



Marie Demarchi

METHODOLOGICAL GUIDE

DATA PRODUCTION

FOR THE ARBOCARTO V2 TOOL

Marie Demarchi, Mathieu Castets, Renaud Marti, Annelise Tran

Table of Contents

Introduction	3
Data storage for the Arbocarto application.....	4
Formatting fields / files	5
Geographical coordinates	5
Environmental file.....	5
Tables of correspondence of the environmental carrying capacities.....	6
Meteorological survey	7
Location of the weather stations	7
Vegetation.....	8
Density classes in the simulation results files	8
Input data required for the Arbocarto V2 application.....	9
The environmental file	9
The division (geographical limits of the sectors)	9
IRIS division	9
Finer cut	9
The attributes: estimation of the carrying capacity values of the environment (KLfix, KLvar)	12
Land use data	12
Use of other sources of information to refine the estimation of environmental carrying capacity values (KLfix and KLvar).....	16
Vegetation layer	18
The attributes: altitude estimation (Altitude).....	27
Weather data	28
Weather station location file	28
Weather station survey.....	29
APPENDICES	30
Appendix 1 - Methodology for the creation of the environmental file at the fine cut	30
Appendix 2 - Elevation Assignment	33
Appendix 3 - Access to the weather report	34

Introduction

The purpose of this document is to describe the data required for the Arbocarto V2 application, and to supplement the User Manual. It is a question of listing them and detailing for each the mode of acquisition and the preliminary treatments necessary in certain cases, before being able to be integrated into Arbocarto v2. This document also recalls the modes of storage and formatting of input data for the application.

The opening up of data (*Open data*) opens up the field of available data sources that can be integrated into the tool and may make it possible to avoid launching a land use production contract for an entire territory, or only for a land use class as initially defined at the start of the project. The guide will therefore focus more on the methods of integration and construction of the environmental file from different sources, in order to take into account the updates of the dynamics of the territories. If the organization using Arbocarto V2 does not have a dedicated Geographic Information System (GIS) department, the guide can be used to support a contract for the creation of preparatory data for the constitution of environmental files.

The document is organized as follows: the first two parts (Storage and Formatting of fields/files) are reminders on the organization of data and the expected format for the different entries. Indeed, storage and formatting have not changed between the Arbocarto v1 demonstrator and the new v2 version. The third part (Input data needed for the Arbocarto v2 application) describes the available data sources and the pre-processing needed to create the input data files.

Data storage for the Arbocarto application

The file tree consists of a set of subdirectories with free names, **except for those in red** in the *Tablea* below. Subdirectories can be created except for the "red" paths (e.g. CSV/METEO not modifiable)

Directory names must not include spaces, accented letters or punctuation characters.

ARBORESCENCE	Description	Type of file contained in the directory		
<i>*required directory / **recommended directory</i>		text	vector	raster
NEW_METEO_REVEAL **	new weather readings	(.csv) one per weather station		
WORK *				
• CSV *				
– KL	<i>Correspondence tables load capacity</i>	(.csv)		
– METEO *	<i>Compiled weather data history</i>	(.csv) one per weather station		
• NDVI	<i>Vegetation index</i>			X
• SHP *				
– PARCELS	<i>Environmental" file</i>	.txt file accompanying the environmental file	X	
– IRIS				
– NEIGHBORHOOD				
– POINTS_NOIRS	<i>Additional data</i>		X	
– POINTS_VERTS	<i>Location of action areas</i>		X	
– SAUV_ANNUAL	<i>Saving interim annual statements</i>		X	
– STATIONS *	<i>Location of the stations</i>		X	
REFERENCES**	Reference directory (source data)			
RESULTS **	Directory of simulation results			

Tablea u 1: Data storage tree

Formatting fields / files

Geographical coordinates

All geographical data used in the Arbocarto application (environmental files, location of weather stations, black points, green points) must be projected (defined in a projected spatial reference system - the units of the coordinates are expressed in meters), because surface calculations are performed (for the estimation of the carrying capacity of the environment, for the calculation of the effects of control actions). In this case it is:

- of the Lambert 93 projection for metropolitan France;
- of the RGR92 projection for Reunion
- UTM Zone 38 S projection for Mayotte
- UTM Zone 20 N projection for Martinique
- etc.

Environmental file

Vector layers used as environmental files must have at least the first 5 fields of the *Table 2*. The last four are useful for future analysis. The case (upper/lowercase) of the field names must be respected as well as the field type.

Name of the field	type	description
id	String	Identifier of graphic objects
SurfHA	Double float	Area in hectares of graphic objects
Altitude	Integer	Average altitude in meters
KLfix	Integer (10)	Environmental carrying capacity impounded by human action
KLvar	Integer (10)	Environmental carrying capacity impounded by precipitation
CODE_IRIS	String	INSEE identifier of the IRIS block
INSEE_COM	String	INSEE identifier of the municipalities
COM_NAME	String	Name of the municipality
NAME_IRIS	String	Name of the block IRIS

Table 2 Structure of the environmental file

Tables of correspondence of the environmental carrying capacities

The **structure of these csv files MUST NOT be changed** and must match the structure described in the *Table 3*.

Fields	CODE	Libellé	Kltot (nb L / ha)	% gites anthropique	Klfix	Klvar
Description		<i>Land use class or type of 'black spot'</i>	<i>Total carrying capacity (number of larvae per hectare)</i>	<i>% of breeding sites impounded by humans (in the absence of precipitation)</i>	<i>Environmental carrying capacity impounded by human action</i>	<i>Environmental carrying capacity impounded by precipitation</i>

Table 3 Structure of the .csv files of the environmental carrying capacities (KL)

The values can be adapted as knowledge of the field increases. Correspondence tables exist for the different land use classes and for each type of "black spots". The modification of the environmental carrying capacity values is done via the environmental carrying capacity (KL) configuration modules "KL parameters". ARBOCARTO will produce a new csv file.

You can also create a new type of "black point". In this case, in a text editor, simply add a line and fill in at least the "CODE" and "Label" fields. The values of the other fields can be modified via ARBOCARTO (module " *Configuration of the environmental carrying capacity* ").

Delivered file	Example of modification: addition of another type of "black spot"
<pre> 1 CODE;Libellé;Kltot (nb L / ha);% gites anthropique;Klfix;Klvar 2 1;cimetière;5000;0.8;4000;1000 3 2;serres;3000;0.8;2400;600 4 3;ventes plantes;5000;0.8;4000;1000 5 4;maisons abandonnées;5000;0.3;1500;3000 6 5;terrasses à plot;3000;0.3;900;2100 7 6;garage;1000;0.1;100;900 8 7;jardins familiaux;3000;0.8;2400;600 9 8;déchetterie;2000;0.1;200;1800 10 9;camping;2000;0.5;1000;1000 11 </pre>	<pre> 1 CODE;Libellé;Kltot (nb L / ha);% gites anthropique;Klfix;Klvar 2 1;cimetière;5000;0.8;4000;1000 3 2;serres;3000;0.8;2400;600 4 3;ventes plantes;5000;0.8;4000;1000 5 4;maisons abandonnées;5000;0.3;1500;3000 6 5;terrasses à plot;3000;0.3;900;2100 7 6;garage;1000;0.1;100;900 8 7;jardins familiaux;3000;0.8;2400;600 9 8;déchetterie;2000;0.1;200;1800 10 9;camping;2000;0.5;1000;1000 11 10;embarcation; 12 </pre>

Meteorological survey

The weather reports are in csv format and are composed of 5 fields (weather station number, date of daily report, daily precipitation, minimum temperature, maximum temperature). The decimal values are formatted with points as separator.

There is one file per weather station and the file name corresponds to the station number.

Example: Extract of a weather report from the Bordeaux station named 33281001.

POSTE	DATE	RR	TN	TX
33281001	01/01/2014	6.1	8.5	12.8
33281001	02/01/2014	4.4	9.1	14.7

Note: ARBOCARTO automatically aggregates the new readings to the historical readings for each station, located in a directory for example named NEW_METEO_REPORT. The aggregated result of the surveys is stored in the directory WORK/CSV/METEO. This aggregation is done when in the SIMULATION module, the box "New weather data" is checked. Since the model uses daily data, there should be no "gaps" (missing data) in the time series. If this is the case, ARBOCARTO will display a message in the information area at the bottom of the window and the simulation cannot be performed.

How to access the weather records is detailed in the section Weather data.

Location of the weather stations

The geographic file for the location of the stations must be in Shapefile (.shp) format and be named "WeatherStations". It must be in the same projection system as the other data. It must contain at least the five fields described in Table 4. It must be case sensitive.

Name of the field	Description	type
Nom	Name of the weather station	String
Numero	Station identification number	Long integer
Altitude	Altitude expressed in meters	Integer
Ouverture	Date of opening of the station	String
Fermeture	Possible date of closure of the station	String

Table 4 Structure of the weather station location files

How to access the weather records is detailed in the section Weather data.

Vegetation

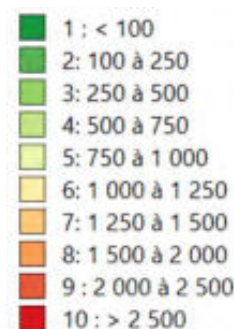
It is a raster data, resulting from satellite image processing. The vegetation layers can be stored in a directory named NDVI. They are in geotiff format, recorded in binary (0: no vegetation, 1: vegetation) and must be defined in the same spatial reference system as all other data (Lambert 93 - datum RGF93 in metropolitan France).

Density classes in the simulation results files

The **shapefile** has a field named **classdens** corresponding to a classification of adult densities per hectare.

The **kml** files present a single piece of information corresponding to the classification of adult densities per hectare. The kml files present directly a legend associated with the 10 proposed classes.

To facilitate the analysis, a QGIS legend file **classdens_arbocartov2_a.qml** was created and delivered with the tool to be used in QGIS software.



For the other stages of mosquito development, a representation by density is also possible thanks to two attributes: the mosquito abundance field at the chosen stage (egg, larva, ...) and the field expressed in hectare (SurfHA).

Input data required for the Arbocarto V2 application

The environmental file

The environmental file characterizes the geographical entities for which the simulations of mosquito densities are performed. This can be an IRIS division¹, or a finer division (e.g. neighborhoods), or a division used by regional health agencies or mosquito control operators for their operational vector control activities (LAV).

This territory grid must be accurate for field use, but also of sufficient size for a coherent modeling of *Aedes albopictus* or *Aedes aegypti* mosquito population dynamics. Following user feedback and tests carried out during the development of Arbocarto V2, it appears that a **minimum size of 3 ha is necessary**.

This led to a review of the "fine" division initially used in the Arbocarto V1 demonstrator. It was therefore necessary to create a new zoning to support the modeling when no territorial division exists to pilot the LAV actions. Thus, except in special cases, it is no longer advisable to directly use the urban land use division. Attribute and geographic information will be used (see the section [Attributes: estimation of environmental carrying capacity values](#)), but the environmental file will no longer follow the boundaries of the polygons making up these land use databases.

The division (geographical limits of the sectors)

IRIS division

These data are available for the whole of France (including the overseas territories). It is an INSEE division of the territory regrouping 2000 inhabitants.

According to the number of inhabitants per commune, this mesh cuts the communes, or uses the administrative limits (=1 single Iris in this case).

We recommend using for the most accurate project: IRIS GE which has been reworked to be used at a scale of 1:10,000 and to be compatible with the geographic contours of other data sources that Arbocarto users may use elsewhere.

To download IRIS GE: <https://geoservices.ign.fr/irisge>

Finer cut

During the ARBOCARTO V2 project, several cuttings with a minimum size of 3ha were produced and are described below in Appendix 1 - Methodology for the creation of the environmental file at the fine cut.

It is up to the users to verify what is most useful and relevant for their territory. We favour one method among others which seems to us to be a good compromise, but the final users are free to create their own division, as long as they follow the minimum size of the meshes and respect the fields and their formatting which are indispensable for the launching of the modelling (cf. previous part).

This new data source supporting the modeling at the "neighborhood" level also presents a new advantage for the updating of the data, which thus becomes less dependent on the evolution of the land cover: it can be updated easily via the Arbocarto v2 tool, without waiting for a complete update of the land cover

¹ IRIS: Îlots Regroupés pour l'Information Statistique, INSEE breakdown

DB: the data sources can be multiple, updated and integrated individually, and also allow easier comparisons between years, as the boundaries do not necessarily evolve, while the attributes will be well updated and taken into account for the modelling.

Thus, we proceed in the same way to update the environmental files whether they represent IRIS blocks or this division to "neighborhoods".

In all cases, to allow future analysis of the results, particularly with INSEE data on the population, whatever the method, the IRIS block boundaries are maintained as well as the identification attributes, and the division is made within each IRIS block.

Comparison of meshes :

- Automatic 3 ha mesh in the form of rectangles or hexagons: easy to implement (either with ArcGIS or QGis), completely independent of land use and occupation
 Negative point: does not reflect the use of the land (more difficult to locate on the ground for example) and consequently generates many unnecessary meshes in natural or agricultural areas and therefore overloads the files produced.
 Improvement element: via a GIS software, select all polygons in the land occupations that are not interesting for LAV, and delete them to generate a new environmental file to be used in the modeling



*In the peri-urban area north-east of Montpellier
 (green IrisGe, orange automatic division rectangle 3 ha)*

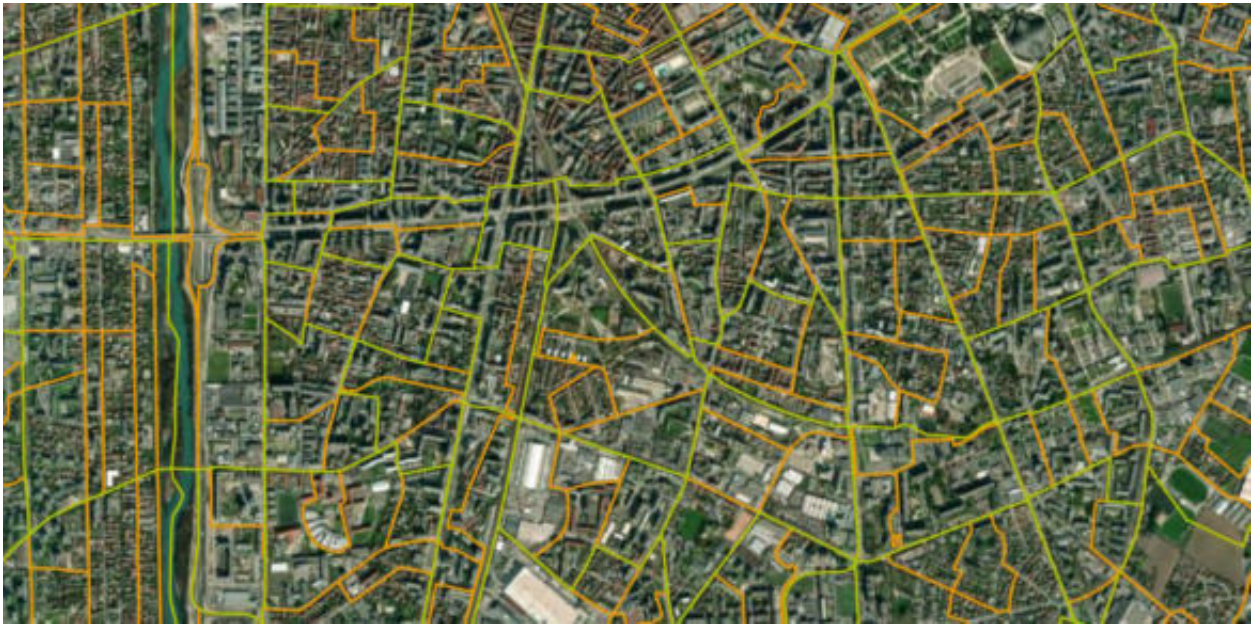


In urban area Montpellier (red IrisGe, black automatic division hexagon 3ha)

- Meshing according to a skeleton of the main roads: a little more complex to produce (see Appendix 1) (ArcGIS required) while waiting for the IGN to make this intermediate step available, also produced internally in their service to constitute the land use databases.
Interest: independence of land use (for a certain number of years) but still reflecting a "human" use of space via the road network, thus allowing a better recognition of the sites by the users.



*In the peri-urban area north-east of Montpellier
(green IrisGe, orange cut 3 ha by roads)*



In the urban center of Grenoble (green IrisGe, orange division 3 ha by roads)

To assign the values of the environmental carrying capacities, one can proceed in the same way, either at the IRIS level or according to this finer division from the CONFIGURATION Module - Creation of a new environmental layer (see User Manual).

This change in fine mesh impacts the interface of the tool, which had to evolve to allow users to build the environmental file by simultaneously taking into account land use classes, vegetation and other important sources of "black spots" such as abandoned houses, or other specifically local data sources, and/or known by field agents.

The attributes: estimation of the carrying capacity values of the environment (KLfix, KLvar)

The CONFIGURATION - Creation of a new environmental file module of the Arbocarto V2 application allows you to update the environmental file created with the chosen mesh with, for example, new parameters for the environmental carrying capacity and/or taking vegetation into account. The initial environmental file will not be overwritten. A new file will be created.

This file can then be used for future simulations of mosquito densities (via the activation of the SIMULATION module).

Land use data

The land use classes of interest to be extracted from the different data sources are the following:

- Mainly residential areas, differentiating the type of habitat
For example, habitat areas in 5 classes:
 - Continuous urban fabrics (1)
 - Discontinuous dense urban fabrics (2)
 - Discontinuous medium density urban fabric buildings (3)
 - Discontinuous low density urban fabric (4)
 - Discontinuous very low density urban fabric and isolated buildings (5)

- Other built-up areas (industrial, commercial, military)
- Other urban spaces (human and mosquito presence):
 - Urban green space
 - Road (because of the presence of street furniture, rainwater network,)
 - Quarry or construction sites

To date, the other occupations of the ground presenting less interest:

Natural areas

Some agricultural areas

It is therefore sufficient to select the different land use classes using GIS software to create files for each type of space that is more or less favorable to mosquitoes (note: it is not necessary to create a file for each land use class: several land use classes can be grouped together in the same file if they have the same environmental carrying capacity characteristics).

Each user is free to use the land cover source of his choice. However, it must be relevant to the scale of data exploitation. This is why we do not recommend the use of Corine Land Cover (although it is free and available for the entire French territory), because it was not designed for use in the neighborhood.

Land use data sources

In the production of Arbocarto v2, the integration of different data sources was analyzed, both in terms of relevance of results and ease/time of implementation.

- Urban Atlas² : Database made available by the European Union through the GMES Copernicus program. It covers European cities with more than 50 000 inhabitants and is updated every 6 years. The 2018 vintage covers 85 French agglomerations (in mainland France, Martinique, Reunion Island). The polygons are aggregated so that the minimum urban area of each object covers 0.25ha. The scale of delimitation of the objects is 1/5 000. The thematic accuracy is more than 85% in urban areas and the graphic accuracy is less than 5m. The database is produced by interpretation of satellite images with a resolution of 2 to 4 m.

It is necessary to create an account (free) to download the datasets.

The database must be reprojected in Lambert 93 for metropolitan France.

- OSO: database produced by the CESBIO on the whole French metropolitan territory and distributed by the CES Theia³. An annual update is produced. It is distributed in raster format and "derived" in vector format. During the Arbocarto V1 project, following integration tests, this data source was discarded, particularly because of its insufficient geographic precision for fine-scale cutting. Within the framework of Arbocarto V2, the integration method being different, we have re-examined the relevance or not of using this data set with regard to its nomenclature.

2 <https://land.copernicus.eu/local/urban-atlas>

3 <https://www.theia-land.fr/product/carte-doccupation-des-sols-de-la-france-metropolitaine/>



Marie Demarchi

- Local urban plans (PLU). Urban zoning is often defined according to the type of habitat. The governmental site "Geoportail de l'urbanisme"⁴ allows you to download urban planning documents at the municipal level.

It should be noted that the pilot sites are also testing databases produced on their territory, the results of which we do not yet have at the time of writing this guide.

Presentation of the results on Montpellier and/or surrounding municipalities according to these 3 data sources

Urban Atlas: the nomenclature is adapted.

Because of the small size of the objects, it is important to integrate all the classes for a better representation of the territory (sufficient carrying capacity of the environment).

The integration is simple (a little longer than in the case of the Arbocarto v1 demonstrator), but allows more flexibility for the update of the data and taking into account the field feedback as indicated previously, it was necessary to enlarge the fine mesh (3 ha minimum)

1 / reprojection to obtain data in the same spatial reference system as the other data (for metropolitan France only, the spatial reference is compatible for the overseas departments)

2 / attribute selection according to the nomenclature chosen for the mosquitoes (which will be included in the .csv files of the environment's carrying capacity)

3 / export in .shp format of each class of the nomenclature. With respect to the carrying capacities of the environment defined in the *configKL_OCS_UAnew.csv* file, 11 files are generated.

OSO: The exploitation method is the same as the one described above: selection of the classes adapted to the nomenclature of this database, then export of each selection in a geographical file. For this database, 4 classes seem relevant:

- | | |
|----|---------------------------------|
| 41 | dense built-up area |
| 42 | diffuse built |
| 43 | industrial and commercial zones |
| 44 | road surface |

However, at this time, we also advise against using this source of information for Arbocarto V2 (see "*Note: data analysis before use*")

Study of the PLU zonings, especially for communes that could not be covered by an urban land use. The PLU zonings make it possible to distinguish between the different building zones.

We have thus retrieved the PLU zonings of two communes in the Montpellier area, using the "Géoportail de l'urbanisme". The urban zonings in the PLU are noted Ux. It should also be noted that the zones concerning isolated buildings in natural areas are called Nh.

- Municipality of the Metropolis of Montpellier :
Prades le Iez (5500 inhabitants) - covered by 2 IRIS
- Rural town of the Community of Communes of the Grand Pic Saint Loup :
Assas (1500 inhabitants) - 1 single IRIS

⁴ <https://www.geoportail-urbanisme.gouv.fr/>

A step of analysis of the regulations (definition of the names of the zonings) and spatial distribution was necessary to establish the correspondence table between the 5 "built" zones of interest for the project and the built zonings of the PLU.

	ASSAS	Prades le Lez
Continuous urban fabrics	UA and UAp	UA1
Discontinuous dense urban fabrics	UD1	UA2,UA3,UA4,2AU1a,2AU1b,2AU2a,2AU2b,2AU3, UC1, UD1
Discontinuous medium density urban fabric	UD2	UC
Discontinuous low density urban fabric		UD2 and UD3
Discontinuous very low density urban fabric and isolated buildings	Nh1 and Nh2	3AU1, 3AU2 and Nh

The following illustrations show the comparative results obtained between simulations performed only with carrying capacities defined on the built classes for Prades-le-Lez (southwest of the area) and Assas (northeast).



In the rural commune of Assas, the results of the Urban Atlas mapping, which only takes into account built-up areas, show that several districts of the commune (low-density housing) are insufficiently informed, which explains why the prediction of mosquito abundance is not carried out for these districts (cf. note above on the use of Urban Atlas, all the land-use classes must be used to characterize a district). In the commune of Prades-le-Lez, where the urbanization is denser, the model runs well for all neighborhoods. However, we observe an overestimation of mosquito abundances for two neighborhoods that correspond to mixed zones (housing and business zones). The finer zoning of the PLU in these areas allows us to take this into account and to obtain more realistic simulations of mosquito populations. The mapping of classes to the 5 "built-up" areas of interest for the project is less automatic.

Note: data analysis before use

Urban Atlas: the nomenclature is adapted and homogeneous. The results on dense urban areas and peri-urban areas are consistent. Nevertheless, this resource is only available for agglomerations with more than 50,000 inhabitants.

OSO land use database: the nomenclature indicates a differentiation among habitat: item 41 Dense urban, 42 Diffuse urban. There are also items 43 for industrial and commercial areas and item 44 for roads. However, in reality, post 41 is only present in a small area of the city of Montpellier. All the rest of the municipality and all the urban areas of the other surrounding municipalities are qualified by item 42. There is therefore ultimately no differentiation of urban habitat within this database.

PLU: there is no standard in the detailed naming of zoning: the same name can define a different typology of habitats. It is necessary to analyze on a case-by-case basis. However, in the absence of a land use database, the PLU is a good alternative.

The OCS GE is a two-dimensional database (land cover and land use) produced and distributed by the IGN. In Martinique, the database benefits from additional classes compared to the metropolitan version (specific vegetation classes and "Residential" use), which are of interest for our project. The OCS GE Martinique database contains 226 combinations of land cover and land use classes. In terms of combinations favorable to the presence of *Aedes aegypti*, we have retained 27 of them. These 27 combinations of land cover - land use were used to establish the correspondence matrix of the carrying capacities of the environment named *KL_ocsge972.csv*. These 27 combinations can be grouped into 7 major land cover-land use types. For these 7 major types, the environmental load capacity matching matrix is named *KL_ocsge972_type.csv*.

Use of other sources of information to refine the estimation of environmental carrying capacity values (KLfix and KLvar)

The load capacity of the environment can be completed via other sources of information (see User Manual, CONFIGURATION Module - Creating a new environmental layer):

OpenStreetMap, BD TOPO from IGN, other locally available sources (e.g. PCRS for networks...) + all the knowledge gathered in the field. We describe here different data sources (non-exhaustive list).

As a reminder, during the ARBOCARTO V1 project, the questionnaire led to the following results:



Openstreetmap⁵ (OSM) is a free collaborative production database (anyone can contribute, there is no certified body to validate the production). However, the complementary themes identified in the framework of the ARBOCARTO project (Table 5) have been verified on the different pilot sites and are consistent with reality.

Interesting theme	OSM Terminology	Available (extraction)
cemetery	"cemetery »	To be extracted from the "landuse_a" layer (polygon)
Allotment gardens	"allotments"	To be extracted from the "landuse_a" layer (polygon)
Sales areas for plants	"florist" and "garden_centre »	To be extracted from the "pois_a" layer (polygon)
Waste Management Center	"recycling" with different spellings of "waste disposal"	To be extracted from the "pois" layer (point)
Campsite	"camp_site"	To be extracted from the "pois_a" layer (polygon)

Table 5 Additional data of the "black points" type

BD TOPO IGN⁶, BD OCSGE IGN⁷

Direct or indirect use of the information, e.g. to extract the urban stormwater network or to extract roads in case of absence of polygonal objects representing roads in an urban land use database. This information can be approximated with the urban road section network of the TOPO DB which is transformed into

5 <https://www.openstreetmap.org>

6 <https://geoservices.ign.fr/bdtopo>

7 <https://geoservices.ign.fr/ocsge>

polygonal objects thanks to the information on the width of the road and/or the number of roads described in the TOPO DB.

Use field knowledge to improve the mosquito's impact range:

For example, in the case of cemeteries or terraces with plots, identified as sites productive of larvae breeding sites: the nuisance may go beyond their strict right-of-way. In this case, using GIS software, it is recommended to enlarge the extent: simply by calculating a buffer zone around it. Once the extent of a cemetery A has been enlarged, create a new layer "cemetery.shp" with all the funeral centers collected for example via OSM and replace cemetery A by the result of the previous treatment. It is then this new file that will be used to assign the carrying capacities of the environment to the environmental file (on the "black spot" line of the cemetery type), in the module CONFIGURATION - Creation of a new environmental layer.

The identification of abandoned houses, private or public spaces built with flat terraces or poorly constructed roof terraces, are neither referenced in the land use databases nor in the files to be extracted from OSM. On the other hand, these places may have been identified by mosquito control operators during home visits, ... in this case, if we wish to improve the environmental file with this local information, we must create the corresponding geographic layer (*see training material - refer to the training material to create the data corresponding to all sites identified as sources of numerous egg-laying sites: abandoned house, junkyard, terraces with blocks, gullies, etc.*).

Vegetation layer

A vegetation layer (binary image in TIFF format in which the pixels have the value "1" if they are vegetation pixels, and "0" otherwise) can be added to modify the carrying capacity of the environment, via the CONFIGURATION - Creation of a new environmental layer module.

Different solutions exist to create or update this layer. In particular, we detail in this section different solutions based on the exploitation of Sentinel-2 images.

Indeed, during the development of the Arbocarto-V1 demonstrator, several sensors with different spatial resolutions had been tested to generate the vegetation layer. The Sentinel-2 satellite images present perfectly satisfactory results for the project. We therefore recommend to use Sentinel-2 images for a first coverage or an update:

- In mainland France: choose a Sentinel2 Level2A archive image, preferably acquired from April to June with the lowest possible snow, flood and cloud cover.
- In the overseas departments: prefer level 3A (monthly synthesis) to obtain a cloudless image.

Solution 1: From Sentinel-2 images, downloaded from THEIA website and processed with QGIS

You must first create a THEIA account⁸, free and accessible to all <https://sso.theia-land.fr/theia/register/register.xhtml>, then retrieve a satellite image from the THEIA catalog <https://catalogue.theia-land.fr> by following the steps below⁹:

1. Position yourself on the geographical area of interest, the map is synchronized with the data in the list;



2. Filter image features:
a) "Theme" tab, select "Reflectance" theme and products;
b) "Satellite" tab, select the "SENTINEL 2" type, and the product level "LEVEL 2A" (mainland France) or "LEVEL 3 A" (monthly synthesis¹⁰) for very cloudy regions (DROM-COM in particular);



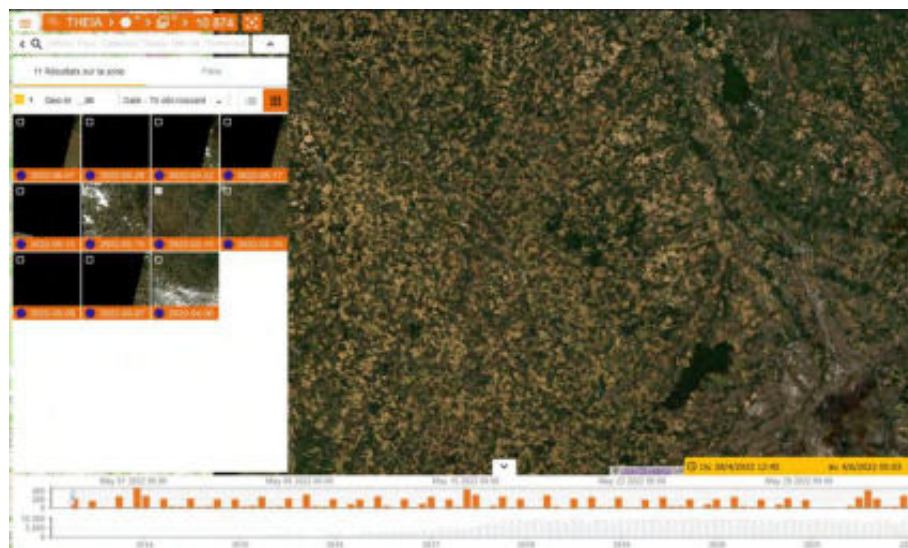
⁸ <https://www.theia-land.fr/comment-avoir-un-compte-theia/>

⁹ For a first visit, follow the tutorial of the online application

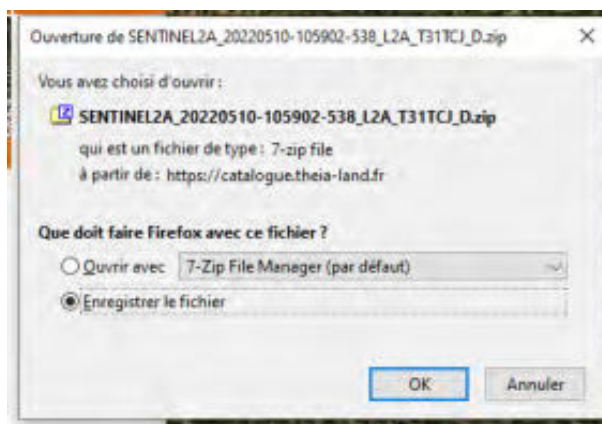
¹⁰ <https://www.theia-land.fr/les-produits-3a-syntheses-mensuelles-de-reflectances-sans-nuages-ou-presque/>

3. Define the temporal filter using the bounds of the orange histogram (at the bottom of the image);

4. View the results: select the desired image, and download it (requires login), by clicking on the dedicated icon;

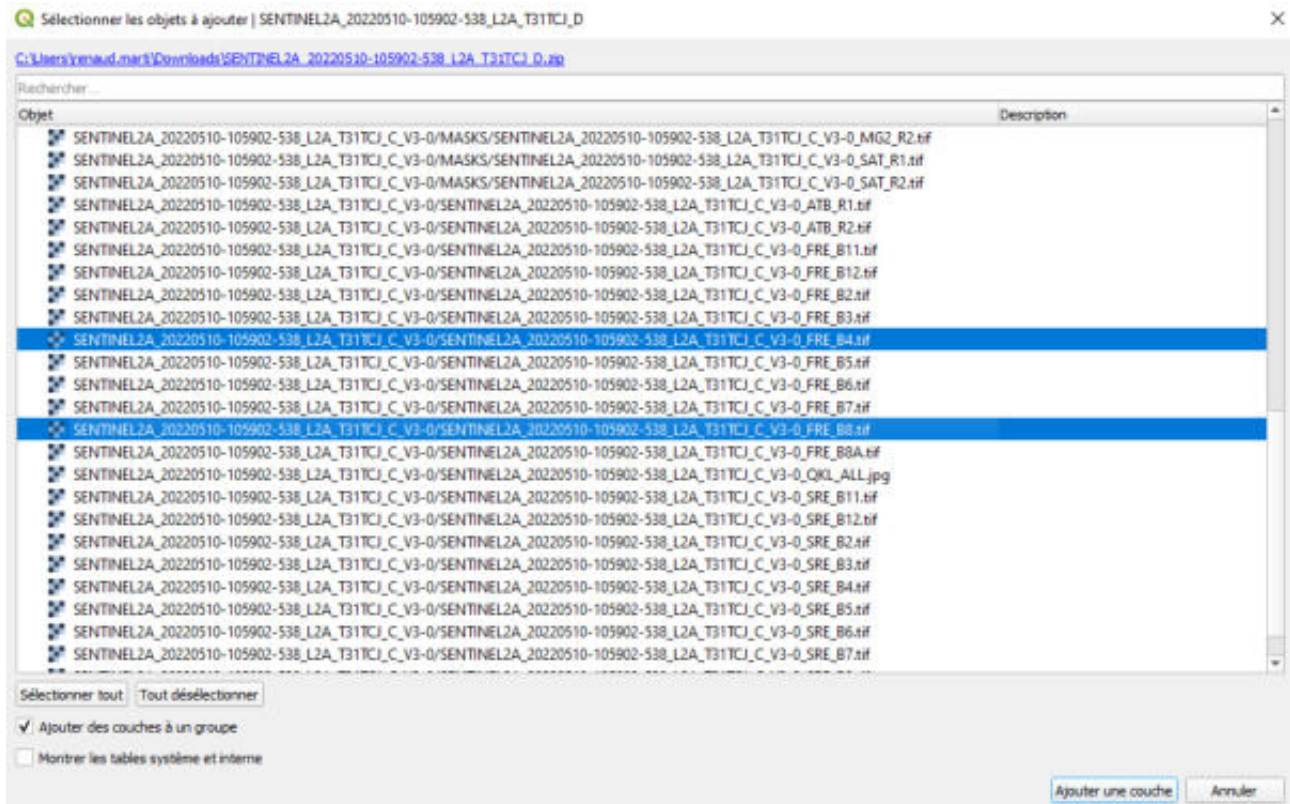


5. The downloaded archive is of the form "SENTINEL2X_20220515-000000-000_L2A_T31TCJ_D.zip" (or L3A instead of L2A depending on the level selected in step 2.b)

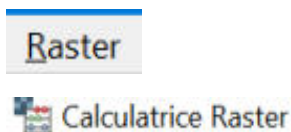


The satellite image archive can now be read directly within a dedicated software, such as QGIS 3.X LTR¹¹, and the NDVI (normalized difference between the red and near infrared channels) can be calculated as follows:

1. Drag and drop the image file (.zip) directly into QGIS, then, in the proposed window, select only the 2 bands ending with "..._FRE_B4.tif" (red R) and "..._FRE_B8.tif" (near infrared PIR) image (screenshot below). Note: the projection system delivered by THEIA is WGS84/UTM ;



2. Once the images are displayed in QGIS, open the Raster/Calculator tool (tab on the top banner of QGIS);

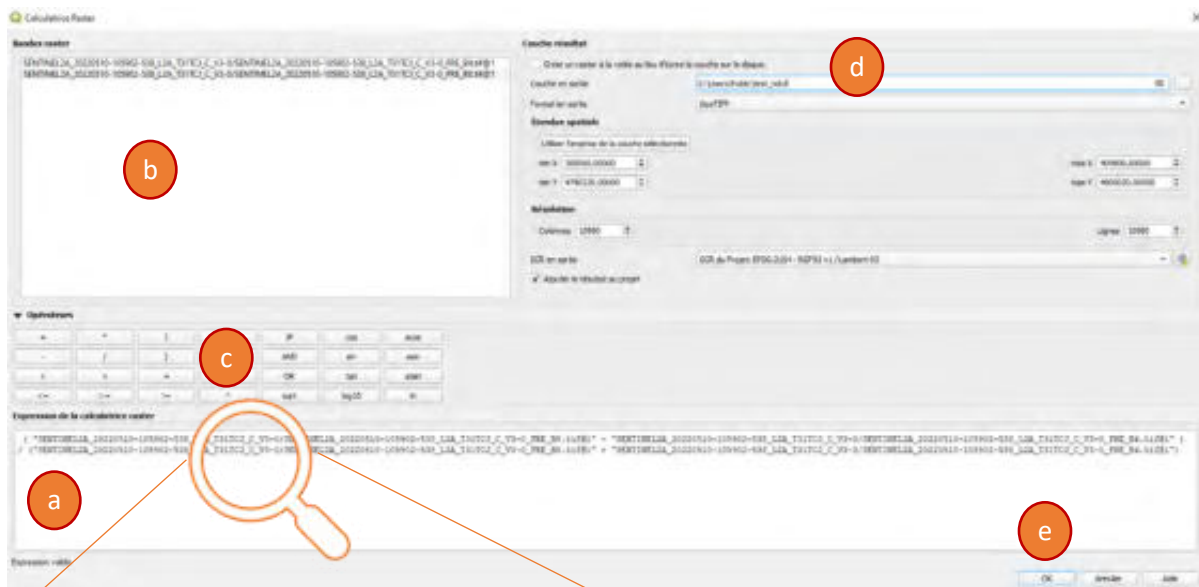


11. <https://www.qgis.org/fr/site/forusers/download.html>

- Enter the NDVI formula **a** in the Raster Calculator ([next image](#)) by double clicking on the layers **b** and clicking once on the operators **c** of the calculator tool interface:

$$\text{NDVI} = (\text{PIR} - \text{R}) / (\text{PIR} + \text{R})$$

or : `("..._FRE_B8.tif@1" - "..._B4.tif@1") / ("..._FRE_B8.tif@1" + "..._B4.tif@1")`



Expression de la calculatrice raster

```
( "SENTINEL2A_20220510-105902-538_L2A_T31TCJ_C_V3-0/SENTINEL2A_20220510-105902-538_L2A_T31TCJ_C_V3-0_FRE_B8.tif@1" -  
"SENTINEL2A_20220510-105902-538_L2A_T31TCJ_C_V3-0/SENTINEL2A_20220510-105902-538_L2A_T31TCJ_C_V3-0_FRE_B4.tif@1" )  
/  
( "SENTINEL2A_20220510-105902-538_L2A_T31TCJ_C_V3-0/SENTINEL2A_20220510-105902-538_L2A_T31TCJ_C_V3-0_FRE_B8.tif@1" +  
"SENTINEL2A_20220510-105902-538_L2A_T31TCJ_C_V3-0/SENTINEL2A_20220510-105902-538_L2A_T31TCJ_C_V3-0_FRE_B4.tif@1" )
```

- Indicate a directory and an output file in the interface **d**, then launch the processing (click on "OK" **e**);

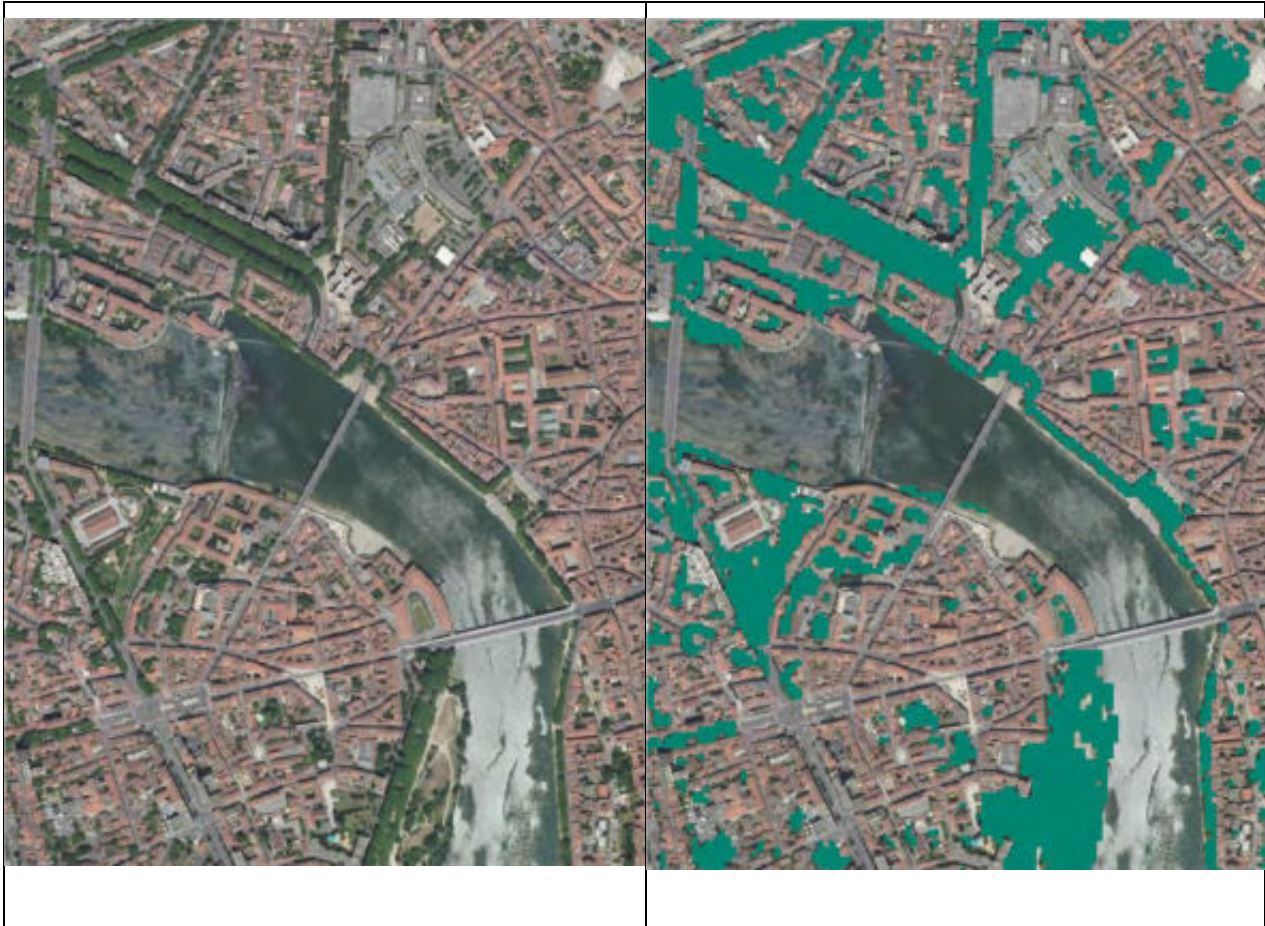
- After processing, the generated raster layer must contain values between [-1 ; 1] ;

- Apply a threshold to the index using the Raster Calculator tool (e.g. threshold value 0.55), to keep only the surfaces corresponding to the vegetation: `"layer_ndvi@1" > 0.55` ([image below](#)) ;

Expression de la calculatrice raster

```
"couche_ndvi@1" > 0.55
```


7. Visually compare the generated raster layer with a satellite image or orthophoto background to adjust the vegetation extraction threshold ([next image](#)). To facilitate the comparison, change the value to 0 in transparency, and assign a color to the value 1 of the band (e.g. green); to change the threshold value, return to step 6 ;



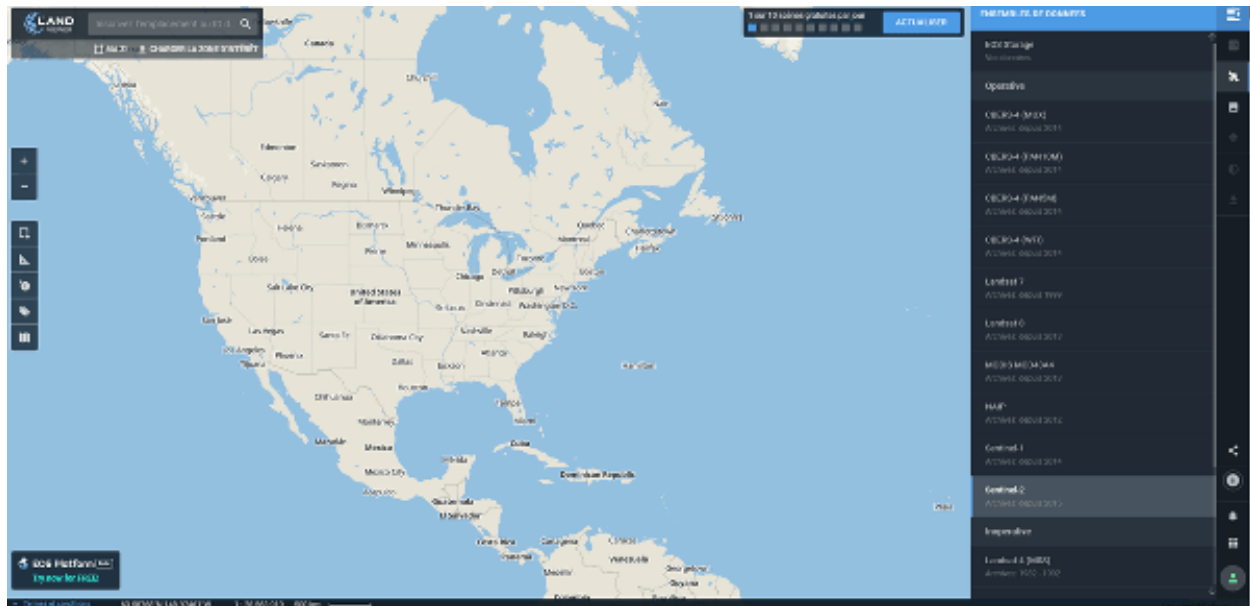
8. Save the thresholded NDVI image in the appropriate projection system (SCR) for further processing (e.g. Lambert 93 in metropolitan France).

Solution 2: Online NDVI production tool - example of the LandViewer platform

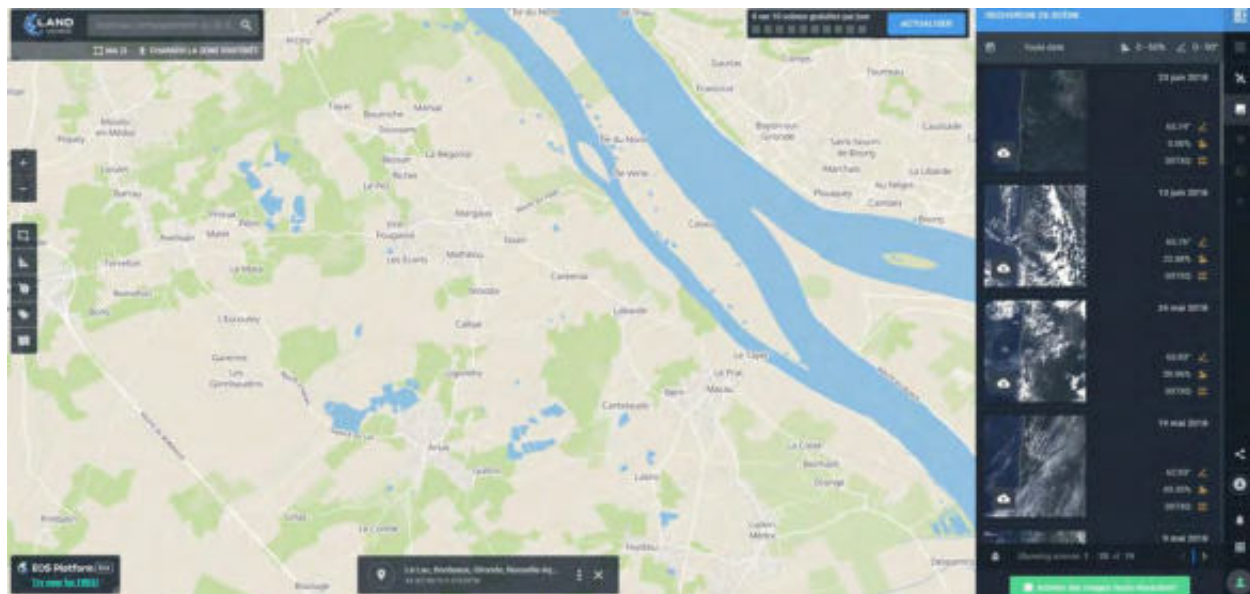
LandViewer (<https://eos.com/landviewer/?lat=43.61070&lng=3.88090&z=11>) is a satellite image analysis platform developed by EOS Inc (Earth Observation System founded in 2015). It allows to visualize for free in a few seconds NDVI images from Sentinel 2 images [source : Extract from <http://www.agrodata-consulting.com/blog/landviewer-viticulture-precision-ndvi/>]

When you arrive on the LandViewer platform :

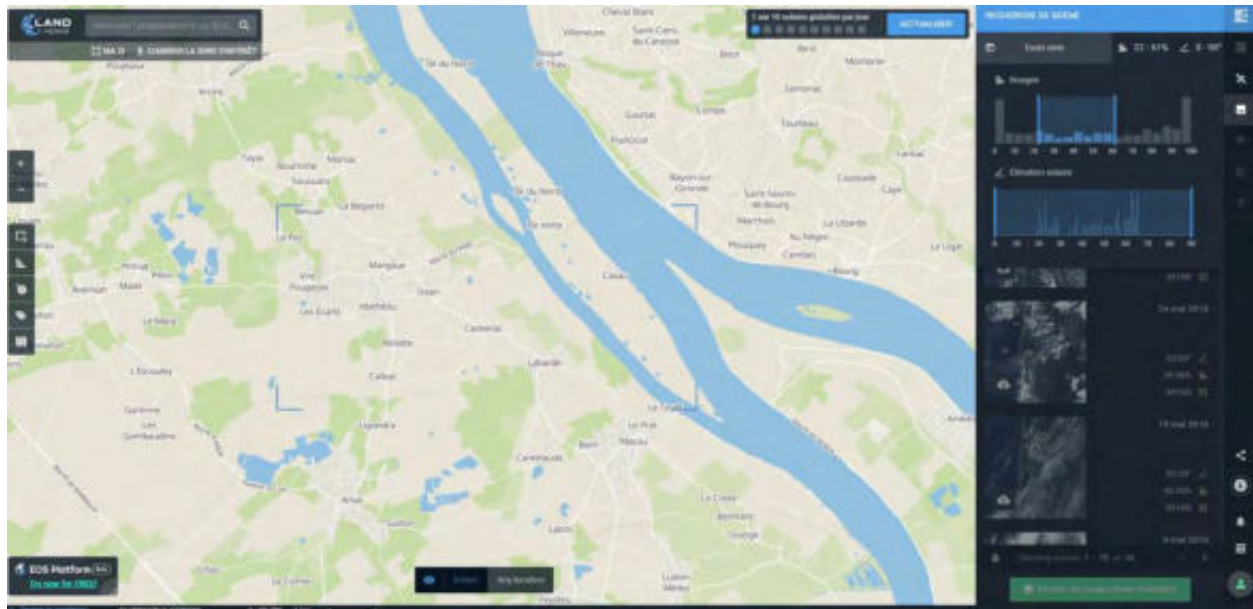
1- Choose the satellite you are interested in, here Sentinel-2.



2 - Then position yourself at the place where your parcels are located: here the Margaux appellation



3 - Select the date you are interested in on the bar on the right in the scene search tab. You can filter the dates according to the cloud rate and the solar elevation level.



4- Then select the NDVI index in the band combination tab. You can also select other indices available on the platform.



5- The downloading of images is possible in geotiff format via the scene download tab.



Once the images are downloaded, you can analyze them on your computer with QGIS: NDVI thresholding, coordinate conversion for use with Arbocarto-V2 (see Solution 1 Steps 6 and 7).

[Solution 3 : By a service provider - example of production by the company Kermap](#)

Kermap service: production for ARS Occitanie Nîmes, Perpignan, Toulouse, via automated processing by deep learning (artificial intelligence - Deep learning).

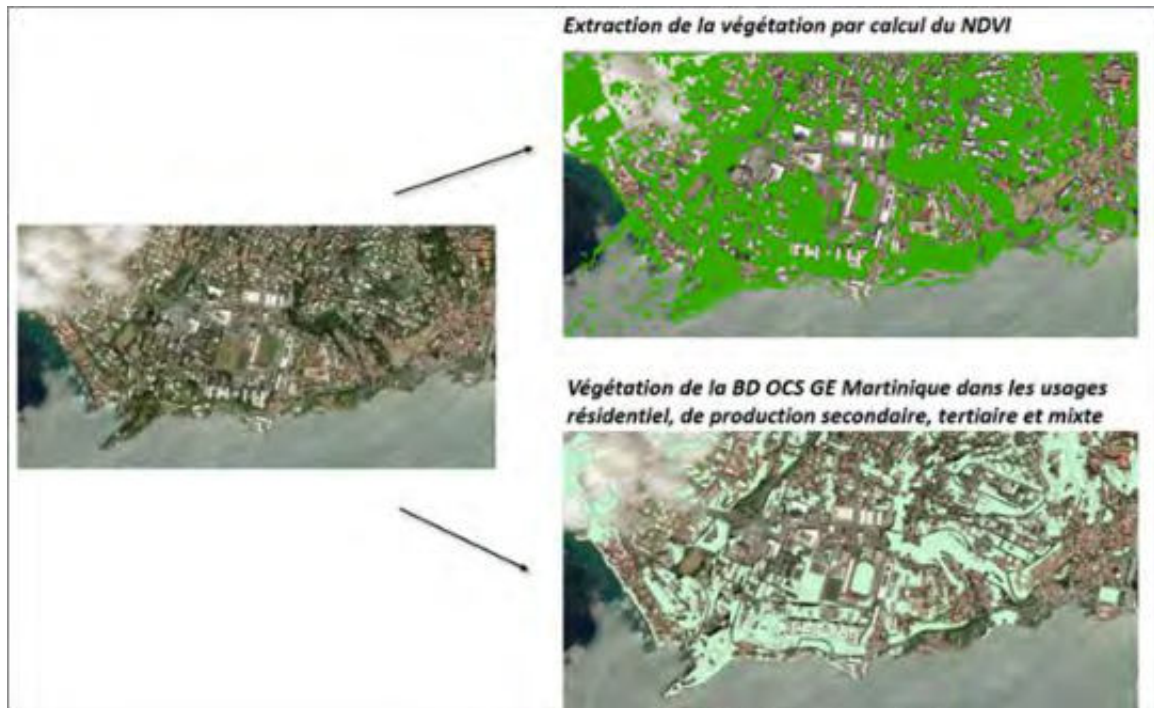
Examples of results for the Dordogne department are available here:
<https://cartographie.mnb24.fr/forets.php>:

https://cartographie.mnb24.fr/mviewer/?x=79665&y=5589899&z=15.777694211641867&l=masque_de_p24*%252CI_enveloppe_foret_1950_s_24_ai*%252CI_enveloppe_foret_1950_s_24_haie*%252CI_enveloppe_foret_1950_s_24_bois*%252Ccommune&lb=ortho-histo&config=apps%2Fmnb_foret_data.xml&mode=u

[Solution 4: Use of a vegetation vector layer distributed by the IGN \(OCS GE\)](#)

The land use data distributed by the IGN (OCS GE) allows the creation of a vegetation layer in vector format, constituted from the polygons of "land cover with vegetation" and whose use is either "Residential", "secondary production", "tertiary production" or "mixed" of the 3 previous uses. Indeed, vegetation in agricultural classes are not of interest for the project. This information can then be rasterized (with QGIS: Menu Raster / Conversion / Rasterize (Vector to Raster)) at a resolution of 5m and reprojected into the same spatial reference system as the other data.

Considering the cloud cover in some regions like Martinique, which is sometimes difficult to fill perfectly with the different Sentinel images (solutions 1 and 2), and the quality and accuracy of the OCS GE, our tests have shown that this layer created by the IGN can be used as input data for Arbocarto-V2. However, the update of this layer depends on the IGN.



A solution to follow

The SCO GreenUrbanSat project (<https://www.spaceclimateobservatory.org/green-urban-sat-grand-nancy>) aims to create and develop, from the exploitation of satellite images, methods and tools to generate a geospatial database of fine description of vegetation in urban areas. For the moment, the data are produced on a pilot area (Nancy) but in the long term the vegetation layers produced could be interesting to use as input for the Arbocarto application.

The attributes: altitude estimation (Altitude)

The altitude is taken into account to allow modulation of the "temperature" parameter collected at the weather stations, at any point of the territory, especially in mountainous areas. The goal is to extract the average altitude for each "parcel" of the environmental file. We used the free SRTM (Shuttle radar topography mission) data source: it is a raster information, originally produced by NASA and USGS, reworked by the CGIAR-CSI (consortium for spatial information of the consultative Group for international agricultural research) downloadable at this address <http://srtm.csi.cgiar.org/srtmdata/>.

The chain of processing required to prepare the data is described in the Appendix 2 - Elevation Assignment.

Weather data

Weather station location file

The information is available on the *Publithèque* site (<https://publitheque.meteo.fr>) of Météo France. It is not necessary to be connected.

Click in the left menu on **Tout sur les stations météo** [All about weather stations]

Fill in the central part of the browser window:

Select location, measured parameters (temperature and precipitation), type 0 1 2 stations

Click on the **Chercher directement les stations météo...** [Search directly for weather stations] button...

Select in the new window the relevant station(s) for the site

Validate the selection

Click on the **Exporter les métadonnées** [Export Metadata] button

Then on **Copier dans le presse papier...** [Copy to Clipboard...]

Open the application *blocnote* (text editor)

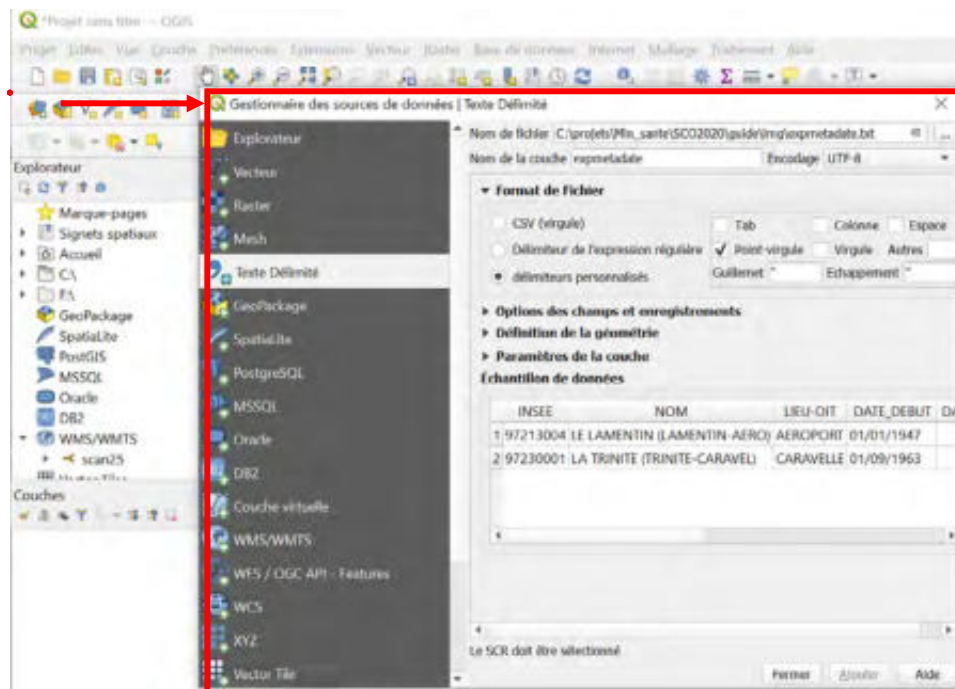
Paste the content and save the file.

Integrate the file in a GIS software

For QGIS, click on **data source manager**

Then select **Delimited Text**

Select the txt file and specify the separator type (semicolon)



Define the SCR (Spatial Reference System.) Meteo France delivers the information with non-projected coordinates (EPSG code 4326)

Click on the **Add** button

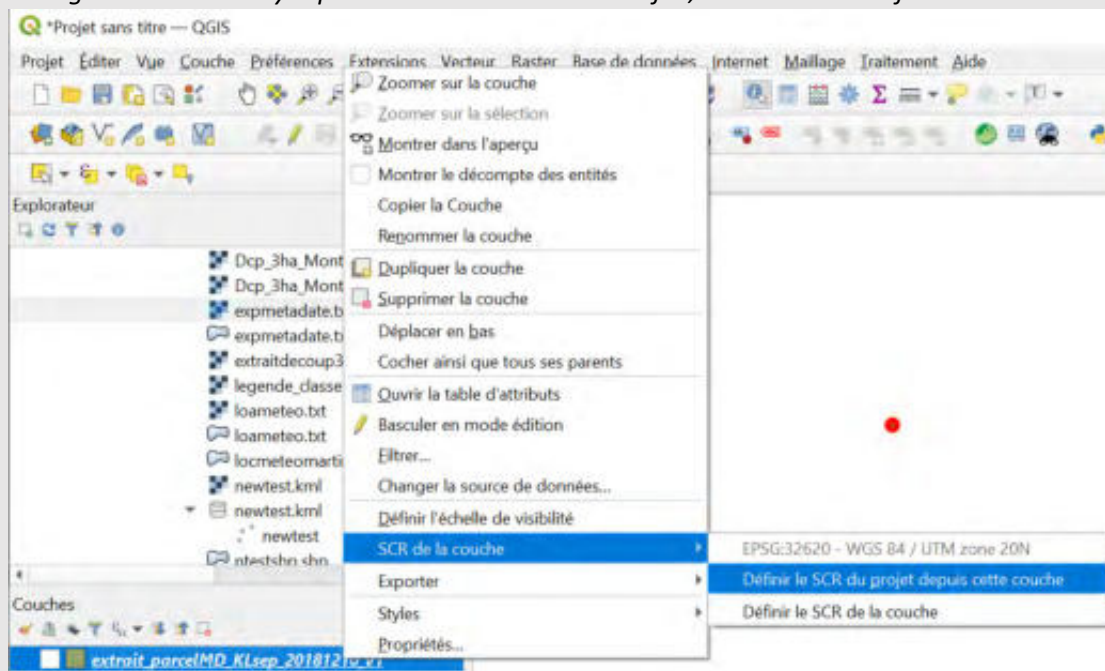
The layer is then integrated into the QGIS project.

Export it in shapefile (.shp) format by projecting it in the same spatial reference system as the environmental file

Tip: to select the right projection, retrieve the one from the environmental file

In the QGIS project, integrate the environmental file

Right click in the Layer panel on the environmental file, and select SCR of the source



By doing this setting before exporting, the projection will then be available in the SCR list.

Weather station survey

Collection after logging on to <http://publitheque.meteo.fr>

The detailed step-by-step instructions are provided in the Appendix 3 - Access to the weather report.

APPENDICES

Appendix 1 - Methodology for the creation of the environmental file at the fine cut

1. Input data

- irisGE
- BDTOPO (*Road* layers, plus possibly Built-up area and *building*)

2. Preliminary steps

- Applied to the **Road** Layer of the *BDTOPO*

Trails will not be considered.

- 1 Keep all sections except those of importance 6 (=trail) :

PROCESSING: SELECTION BY ATTRIBUTE

- 2 Merge all retained sections (*singlepart*)

PROCESSES: create an "ident" field, apply an identical value and MERGE all the road sections based on the value of this field, do not check the option "create multipart entities".

- **irisGE**

Convert the IRIS boundaries online for later use in the **eliminate** function. This will allow the processing to be executed with the perimeter of each IRIS block as a barrier: the IRIS boundaries will thus always be preserved.

We will also use IrisGE with the initial polygonal entities for other treatments.

Later in the description of the treatments, it will be specified IrisGe (lines) or IrisGE (polygons) to identify which type of layer to use.

- **built-up area and buildings (optional)**

This step is not mandatory, if the use of Arbocarto is exclusively for very urban territories. In the case of the study site "Montpellier agrandi" with many rural outlying communes, there is however an interest:

The BD TOPO provides a **built-up area** layer and a **buildings** layer. We found that the **built-up area** layer does not include isolated buildings.

The objective of the following treatments is to be able to identify potentially all inhabited areas.

1. apply on the *building* layer a 50m buffer, with the option to merge the produced buffers
2. aggregate the *built-up area* and *buildings_buffer* layers that have just been produced

3. create a "fus" field and fill the new field with the same value for all polygons
4. merge according to the "merge" field and choose that each polygon built is independent of the others (= do not check the "create multipart entities" option)

Note: for the following treatments to work, remember to deactivate the Z and M dimensions. Indeed, the entities of the BD TOPO layers have intrinsic construction information of height or depth (Z dimension) and of location for the use of dynamic segmentation in a linear reference frame (M).

3. Sequencing of the processes for the creation of the environmental file with fine cutting

Once the pre-processing is done (*Preliminary steps*), the mesh supporting the models can be produced. In order to keep a maximum of inhabited areas, it is necessary to proceed by iteration.

1. Stages of construction of the temporary mesh

T1 Entities to polygons (road, irisGE (polygon) + constructed complete area) > **provisional mesh construction**

T2 Identity with IRISGE > **assignment of IRIS attribute information to the created grid cells**

2. Steps to identify too small meshes, in order to aggregate them together if they simultaneously meet the 3 following conditions: insufficient surface, adjacent and located in the same IRIS block.

T3 Select polygons with an area of less than 3750m².

T4 Create an "id" field and give the value "[CODE_IRIS]_s" where [CODE_IRIS] is the IRIS identifier defined in the CODE_IRIS field

T5 Reverse selection

T6 Assign for all other polygons in the "id" field the internal identification number

T7 Cancel selection

T8 Merge by attribute thanks to the "id" field by specifying that we do not want multipart and by keeping the attributes characterizing the IRIS blocks

3. Final steps for the creation of meshes within IRIS of sufficient area

By the previous steps, the cut contains more meshes greater than 3750 m² than when the provisional mesh was constructed, but there are still small isolated polygons of insufficient size. This time, these small isolated polygons will be "aggregated" to the adjacent polygons of sufficient size, and located in the same IRIS.

T9 Select the result layer (T8) according to the minimum surface of 3750 m².

T10 Eliminate taking into account the IRIS boundary exclusion layer (in line format) so that the removal of small polygons respects the iris boundaries

4. Repeat the previous steps (T3 to T10) for an area of 5000m², then 1ha and 3ha
5. At the end of the processing chain, check the geometry and correct if necessary (in ESRI and OGC mode).

Before using the file produced, in Arbocarto, it is now necessary to add :

- fields (see chapter *Formatting fields/files* and paragraph *Environmental file*)
- the altitude values (see next appendix *Altitude assignment*)

The values of the environmental carrying capacities can be entered directly in a GIS software or via the Arbocarto **configuration** module - **creation of a new environmental layer**.

Appendix 2 - Elevation Assignment

Download the data from <http://srtm.csi.cgiar.org/srtmdata>.

Then perform the following operations:

Zonal statistics between the environmental file produced and the elevation raster layer (*SRTM*) to extract the [average elevation per grid cell](#)

Join on the mesh identifier between the geographic layer and the statistics table.

Calculation on field: Retrieve in the previously created field (or create it now) the average altitude value

It may happen that, depending on the shape of the environmental file, the *zonal statistical* operation does not apply to certain entities. In this case the entities are not informed by the altitude. To obtain the altitude information, it is necessary to perform the following operations:

- Op1 select the uninformed entities,
- Op2 determine the coordinates of the centroid of these entities.
- Op3 Export the attribute table,
- Op4 create a new geographic layer (point) from the coordinates of the centroids (of uninformed entities).
- Op5 Extract for these points, the altitude present in the pixel of the SRTM layer covering the point
- Op6 Join between the obtained file and the environmental file always thanks to the mesh identifier
- Op7 Check that the entities not previously filled in at altitude are still selected, otherwise redo the selection
- Op8 Recover by calculation on the fields the altitude value of the centroid obtained in step Op5 in the field *Altitude*

Delete all joins

The environmental file is now ready to be used in Arbocarto.

Appendix 3 - Access to the weather report

The following elements are taken from the training materials provided in April and May 2022 and the feedback seminar in June 2022.

Formation ARBOCARTO V2 – Avril –Mai 2022

Structuration des données météo

• Relevé météo :

- fichier texte : .csv
- Nom du fichier : numéro de la station
- Paramètres et formatage
 - Relevé quotidien pluie et température

```

POSTE;DATE;RR;TN;TX
97210001;2018-12-23;0.8;26.3;26.3
97210001;2018-12-24;0.6;25.5;25.5
97210001;2018-12-25;3.8;26.8;26.8
97210001;2018-12-26;0;25.7;25.7
97210001;2018-12-27;0.4;26.8;26.8
97210001;2018-12-28;4.8;25.6;25.6
97210001;2018-12-29;6;25.3;25.3
97210001;2018-12-30;6.4;24.6;24.6
97210001;2018-12-31;0;26.6;26.6
    
```

- Séparateur des décimales : point
- Champ date formaté : aaaa-mm-jj

= *spécifications démonstrateur ARBOCARTO V1*

ARBOCARTO V2 autorise également :

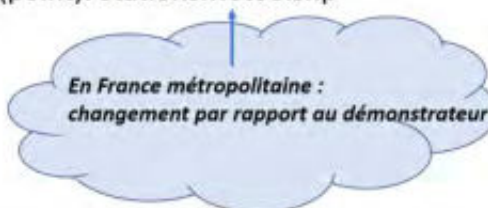
- Séparateur des décimales : virgule
- Champ date formaté : aaaa/mm/jj ou jj-mm-aaaa

Formation ARBOCARTO V2 – Avril –Mai 2022

Structuration des données météo

• Localisation des stations météo

- Fichier géographique (point): StationsMeteo.shp

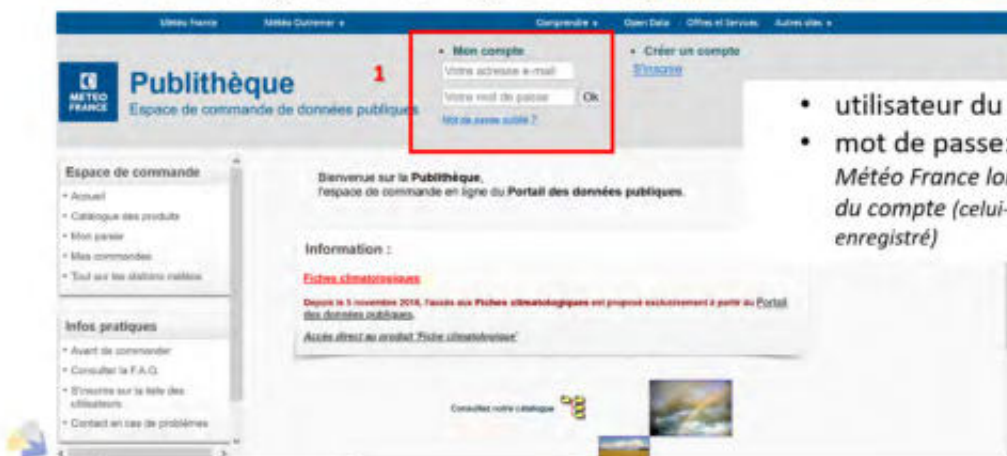


Nom du champ	Description	type
Nom	Libellé de la station météo	Chaîne de caractères
Numero	Numéro identifiant la station	Entier long
Altitude	Altitude exprimée en mètre	Entier
Ouverture	Date d'ouverture de la station	Chaîne de caractères
Fermeture	Date éventuelle de fermeture de la station	Chaîne de caractères

TP Collecte relevé météorologique

Etape 1

Dans un navigateur internet, saisir <http://publitheque.meteo.fr>

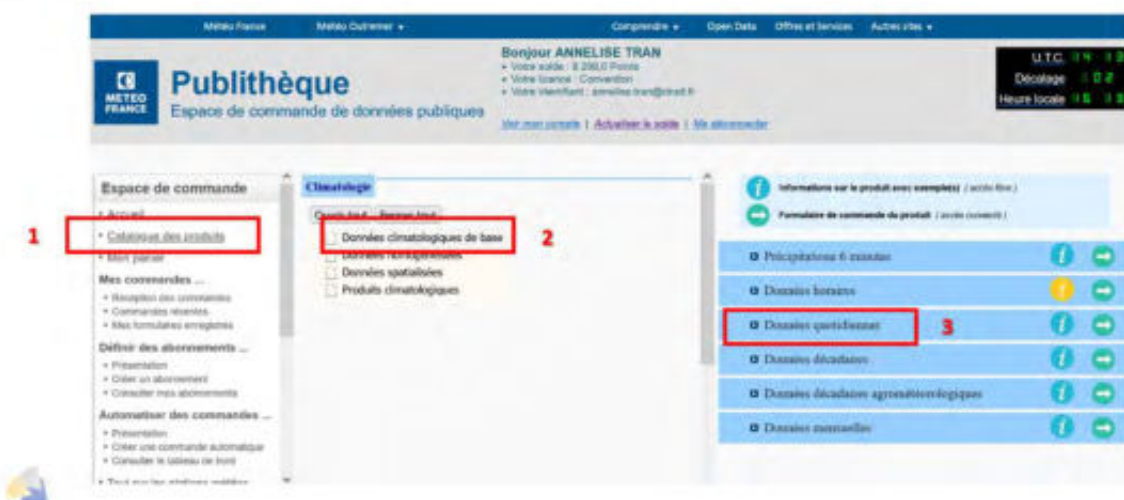


- utilisateur du compte : *mail*
- mot de passe: *celui fourni par Météo France lors de la création du compte (celui-ci peut déjà être enregistré)*

Description for the first collection of weather records

TP Collecte relevé météorologique

Etape 2



To go further on the Météo France website, you must be connected.

TP Collecte relevé météorologique

- 3 - renseigner la partie centrale de la fenêtre (type formulaire)

3.1 mode de livraison :
envoi par mail (pour transmettre fichier)

*l'option initiale "boîte de réception"
= récupérer le fichier sur le site de Météo
France
Choix enregistrable dans « Mes
Préférences ».*

TP Collecte relevé météorologique

DIAPOSITIVE POUR INFORMATION

PAS DE MODIFICATION PENDANT LE TP

DIAPOSITIVE POUR INFORMATION

- *Choix enregistrable dans « Mes Préférences ».*

TP Collecte relevé météorologique

• 3.2 Définir les fichiers constituant le produit

Ne garder cocher que : "Data"

Choix des formats

- ☒ Data
- ☐ Excel "Data"
- ☐ Excel "total"
- ☐ HTML
- ☐ PDF

Personnalisation (concerne principalement les formats Data et Excel)

- Ajouter des identifiants de la station après le N° INSEE :
 - Nom : Pas de nom
 - Altitude : Non
 - Coordonnées : Pas de coordonnées
- Définir le format des séparateurs :
 - Séparateur décimal : Virgule
 - Séparateur de champs : Point-virgule
- Définir le format des dates (avec un format texte uniquement) :
 - Date : aaaa/mm/jj

modifier le type de caractère pour le séparateur des décimales : point

Conservé "séparateur des champs : point-virgule"

Modifier le format des dates de la manière suivante : aaaa/mm/jj

TP Collecte relevé météorologique

• 3.3 référence temporelle

- saisir la période des relevés météo (date de début et date de fin)
- penser à décocher toutes les infos complémentaires concernant l'option :
 - mois et décade :
 - décocher les 3 cases de la colonne "décade"
 - cliquer sur le bouton "rien" permettant de décocher toutes les cases des numéros de mois.
 - semaine : cliquer sur le bouton "rien"

Référence temporelle

Période

Date Début

Date Fin

Format : JJMMAAAA

Précisez éventuellement la référence temporelle en définissant :

- ☐ La sous-période ...
- ☐ Les mois, les décades ...
- ☐ Les semaines ...

Case 1: a single weather station on the study site or decision to obtain a file per station directly as required for Arbocarto

Formation ARBOCARTO V2 – Avril –Mai 2022

TP Collecte relevé météorologique

• 3.4 référence géographique

- sélectionner le département
- sélectionner "précipitation" et "température"
- cocher les cases 0 1 2, pour le type de station
- recherche **la** station :

- via le bouton "chercher directement",
- Ou en saisissant le début du nom de la station

• sélectionner **la** station

- la ligne devient grisée,
- puis cliquer sur le bouton "valider ma sélection"
- la fenêtre se ferme.

- **la station apparaît alors dans le formulaire que l'on est en train de renseigner**



Formation ARBOCARTO V2 – Avril –Mai 2022

Illustrations 3.4 référence géographique



Table des stations du département H, auvergne de 1/1/2022 au 31/12/2022

Paramètres : Paramètres : (PRÉCIPITATION, TEMPÉRATURE)

Type de station : 0, 1, 2

Altitude : 0m - 1000m

Filter :

Numéro	Nom de station	Altitude	Type	Stat	Lat	Long
3420002	BEZERS-VIAS	1015 m	0	0	43.322	3.3525
3415001	BOYTELLE-AEROPORT	1011 m	0	0	43.591	3.0945
3420002	BETE	1075 m	0	0	43.3973	3.0921
3420001	LES PLANS	0849 m	1	0	43.796	3.3481
3417001	NAUVEL LES BEZERS	0140 m	1	0	43.470	3.140
3421001	PEZENAS-TOURDES	0040 m	1	0	43.4378	3.4303
3421001	THAMERIE LE LAC	0000 m	1	0	43.7465	3.8886
3420003	ST ANDRE DE SAVIGNES	0086 m	1	0	43.0841	3.5578
3421001	ST MARTIN DE LONDRES	0214 m	1	0	43.7705	3.7245
3420003	REDABRIER	0073 m	0	0	43.6308	3.148



TP Collecte relevé météorologique

- 3.5 paramètres météorologiques

- **choisir :1..., 19... et 21**

(garder la touche Ctrl du clavier enfoncé pour permettre la sélection multiple de ces 3 paramètres)

- 3.6 cliquer sur le bouton "ajouter au panier "



- 4- **vérifier** la commande puis cliquer sur le lien "**confirmer la commande**"

Un mail est automatiquement envoyé. Il contient des pièces jointes



Renouveler les opérations à partir de l'Etape 2 pour chaque station météo

TP Collecte relevé météorologique

RECUPERATION DU RELEVÉ DEPUIS LE MAIL

- cliquer sur la pièce jointe au mail qui se nomme "xxxx.data" et choisir "**enregistrer sous**" [**ne pas essayer de l'ouvrir**]
 - naviguer dans votre arborescence de fichiers (ex : NOUVEAURELEVÉ)
 - le renommer à l'aide du **numéro_de_station** + ".csv" (numéro de station est à adapter en fonction de la commande choisie)
- se déplacer dans l'explorateur Windows jusqu'au répertoire de stockage du fichier (NOUVEAURELEVÉ)
- vérifier sa présence et supprimer la fin du nom du fichier ".data".
exemple il ne doit être écrit que 30189001.csv
- Un message va apparaître, cocher "oui "



Case 2: several weather stations are available on the study site and we want to retrieve via PubliThèque only one file with the readings of all the stations.

In this case proceed in the same way as for a single station, and use a sub-Excel macro to extract from the global file the x weather report files for the x stations present.

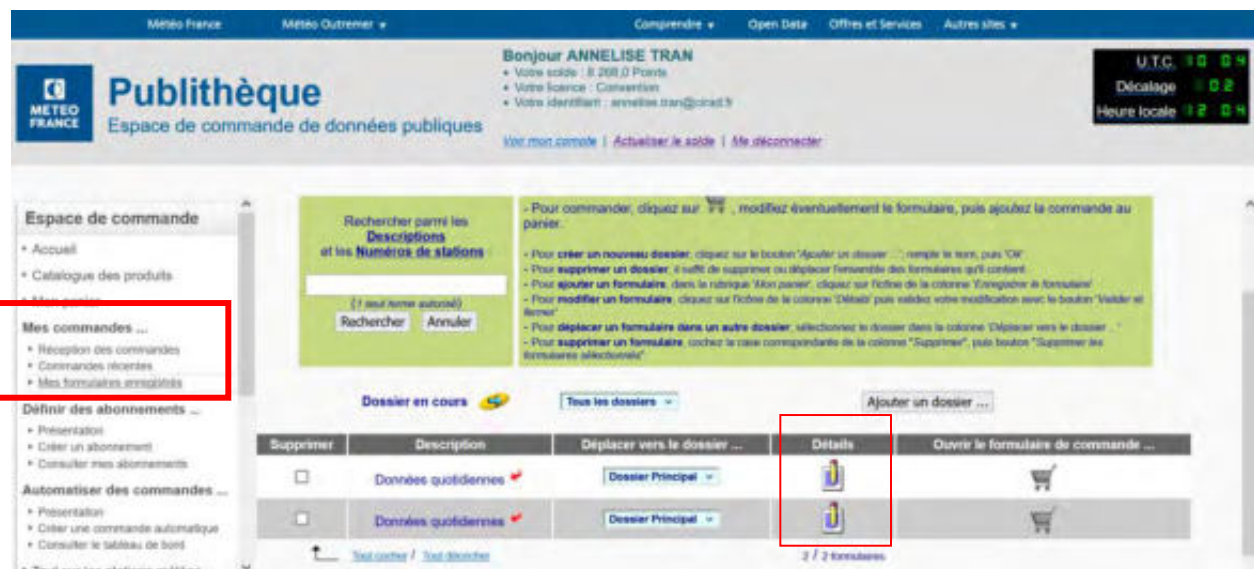
The regional health agency of Reunion Island has made available its macro, which must be adapted to the new site. To speed up processing: only apply it to new data and some functions are no longer necessary for Arbocarto v2 which allows more flexibility in the formatting of fields (see 1st illustration in this appendix)

Alternative to case 2 and also applicable for future collections in case 1

It is possible to automate directly from the *PubliThèque* site of Météo France :

After the first command (=1^{ère} collection of a station):

1. click in the left menu on **Mes commandes...** [My Orders...], then on **Mes formulaires enregistrés** [My Saved Forms]
2. Click on **Détail** [Detail]



All the parameters of the 1st command appear.

The connected user can then choose to modify an element:

Example 1 several stations: the only modification is to change the reference station (see slide about 3.4 geographical reference) and then run the command

Example 2: need to retrieve the readings for another period: this time it is necessary to modify only 3.3 Time reference and only the period of the weather readings (start date and end date) and then to launch the command.

The history of all orders (not deleted by the user) appears in **Mes commandes...** [My Orders...] The following collections are therefore faster.

Note: it is also possible to launch an automaton which, according to the frequency chosen by the user, will automatically carry out future commands. In this case, in the left-hand menu click on **Automatiser des commandes** [Automate commands]

