact in the

¹ Please note that this use case is based on a [fictitious story](#) built for instructional purposes.



Blue Group

UC1 - AQUATIC INVERTEBRATES



- Seasonal study of brackish water invertebrates
- 7 stations
- Van Veen Grab, hand operated
- Abundance, biomass, environmental variables

- abundance/ biomass comparison, distribution of species in geometric abundance and geometric size classes
- structure and dynamics of the communities
- assessment of anthropogenic impact



UC1 - AQUATIC INVERTEBRATES - DATA

	A	B	C	D						
1	Sampling area:	Gialova lagoon (SW Pelopponese), Ionian Sea, Greece								
2	Sampling period:	4: I: 27 May-3 June 1995; II: 26 September-3 October 1995; III: 28 November-5								
3	Stations:	7, Stations; Sumbols: A, B, C, D, E, F, G								
4	Replicate samples:	5; Symbols: 1, 2, 3, 4, 5.								
5	Co-ordinates:	A: 36o, 57' 385 N; 21o 39' 967; B:36o, 57' 380 N; 21o 39' 965; C: 36o, 57' 251								
6	Sampling	A	B	C	D	E	F	G	H	I
7		1	General classification of sampling area:		brackish water lagoon					
		2	Classification by Station:							
		3		A	zone III, according to confinement scale proposed by Guelorget & Perthuisot (1992)					
		4		B	zone III, according to confinement scale proposed by Guelorget & Perthuisot (1992)					
		5		C	transitional zone					
		6		D	transitional zone					
		7		E	zones IV-V, according to confinement scale proposed by Guelorget & Perthuisot (1992)					
		8		F	zones IV-V, according to confinement scale proposed by Guelorget & Perthuisot (1992)					
					(1992)					

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	parentEventID	sampleSize	samplingProtoc	habitat	startDayO	endDayOf	year	locality	municipali	waterBody	country	verbatimL	verbatim	
2		IA	0.25	hand operated	brackish w	147	154	1995	Gialova lag	SW Pelopp	Ionian Sea	Greece	36.9564	21.66
3	IA	IA1	0.05	hand operated	brackish w	147	154	1995	Gialova lag	SW Pelopp	Ionian Sea	Greece	36.9564	21.66
4	IA	IA2	0.05	hand operated	brackish w	147	154	1995	Gialova lag	SW Pelopp	Ionian Sea	Greece	36.9564	21.66
5	IA	IA3	0.05	hand operated	brackish w	147	154	1995	Gialova lag	SW Pelopp	Ionian Sea	Greece	36.9564	21.66
6	IA	IA4	0.05	hand operated	brackish w	147	154	1995	Gialova lag	SW Pelopp	Ionian Sea	Greece	36.9564	21.66
7	IA	IA5	0.05	hand operated	brackish w	147	154	1995	Gialova lag	SW Pelopp	Ionian Sea	Greece	36.9564	21.66
8		IIA	0.25	hand operated	brackish w	269	276	1995	Gialova lag	SW Pelopp	Ionian Sea	Greece	36.9564	21.66
9	IIA	IIA1	0.05	hand operated	brackish w	269	276	1995	Gialova lag	SW Pelopp	Ionian Sea	Greece	36.9564	21.66
10	IIA	IIA2	0.05	hand operated	brackish w	269	276	1995	Gialova lag	SW Pelopp	Ionian Sea	Greece	36.9564	21.66
11	IIA	IIA3	0.05	hand operated	brackish w	269	276	1995	Gialova lag	SW Pelopp	Ionian Sea	Greece	36.9564	21.66
12	IIA	IIA4	0.05	hand operated	brackish w	269	276	1995	Gialova lag	SW Pelopp	Ionian Sea	Greece	36.9564	21.66

UC2 - LEPIDOPTERA



Session 08, Sample-data use case 2:

Lepidoptera¹

Butterfly counts using Pollard walks in Israel



Description

The Israeli Butterfly Amateur Network (IBAN) has been capturing data about the occurrence of Lepidoptera species in the country since 2009. An extensive network of amateur observers use a standard protocol based on Pollard walks² to capture the information. They send this information in the form of paper sheets to a central office where it should be digitized and stored in a local database. The main product of the sampling effort is a paper annual report based on the sightings.

The central office is staffed mainly with volunteers and the data received exceeds their digitization capacity, so the paper data sheets are piling up, undigitized. The steering committee of the organization is considering more agile production of digital data as they would like to start publishing the data online regularly. They would also like to start processing pictures that their volunteers are already capturing with their mobile phones. Their ultimate objective is to raise the profile of the network and strengthen collaborations with local and regional governments to influence the conservation policies for lepidoptera in the country.

At the moment, there is no formal agreement with the amateurs capturing the data about data use. The steering committee has some concerns that when they start publishing the data online, they will have to formalize that component.

¹ Please note that this use case is based on a fictional story built for instructional purposes.

² Image from "Hipparchia psidice 1" by Gideon Pisanly (Gidip) גידון פיסנלי - Own work. Licensed under CC BY 3.0 via Commons - <https://goo.gl/4H-yut>

³ More information about Pollard walks can be found at http://faculty.wvu.edu/imcl/Methods/billy_walk.pdf



Green Group



Red Group

UC2 - LEPIDOPTERA

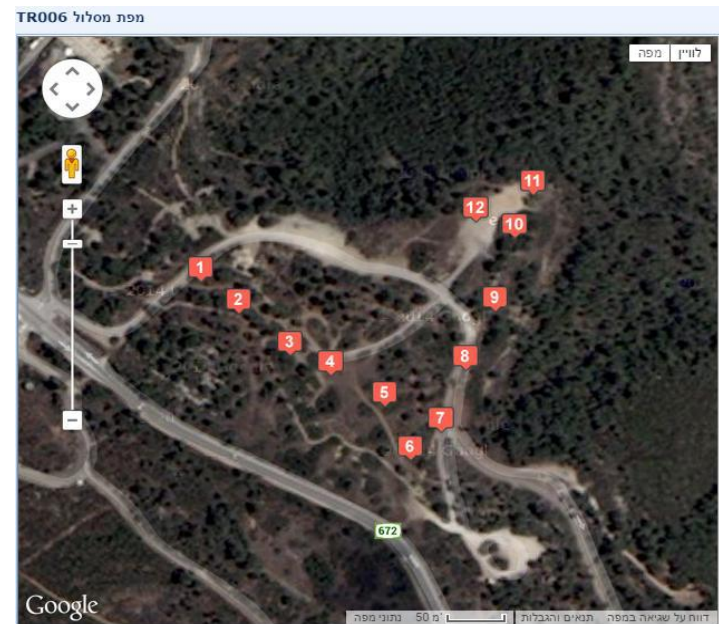
Israel Systematic Butterflies Monitoring Scheme (BMS-IL)

- “Pollard Walks” protocol: Observations along fixed transects.
- Transects lengths range between 300 to 600 m in length, and divide into 50 m sections.
- For each species: record the number of individuals seen in each section

The screenshot shows the 'Add a sighting' form in the BMS-IL application. It includes tabs for 'Home', 'Sightings Data', 'Add a sighting', 'Add Transect report', and 'Transect data'. The 'Add a sighting' tab is active, showing options to 'Add Observation', 'Save and add another species', 'Save and exit', 'Clear All', and 'Report as Sporadic'. Below these are tabs for 'General', 'Butterflies', and 'Help'. The 'Butterflies' tab is selected, showing a form for 'Adults only' with a dropdown for 'by list' (currently 'Satyrium ilicis') and a dropdown for 'by pictures' (currently 'Lycaenidae'). Below this is a table for recording sightings in 12 sections. The table has columns for 'Section' and 'Quantity'. The data entered is as follows:

Section	Quantity	Section	Quantity	Section	Quantity
1		7		12	
2	2	8	3		
3		9			
4		10			
5		11	1		
6		12			

Below the table is a 'Comment on behavior' text area and a small image of a butterfly. At the bottom, there is a link that says 'Select 'None Seen' to report none'.



Transect map

UC2 - LEPIDOPTERA - DATA

Transect Data: Tr. Code: TR006, Event Date: Mar-09-2015
Results: Abundance=13, Species richness=5, Time=45 min

<http://www.gluecad.com/buttdb/hompage.asp?lng=eng>

Home

Sightings Data

Add a sighting

Add Transect report

Transect data

עברית

Contact

Login

📍

מיקום

EMS-Israel

▼

Select Transect

▼

Show in graph

▼

Select Butterfly

▼

📍

Transect map

כל המסלולים, כל המצפים

BMS-Israel

#	Date	Transect	Time		Species Name	Section code																							
			From	To		1	1e	2	2e	3	3e	4	4e	5	5e	6	6e		7	7e	8	8e	9	9e	10	10e	11	11e	12
17	Mar,11 2015	TR027	11:20	12:00	Pieris brassicae																1								
18	Mar,11 2015	TR027	11:20	12:00	Pieris rapae							3	2	1			1			1	2	2							
19	Mar,11 2015	TR027	11:20	12:00	Complex 13-14							2									2								
Gilad yaar habanin						Abundance/Richness:												17 / 3											
20	Mar,9 2015	TR006	09:45	10:30	Archon apollinus	1															1					3			
21	Mar,9 2015	TR006	09:45	10:30	Pieris rapae	2																							
22	Mar,9 2015	TR006	09:45	10:30	Pontia daplidice																					1			
23	Mar,9 2015	TR006	09:45	10:30	Anthocharis cardamines																1								
24	Mar,9 2015	TR006	09:45	10:30	Gonepteryx cleopatra																					4			
Carmel Hurshan haarbaim						Abundance/Richness:												13 / 5											
25	Mar,8 2015	TR007	10:13	10:35	Archon apollinus																		1						
26	Mar,8 2015	TR007	10:13	10:35	Pieris brassicae									1				1											
27	Mar,8 2015	TR007	10:13	10:35	Gonepteryx cleopatra																2			2					
28	Mar,8 2015	TR007	10:13	10:35	Lasimemata megera emilyssa																				1				
Kibutz Sasa						Abundance/Richness:												8 / 4											
29	Mar,7 2015	TR054	12:00	12:30	Papilio machaon											3					1	1							
30	Mar,7 2015	TR054	12:00	12:30	Anthocharis cardamines									2	1						3					1		1	
31	Mar,7 2015	TR054	12:00	12:30	Vanessa atalanta																1								
32	Mar,7 2015	TR054	12:00	12:30	Complex 10-12	3											3	2	6	4	3		3		1		1		
Nachshonim Kakal forest						Abundance/Richness:												43 / 4											

Page 2

Of 364

Displaying 17 To 32 Of 6190 Items

UC3 - MACROPHYTES



Session 08, Sample-data use case 3: Macrophytes¹

Riparian vegetation coverage in reaches of the Kinzig river (Germany)



Description

A research team within the Senckenberg Nature Research Society (*Senckenberg Gesellschaft für Naturforschung*, SGN) maintains a monitoring programme in different reaches of the Kinzig river in Germany. They have set permanent plots where vascular plants are surveyed twice every year (early spring and summer) using the Braun-Blanquet cover-abundance scale (also known as 'the relevé method'²).

This monitoring effort is funded by an international long-term monitoring project which also provides the means to publish the resulting information. The German team sends their data to the data unit via email every year and some time afterwards the information appears in a specialized online portal.

The team is making a big effort in both fieldwork and data preparation and would like to obtain more benefits from it. They are exploring other ways to make the data available, but they are unsure of which requirements in terms of data licensing and restrictions to re-publishing they are subject to as part of the project that is funding their research.

¹ Please note that this use case is based on a fictitious story built for instructional purposes.

² Image "Kinzig 292" by Marion Schneider & Christoph Aistleitner - own picture - User:Mediocrity. Licensed under Public Domain via Commons -

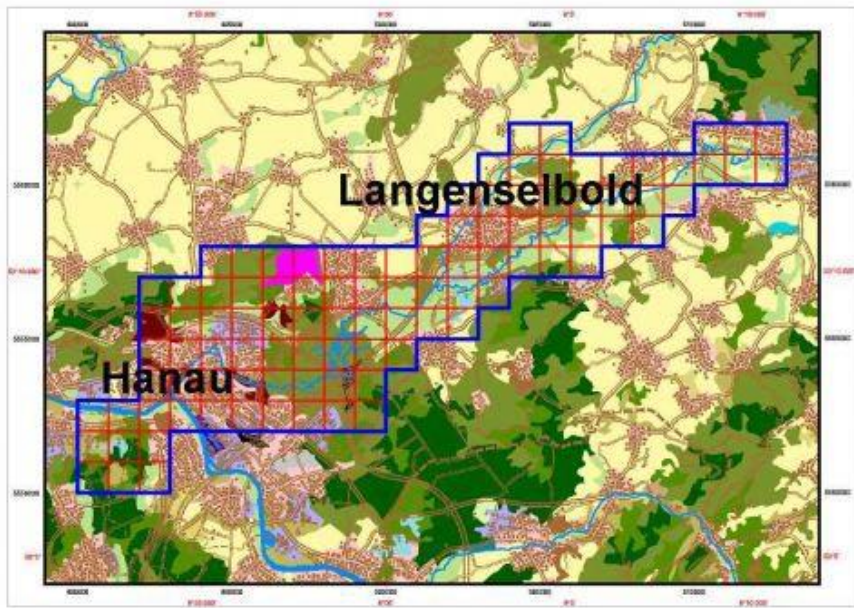
https://commons.wikimedia.org/wiki/File:Kinzig_292.jpg#/media/File:Kinzig_292.jpg

³ For a quick description of the relevé method, please check the section '1. Introduction, Description' in http://files.dnr.state.mn.us/ecoincbs/relve/relve_sintolepage.pdf by the University of Minnesota..



Yellow Group

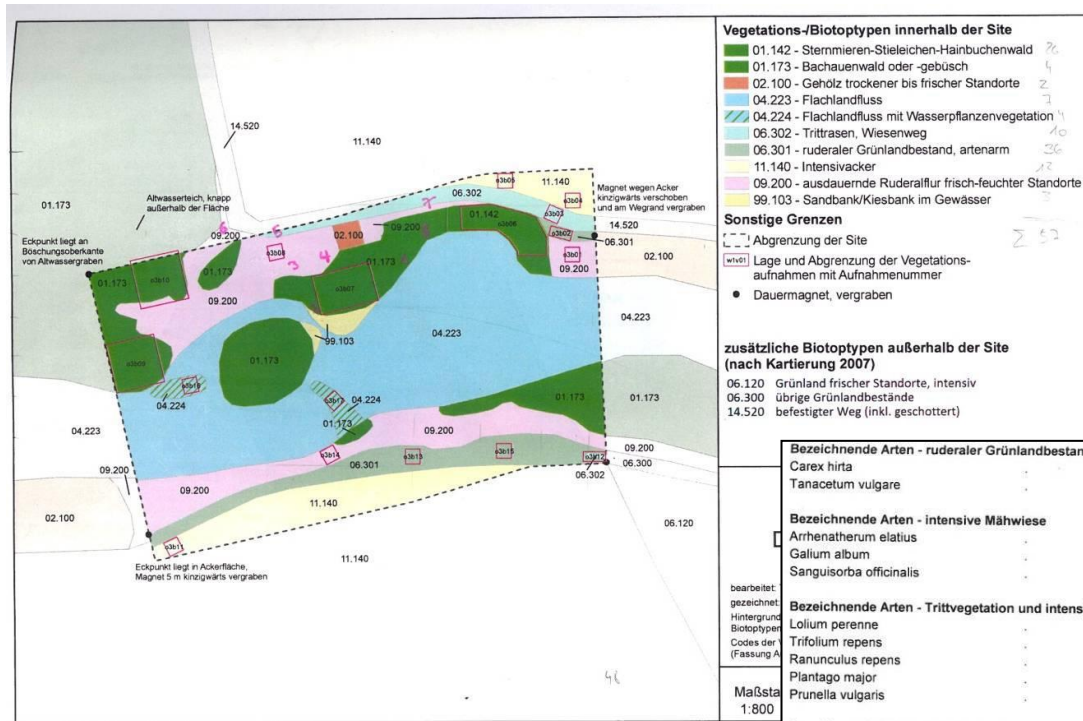
UC3 - MACROPHYTES



Rhine-Main-Observatory (RMO) macrophytes biodiversity

- Percentage coverage of riparian vascular plant species at 21 stream reaches of the Kinzig river
- Braun-Blanquet vegetation relevée; annual vegetation relevées (spring for early-bloomers, summer for rest of vegetation)
- Sampled annually

UC3 - MACROPHYTES - DATA



Biotic components:

Vegetation coverage

Average height per biotype

Species abundance information following the Braun-Blanquet scale

Abiotic components:

Land use

Meteorological data

Hydrological data

Physico-chemical data

Bezeichnende Arten - ruderaler Grünlandbestand											
Carex hirta											
Tanacetum vulgare											
Bezeichnende Arten - intensive Mähwiese											
Arrhenatherum elatius											
Galium album											
Sanguisorba officinalis											
Bezeichnende Arten - Trittvegetation und intensive Weide											
Lolium perenne											
Trifolium repens											
Ranunculus repens											
Plantago major											
Prunella vulgaris											
Bezeichnende Arten - Grünland- und Rasenvegetation											
Agrostis stolonifera											
Agrostis capillaris											
Holcus lanatus											
Plantago lanceolata											
Taraxacum Sec. Ruderalia											
Trifolium pratense											
Ranunculus acris											
Stellaria graminea											
Achillea millefolium											
Festuca rubra											
Rumex acetosa											
Veronica chamaedrys											
Vicia cracca											
Deschampsia cespitosa											
Ajuga reptans											
Cardamine pratensis											
Lysimachia nummularia											
Rumex obtusifolius											

UC4 - CORAL REEF SURVEYS



Session 08, sample-data use case 4: Coral Reef Surveys¹

Effects of Climate Change in coral reefs



Description

The Reef Life Survey (RLS) has been collecting information from 1,879 sites in coral and rocky reefs distributed worldwide. They use standardized methods, offering new opportunities to assess broad-scale spatial patterns in community structure. The group benefits from contributions from investigators associated with science and conservation agencies worldwide, and the assistance of a team of over 100 recreational scuba divers, who undertook training in scientific techniques for underwater surveys and voluntarily contributed skills, expertise and their time to data collection.

A small team of researchers collaborating with RLS would like to assess the effects of climate change and the increasing acidification of the seas in coral reefs worldwide. They have been making a special effort to clean and complete a dataset that originally came from the volunteers' network.

The board of RLS would like to publish their data online, including the clean dataset produced by the climate change group. Unfortunately, this group believes that they have increased the value of the data substantially and they do not agree on sharing those data openly till they gather enough evidence to get to conclusive results and their studies are published.

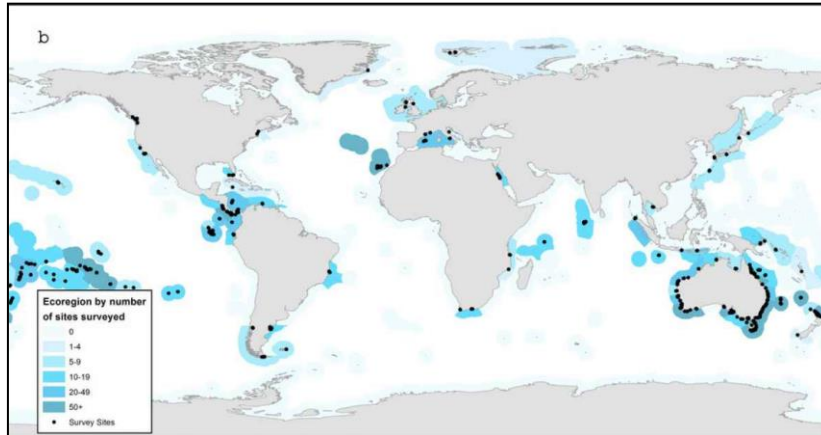
¹ Please note that this use case is based on a [fictitious story](#) built for instructional purposes.

² Image by Fernando Alvarez, obtained via <http://www.freemages.com/photo/lovin-scuba-1402126>.



Purple Group

UC4 - CORAL REEF SURVEYS



- 2006-2013
- 1,879 sites in coral and rocky reefs distributed worldwide
- Change in ecosystem, effects of climate change and the increasing acidification of the seas

- visual census by scuba divers along 50 m underwater transects
- fish and invertebrate species are recorded with abundance estimates for less abundant species
- photographs are also taken of the coral or seaweed cover
- envi parameters: water temperature, pH and salinity



UC4 - CORAL REEF SURVEYS - DATA

ID	Diver	Buddy	Site No.	Site Name	Latitude	Longitude	Date	vis	Directi	Time	P-Qs	Depth	Method	Block	Code
										12:00		0, 1, 2	0, 1, 2		
1	DIVER	BUDDY	SITE	Site Name	Latitude	Longitude	1/01/2012	VIS	DIREC	13:30	P-Qs	DEPT	METHO	BLOCK	CODE
2	DIVER	BUDDY	SITE	Site Name	Latitude	Longitude	1/01/2012	VIS	DIREC	13:30	P-Qs	DEPT	METHO	BLOCK	CODE
3	DIVER	BUDDY	SITE	Site Name	Latitude	Longitude	1/01/2012	VIS	DIREC	13:30	P-Qs	DEPT	METHO	BLOCK	CODE
4	DIVER	BUDDY	SITE	Site Name	Latitude	Longitude	1/01/2012	VIS	DIREC	13:30	P-Qs	DEPT	METHO	BLOCK	CODE
5	DIVER	BUDDY	SITE	Site Name	Latitude	Longitude	1/01/2012	VIS	DIREC	13:30	P-Qs	DEPT	METHO	BLOCK	CODE
6	DIVER	BUDDY	SITE	Site Name	Latitude	Longitude	1/01/2012	VIS	DIREC	13:30	P-Qs	DEPT	METHO	BLOCK	CODE
7	DIVER	BUDDY	SITE	Site Name	Latitude	Longitude	1/01/2012	VIS	DIREC	13:30	P-Qs	DEPT	METHO	BLOCK	CODE
8	DIVER	BUDDY	SITE	Site Name	Latitude	Longitude	1/01/2012	VIS	DIREC	13:30	P-Qs	DEPT	METHO	BLOCK	CODE
9	DIVER	BUDDY	SITE	Site Name	Latitude	Longitude	1/01/2012	VIS	DIREC	13:30	P-Qs	DEPT	METHO	BLOCK	CODE
10	DIVER	BUDDY	SITE	Site Name	Latitude	Longitude	1/01/2012	VIS	DIREC	13:30	P-Qs	DEPT	METHO	BLOCK	CODE
11	DIVER	BUDDY	SITE	Site Name	Latitude	Longitude	1/01/2012	VIS	DIREC	13:30	P-Qs	DEPT	METHO	BLOCK	CODE
12	DIVER	BUDDY	SITE	Site Name	Latitude	Longitude	1/01/2012	VIS	DIREC	13:30	P-Qs	DEPT	METHO	BLOCK	CODE
13	DIVER	BUDDY	SITE	Site Name	Latitude	Longitude	1/01/2012	VIS	DIREC	13:30	P-Qs	DEPT	METHO	BLOCK	CODE
14	DIVER	BUDDY	SITE	Site Name	Latitude	Longitude	1/01/2012	VIS	DIREC	13:30	P-Qs	DEPT	METHO	BLOCK	CODE
15	DIVER	BUDDY	SITE	Site Name	Latitude	Longitude	1/01/2012	VIS	DIREC	13:30	P-Qs	DEPT	METHO	BLOCK	CODE
16	DIVER	BUDDY	SITE	Site Name	Latitude	Longitude	1/01/2012	VIS	DIREC	13:30	P-Qs	DEPT	METHO	BLOCK	CODE

Species	Common name	Total	Inverts	2.5	5	7.5	10	12.5	15	20	25	30	35	40	50	62.5	75	87.5	100	112.5	125	137.5
NOT PRESENT	NOT PRESENT	0	0	1	2	3	4	5	6	8	10	12	14	16	20	25	30	35	40	45	50	55
NOT PRESENT	NOT PRESENT	0	0																			
NOT PRESENT	NOT PRESENT	0	0																			
NOT PRESENT	NOT PRESENT	0	0																			
NOT PRESENT	NOT PRESENT	0	0																			
NOT PRESENT	NOT PRESENT	0	0																			
NOT PRESENT	NOT PRESENT	0	0																			
NOT PRESENT	NOT PRESENT	0	0																			
NOT PRESENT	NOT PRESENT	0	0																			
NOT PRESENT	NOT PRESENT	0	0																			
NOT PRESENT	NOT PRESENT	0	0																			
NOT PRESENT	NOT PRESENT	0	0																			
NOT PRESENT	NOT PRESENT	0	0																			

CHALLENGES AND NEXT STEPS

- Free text metadata for data papers, but also need for controlled vocabularies in the data part ?
- How to choose between Event Core, Occurrence Core, taxonomic Checklist Core ?
- Move existing data at GBIF to “Event Core” ?
- EU-BON IPT, GBIF IPT, Role of the GBIF national nodes ? → Long term visibility of EU BON after the end of the project ?
- New GBIF-EU BON user-guide (final draft) will be available soon
- Next EU BON training courses, Helpdesk support → longer term how to sustain this → together with GBIF, TDWG, Geo-BON; LTER, LifeWatch



GBIF Nodes training 2015 :

<http://community.gbif.org/pg/pages/view/47903/agenda-for-the-gb22-training-event>

Larissa Smirnova