European
Commission

JRC TECHNICAL REPORTS

Introducing E-Water module

Analysing and processing environmental data in the frame of Mékrou Project

Luigi Cattaneo

Guido Ceccherini

Marco Pastori

Vasileios Markantonis

César Carmona-Moreno

2017

Joint
Research
Centre

Contact information
Name: César Carmona Moreno
Address: Via E.Fermi, 21027 ISPRA (VA)
Email: cesar.carmona-moreno@ec.europa.eu

JRC Science Hub
https://ec.europa.eu/jrc

JRC

Abstract

This report is intended to introduce E-Water, a software interface allowing the most possible wide range of users to easily analyse and process observed and simulated data located in a geographical area and to provide an overall view about its features and purposes in the frame of Mekrou Project. Given the large amount and distinct nature of available datasets, it will gather all of them into a common set (derived from a single database - or a cluster of databases) in order to provide a direct overview of suitable inputs for further operations. Apart from simple data visualisation, it can be used to run modules for processing and transforming input data, providing a graphical interface to visualize and save results in many ways, depending on user's choice.

After a short presentation of the Mékrou project, the *E-Water* module is briefly resumed with an overall analysis of his behaviour and the purposes it is designed for. The following chapters illustrate his inner architecture and data system, plus its interactions with databases. Finally, each of its main functionalities are described through the following sections: hydrology, agriculture, climate, socio-economics, agriculture and water demand optimisation. The last part explains final considerations and possible developments for the future.

Contents

The Mekrou Project

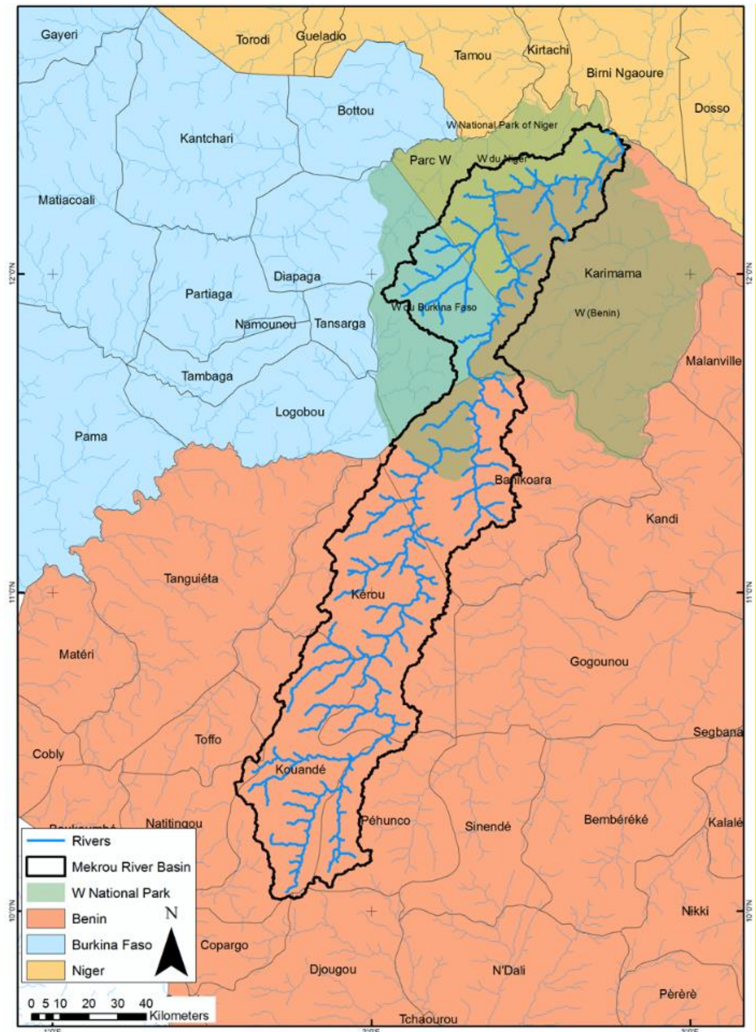


Figure 1: location of the Mékrou River Basin and Area of influence

The name of the project derives by the specific domain of interest where the analysis takes place: the Mékrou river basin, an area of roughly 10000 km^2 shared by the countries of Niger, Burkina Faso and Benin (the latter covering the greatest part of its surface), including the W Park natural reserve, for a total population of 280.000 people. Water scarcity is a common local issue, especially considering that the whole region economy is mainly sustained by agriculture and livestock, but many other aspects are affected too (industry, fishing, mining, energy, personal use, etc.). Promoted by Global Water Partnership of West Africa and Institute for Environment and Sustainability (IES) of EC Joint Research Centre, Mékrou Project has been set in order to safeguard the transboundary basin resources and promote a sustainable development of the whole area under both natural and socio-economic aspects, which can be achieved with an adequate improvement of land use and general economic standards of the population.

1 Main Purposes

E-Water module has been developed in the frame of the Mékrou Basin Project: being primarily conceived within a cooperation program between the European Commission and African institutes (ABN, AGHRYMET), its interface features a particular focus in knowledge sharing on topics related to socio-environmental issues associated with climate change and land use of the local area.

The specific objective of the project is to improve knowledge of decision-makers and the scientific community of Mékrou river basin about the problems and consequences of water resource management and climate change, particularly in view of integrating these issues into sustainable development strategies. Therefore, in order to achieve these goals, it is crucial for both policy makers and researchers to understand climate variability at local-regional-continental scales.

In this context, the software described in this document represents an effort to gather and process climate data available in Mékrou river basin, in order to produce concise and clear information about the variability of key climatic variables, such as precipitation and temperature.

E-Water is set to be addressed to the widest range possible of users, so it has been developed as stand-alone application for Windows systems (the most used in the area) fully relying on open-source software and libraries, so it is possible to use it without purchasing any particular licence. In certain cases, the presence of a GIS interface could help to improve data analysis, but it is not strictly required and, also in this case, free solutions are available.

The overall structure of the module can be represented and divided into three steps: ***input stage***, ***main stage*** and ***output stage***.

1.1 Input stage

Other than common input files selected directly from folder, the module is structured to display and analyse data previously stored into remote or local database(s), which means that most users are not required to execute storage operations, but only to run processes through interface. Since each single input data is conceived to undergo a specific kind of processes, they are arranged into multiple lists, making them available only for determined operations: such organisation helps user to view and manage data in the easiest way possible without any particular SQL database expertise. The import phase can be handled directly by user, when he needs to put new data into the database, but, in case of periodical dataset updates, it could be necessary to add automatic routines.

Other than importing data, it could also be possible to create new datasets by processing and re-aggregating the existing ones. In this way, users will have a wider input choice allowing them to run processes with optimised performances. In fact, there is not a single way to store data inside a database table, but a given format can be chosen whether to save disk space, to grant better performances in terms of speed, or to ensure an output containing the largest possible amount of information.

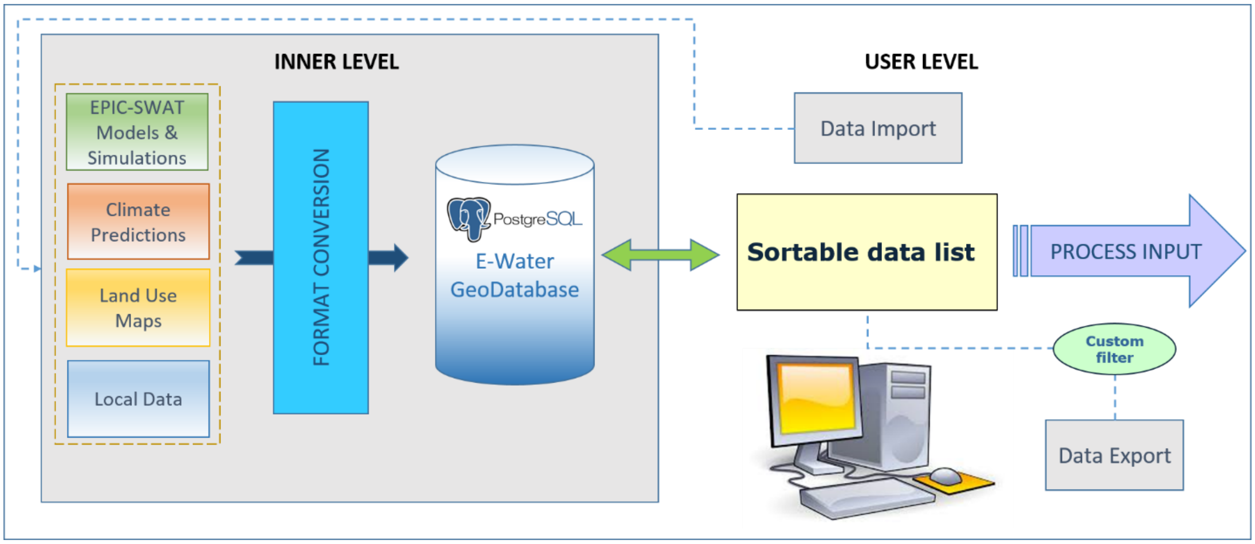


Figure 2: input stage diagram

Also the choice of format is driven by such priorities. Although a unique standard is clearly the best solution for an overall better compatibility among distinct datasets, making them easier to be handled by the module, it is worth considering that a set of multiple standards, each one designed in order to fit properly specific processes, could lead to a better visualisation and improved performances.

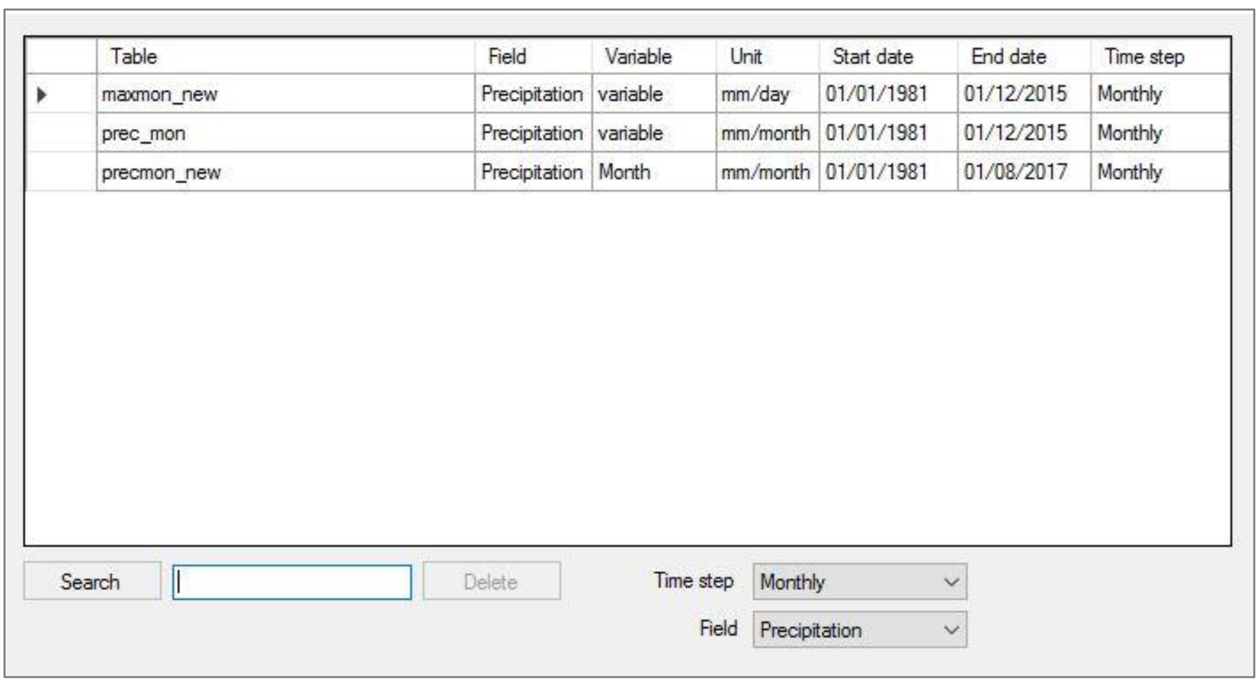


Table	Field	Variable	Unit	Start date	End date	Time step
maxmon_new	Precipitation	variable	mm/day	01/01/1981	01/12/2015	Monthly
prec_mon	Precipitation	variable	mm/month	01/01/1981	01/12/2015	Monthly
precmon_new	Precipitation	Month	mm/month	01/01/1981	01/08/2017	Monthly

Figure 3: sample view of objects list from database

Finally, each data previously imported is displayable from a list directly called from the module interface, making all objects easier to retrieve and analyse. It has to be noticed that such list does not necessarily display datasets as they are imported, but is structured so that user can easily handle each variable for further operations. In this way, one dataset could be visualised as many distinct objects, or multiple input data can be represented in a single row (Figure 3).

1.2 Main stage

From the main menu, it is possible to select the kind of process to carry out, each one related to a distinct tab of the interface: generally, it takes to select the input data along with some parameters in order to specifically define the ongoing process. If needed, operations can be automatized and executed as a streak of multiple runs at once, and input configurations can be loaded from a previous process. Timing is affected by considerable variations: it strictly depends on process and inputs but also relies on the hardware where module is executed.

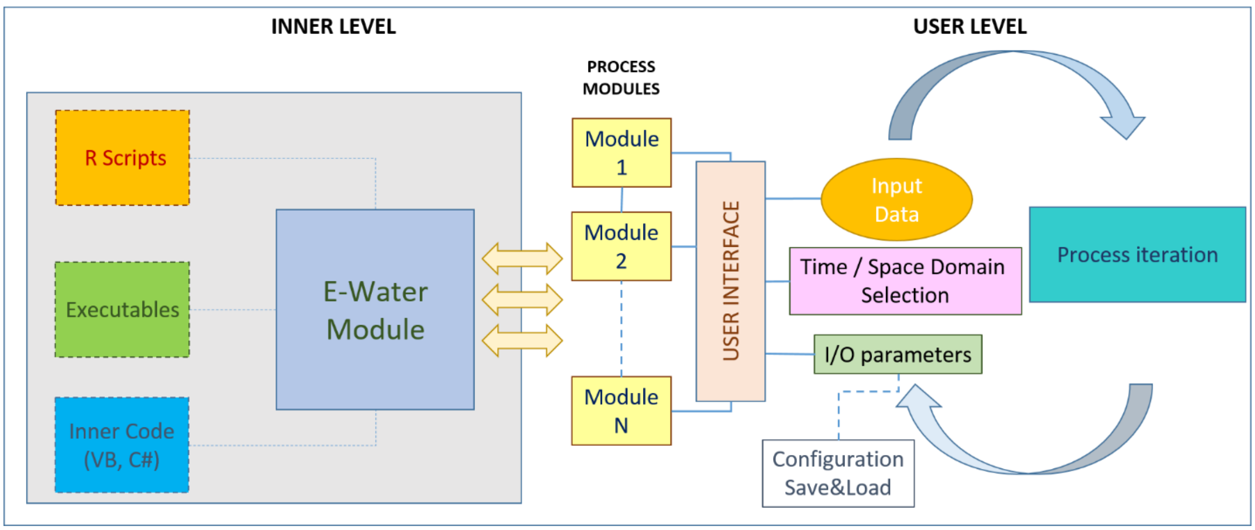


Figure 4: main stage diagram

The module provides a form interface, with different tabs each containing a dedicate process menu in order to grant user an easy and fast access to the functionalities he is more interested in. The available sections are ***Hydrology***, ***Agriculture***, ***Climate***, ***Socio-Economics***, ***Optimisation - Agriculture*** and ***Optimisation – Water Balance*** (Figure 5).

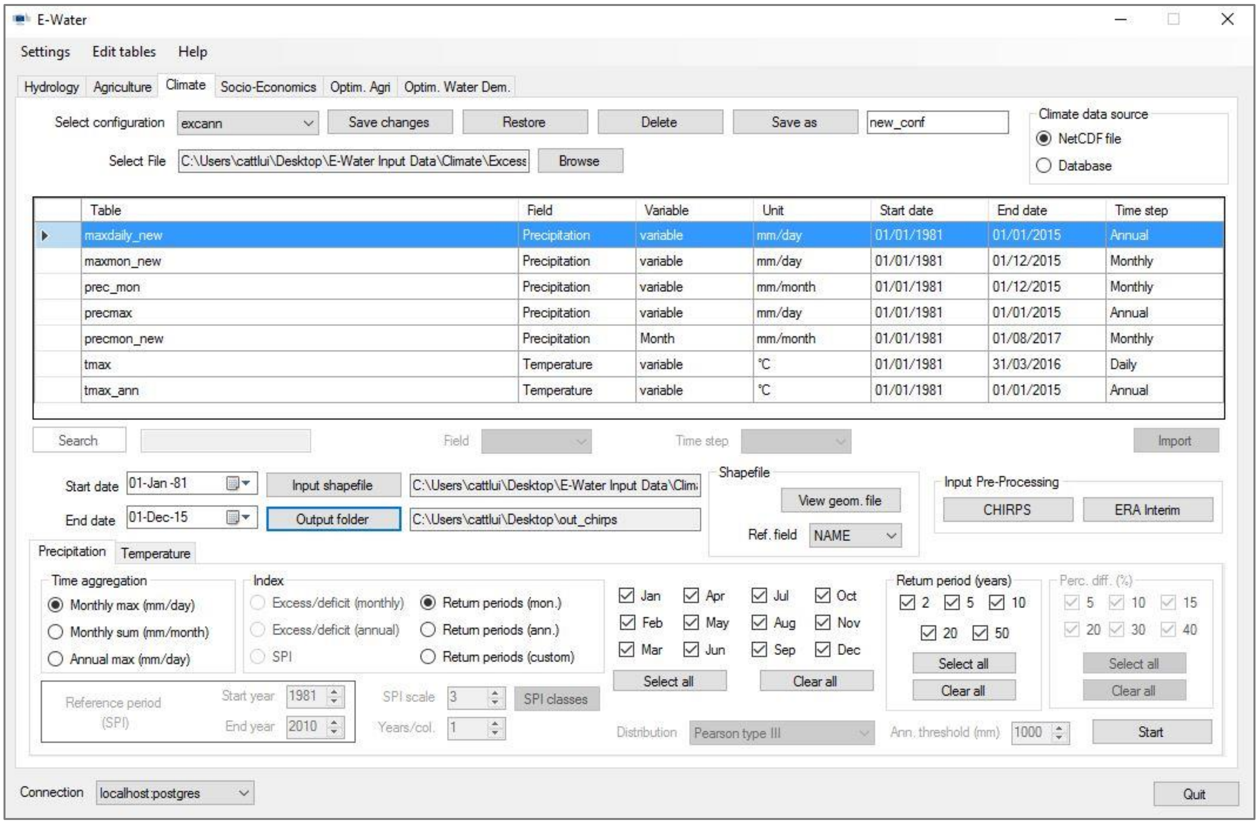


Figure 5: E-Water main interface

1.3 Output stage

Given the different purposes it may be used for, not every process is necessarily associated with a single output format, but user can choose in which way it must be produced by the running module, and results, which are usually represented through a graphic interface, can be saved as images or numerical tables. Furthermore, georeferenced outputs are displayable as maps with the help of a GIS interface. In some cases, a single process generates multiple objects, usually stored inside selected folder or a set of tables named with a common prefix (e.g. climate variability).

Since output data obtained at the end of a given process could be useful for further analysis, it is worth to consider storing it into database in order to have it always available as input. This option may help to avoid redundant operations through similar processes and save a considerable amount of time, but, on the other hand, a large use of this feature may heavily affect disk space availability and require periodical cleaning especially for small storing devices.

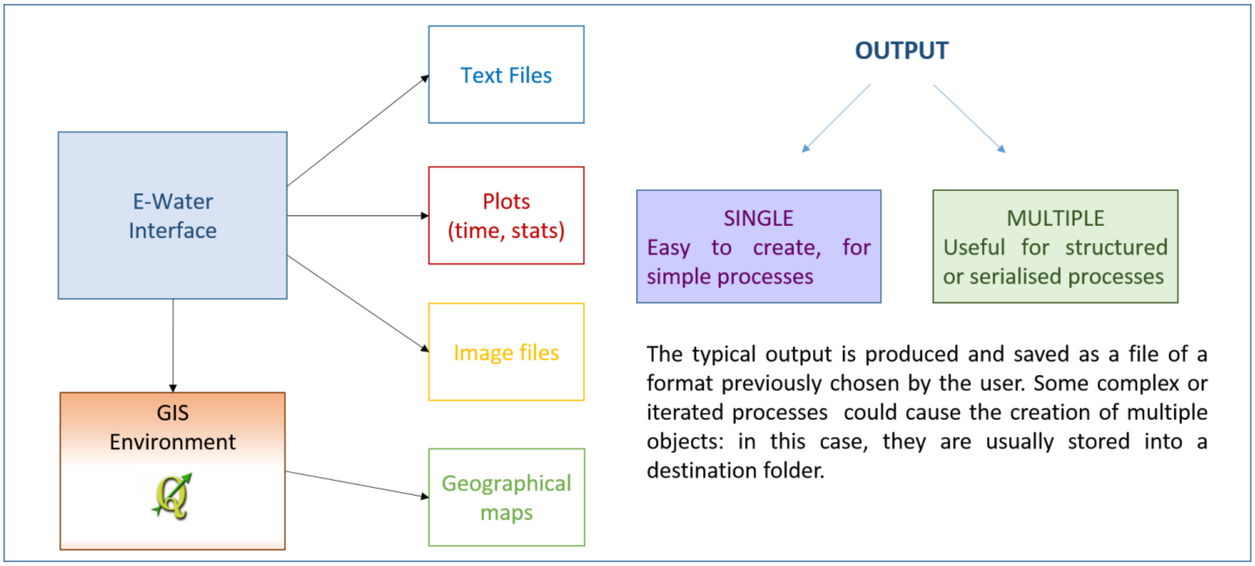


Figure 6: output stage diagram

2 Internal structure

In order to fulfil a wide range of tasks, E-Water has been realised by embedding different environments and languages. Its backbone structure has been compiled with C# language (.NET Framework 4.5) through a Visual Studio 2015 environment, allowing to realise a multiple windows form interface (each form corresponds to a class running all internal processes through its methods), but it also implements:

- R scripts containing statistical and mathematical functions used by its analytical processes
- PostgreSQL interface system allowing to send and retrieve data through command queries
- Other application (*.exe*) for the execution of external models (SWAT and EPIC)

2.1 Class structure

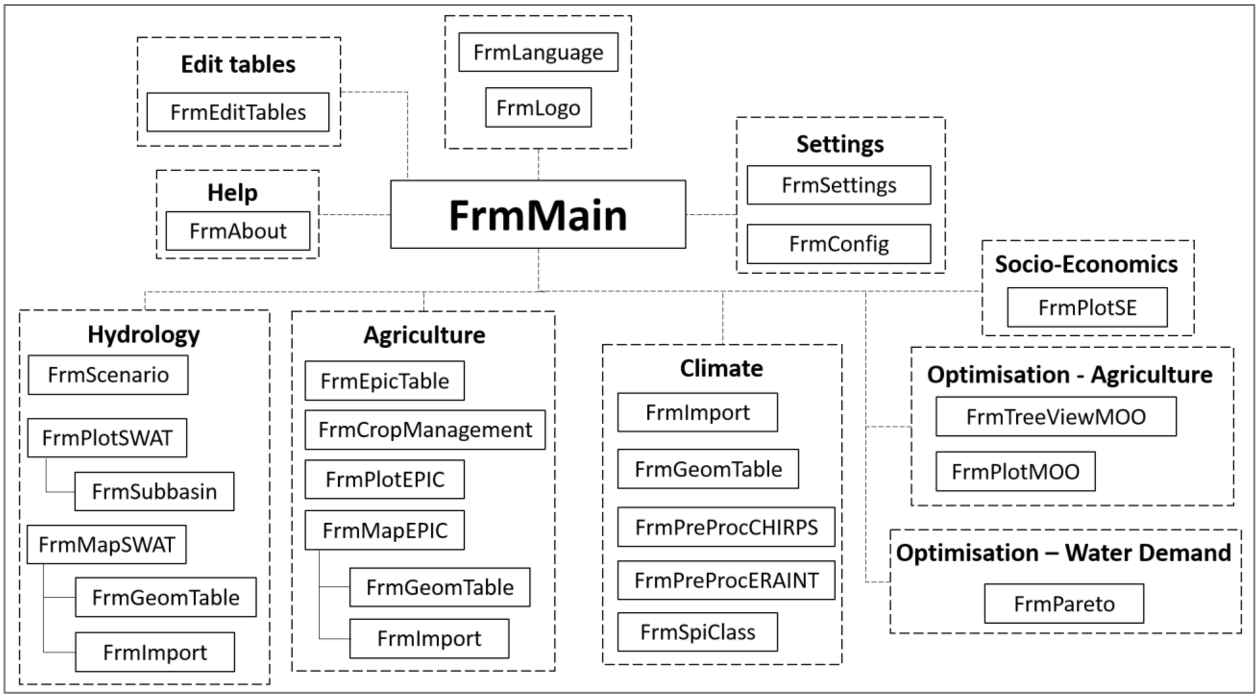


Figure 7: diagram of class hierarchy in E-Water

As already explained, the backbone structure of E-Water consists in a set of windows forms providing a full visual interface and running operations activated through triggered events (e.g. button pushing). **Error! Reference source not found.**shows all forms, while Figure 7 explains their operational diagram.

Following the schema, **FrmMain** is the form related to the main menu interface, from which it is possible to call other ones, each related to specific features. The dashed boxes represent the different sections related to a distinct process, including all involved forms: most of them are directly called from **FrmMain,** but others have a tree-like dependence, meaning they are specifically invoked from another form (e.g. **FrmGeomTable** invoked by **FrmMapSWAT**). The following section will explain the most relevant forms along with their characterising methods.

Name	Description
FrmAbout	Basic information about E-Water frame project
FrmConfig	Configuration folders menu
FrmCropManagement	Crop management tables view from Agriculture section
FrmEditTables	Edit tables menu
FrmEpicTable	Editing menu of input tables lists for Agriculture section
FrmGeomTable	View for tables with geometry (polygon) information
FrmImport	Database import menu for climate data and shapefiles
FrmLanguage	Starting language selection menu
FrmLogo	E-Water banner (loading image)
FrmMain	Main E-Water menu, containing all other sections
FrmMapEPIC	Spatial maps menu for Agriculture processes

FrmMapSWAT	Spatial maps menu for Hydrology processes
FrmPareto	Pareto plot menu from Optimisation – Water demand section
FrmPlot	Generic plot interface menu (parent class of all plot interfaces)
FrmPlotEPIC	Time plots interface menu for Agriculture processes
FrmPlotMOO	Histogram plots interface menu for Optimisation – Agriculture outputs
FrmPlotSE	Histogram plots interface menu for Socio-Economics data
FrmPlotSWAT	Time plots interface menu for Hydrology processes
FrmPreProcCHIRPS	Pre-processing menu for CHIRPS datasets
FrmPreProcERAINT	Pre-processing menu for ERA-Interim datasets
FrmScenario	Editing menu of scenario lists for Hydrology section
FrmSettings	Connection settings menu
FrmSpiClass	SPI classification menu
FrmSubbasin	Subbasins view (Mékrou area)
FrmTreeViewMOO	Region and crop selection menus from Optimisation – Agriculture section

2.1.1 FrmMain

Since it is the starting point of every E-Water processes, this is clearly the biggest class among of the entire application. It contains all general operations executed at start, plus all functions and variables related to every implemented process. Here are some representative functions.

- ***FrmMain_load***: function triggered at launch. First, it checks if it is running for the first time since the installation (*csInitValues.FirstRun* variable) for preliminary operations (language setting, installation of required software), then it loads all saved configurations and state variables, including selected database connection (*conn* variable). Selection menus and checklists are filled according from predefined source files or database tables. A low-level R engine, defined as a global variable, is also initialised along with its libraries and functions, in order to be ready for execution.
- ***initializeComponent_lang***: function called at launch and every time interface language is changed, setting all texts belonging to visual objects (buttons, labels, etc.) to chosen language. It is also included in other forms, despite with different name and implementation (variables to set).
- ***refreshDataGrid***: this function is used to fill and refresh all data grids showing objects from database. Depending on current section (Agriculture, Hydrology, Climate and Socio-Economics) the function retrieves data from related database, also applying filters if present (keywords, data type, etc.). More specifically, each section has its query string, which is used to select tables from database.
- ***saveConfClimate***: this function is called whenever the list of Climate configurations is updated, adding a new one or changing an existing one. Configurations are saved as .xml files into designated folders (see ***Configuration folders***). There are the same functions also for Hydrology (***saveConfSWAT***) and Agriculture (***saveConfEPIC***).
- ***loadConfClimate***: this function is called when a particular Climate configuration is selected, loading all settings from its related file and applying them to all input parameters on the interface . It is also used to restored a configuration in order to

undo all recent changes. There are the same functions also for Hydrology (***loadConfSWAT***) and Agriculture (***loadConfEPIC***).

- ***runSWAT***: when SWAT simulation is launched from Hydrology menu, this function creates a temporary copy of input folder in order to modify it with all parameters and scenario settings chosen by user. Then, it runs SWAT execution file (*E-Water/SWAT/SWAT_64rel.exe*) and finally saves all outputs on designed *PostgreSQL* database. This process also involves R functions contained in *r_swat.r* file.
- ***EpicRun***: when EPIC simulation is launched from Agriculture menu, this function creates a temporary view of input site table in order to get from database only chosen elements and merges it with provided parameters. Then, it launches a custom execution file (*E-Water/EPIC/EpicSrv/EpicRun.exe*), which directly interacts with database by retrieving inputs, processing them with model core executable (*E-Water/EPIC/master/EPIC0810.exe*) and saving results into output tables.
- ***buttonIndStart_Click***: main function gathering all Climate processes. It fetches all input data, regardless of type and aggregation, then process them with a dedicated algorithm related to the index selected by user. Every index evaluation is actually implemented with R functions contained in scripts *ClimateIndices.r* and *spi.r*.
- ***buttonMngRun_Click***: main function for Agriculture Optimisation. It collects data from input scenario (already updated by user) and carries out optimisation according to selected factors and constraints (R functions are involved in the process). After finishing, it automatically calls single report function.
- ***singleReport***: single report evaluation for Agriculture Optimisation. It shows a plot interface where it is possible to analyse output indices.
- ***reportMO***: multi optimisation report for Agriculture Optimisation. It merges optimised data with alternate optimisation set for comparing them and finally produces plot images showing their relationship. Core operations are carried out by R function *report_MOO* from *Load_Scenario.r* script.
- ***waterBalanceSingle***: single scenario report evaluation for Water Demand Optimisation. It collects data from input scenario (already updated by user) and carries out optimisation according to selected factors and water demand reduction constraints (R functions are involved in the process). Core operations are carried out by R function *SingleWater* from *MainFunctionsMOWM.r* script.
- ***waterBalanceMOO***: multi optimisation report for Water Demand Optimisation. It merges user scenario with stress indicator factors and selected proportions of reduction consts, and finally shows all optimisations (solutions) on a Pareto plot. Core operations are carried out by R function *MOOWater* from *MainFunctionsMOWM.r* script.

2.1.2 FrmImport

Depending on the section launching it, this form could serve for multiple tasks. From Climate section, it allows user to select input *NetCDF* files (*.nc*) to be imported into *PostgreSQL* database, while from spatial map menus it is possible to import geometry files (*.shp*, *.tif*). It also provides a view of all currently imported tables. Here are its main functions:

- ***ReadInputfile***: this function reads input file selected by user in order to have a preliminary view of its features and automatically fill some configuration parameters before import process. For NetCDF files, it tries to detect its spatial and temporal fields with name similarity criteria (e.g., *'lon'* and *'lat'* will be considered as coordinates), while for geometry files it determines its coordinates reference system.

- ***importGeo***: implementation of geometry files import process through *PostGIS* commands. Depending on its format – SHP or TIF – the command string is slightly different.
- ***importNC***: NetCDF files import process through *PostGIS* commands. During the process, the last date of climate data is evaluated stating from initial date and time step, both given by the user.
- ***refreshDataGrid***: this function is used to fill and refresh all data imported into database. Depending on the section it has been launched from, it automatically retrieves data from their related database (climate or geometry), also applying filters if present (keywords, data type, etc.).

2.1.3 FrmPlot

This class contains all common features used to plot output data obtained from processes. Despite there are no instances of this class in E-Water, it provides general features and methods to his inherited classes (**FrmPlotEPIC**, **FrmPlotMOO**, **FrmPlotSE**, **FrmPlotSWAT**), which are substantially similar each other. For this reason, most of its functions are actually virtual methods with a different implementation for every distinct inherited plot class. It massively uses objects from **Nplot** library in order to manage graphic plots. Here are its main functions:

- **updatePlot** (virtual): this function is called every time a new plot is drawn on the panel. Given its main role, it interacts with all data variables in order to represent them in the best way possible. Normally, it is executed when plot form is displayed for the first time, and for every further change to visualisation parameters.
- **saveToXLS** (virtual): this function is used to save plotted data into Excel documents. Since the different nature of all output data, the criteria of writing values into the document cells are not the same.
- **printImage**: this function is used to save plotted data as bitmap images. Since it makes essentially a snapshot of the plot panel, taking into account its real dimensions (even in case of resizing), this process is the same for all inherited forms.

2.1.4 FrmMapEPIC / FrmMapSWAT

These two forms have similar interface and behaviour since they both have to produce spatial map from Agriculture (FrmMapEPIC) and Hydrology (FrmMapSWAT) output data. Basically, they link output data produced by model with geometry objects (files or tables) using station identifiers as spatial references in order to create geographical maps. Distinct values are arranged by regions (agriculture) or subbasins (hydrology). Among their features there are:

- **buttonExport_Click**: this function implements the whole process of map creation. First, it retrieves selected indices from *PostgreSQL* database, and then it saves their temporal means on spatial maps using R functions (*exportEPICmap* for agriculture, *exportSWATmap* for hydrology). Geometry sources are treated differently by both functions depending on their format (file or table)

2.1.5 Other classes

E-Water also implements some classes not related to forms, which are stored into the following internal folders:

- ***Catalogs***
- ***Common***
- ***Resource***

Catalogs folder contains general processes dealing connecting interface to database tables. In particular, ***csAdminCatalogDATA.cs*** manages the behaviour of all list interfaces provided by the module (refresh, edit, custom search).

Common folder gathers many types of classes and functions widely used in E-Water.

- ***csInitValues.cs*** contains all functions used by the module during the boot phase (e.g. loading saved data and configurations), along with global variables and functions used throughout the whole interface. For instance, it manages the whole process involving R environment initialisation, loading all packages and scripts, and creating static *Rengine* variable ***engine*** that will be used throughout the whole process. It also shares its private features (mostly folder paths) with all other classes through a set of properties (*get*/*set* logic).
- ***csMiscFunc.cs*** is more oriented on providing multiple sets of functions arranged by specific categories, like math or geospatial functions (static classes). Given its general purpose, this file can be easily adapted and used for other projects.
- ***csPostGRCONN.cs*** identifies *PostgreSQL* connection properties and implements all its core processes during the interactions between software interface and database.
- ***XMLlogFile.cs*** contains internal features related to XML text files and is used mainly for managing internal log files (e.g. read/write).
- ***TriStateTreeView.cs*** specifically implements the tristate tree view object, a variant of original counterpart containing also a third state for its intermediate nodes (other than checked and unchecked).

Resource folder (different from default *Resources*) contains three localised resource sets (English, French and Italian) providing the textual entries used on the user interface, each one managed as a string variable throughout the code.

2.2 Settings (state variables)

All state variables used by the module are stored into his ***Properties.Settings*** class. They are loaded during launch phase and updated when process is terminated. Through *Visual Basic* interface, they can be visualised as a set of strings: the ***Scope*** attribute specifies if user can edit their values from interface (***User***), or only developer have access to them (***Application***) (Table 1).

Name	Description	User access
AppFolder	Name of the application, used for related folders	No
Rfolder	Name of folder containing R scripts	No
intestazMsg	Header of message boxes	No
EpicRunCmd	EPIC executable file (external process)	No
DBfileName	Text file containing connection list	No
packagefilename	Text file containing R packages list	No
EpicInputFileName	XML file containing EPIC input configuration settings	No
SEindFile	Text file containing definitions of socio-economics indices	No
ActiveLanguage	Current active interface language	Yes
EPSGlist	Reference table of coordinate systems	Yes
SWATfile	SWAT executable file	Yes
ClmScenFile	XML file containing SWAT climate scenarios	No
MgtScenFile	XML file containing SWAT management scenarios	No
SWATinputPath	SWAT input folder path	Yes
SWAToutputPath	SWAT output folder path	Yes
ClimateConf	Current input Climate configuration	Yes
HydroConf	Current input Hydrology configuration	Yes
HydroFold	Hydrology configuration folder path	Yes
AgcltConf	Current input Agriculture configuration	Yes
AgcltFold	Agriculture configuration folder path	Yes
ClimateFold	Climate configuration folder path	Yes
ShpFold	Shapefile folder path	Yes
SWATindicesFile	XML file containing localised SWAT output indices list	No
SWATshp	SWAT input shapefile for spatial maps	Yes
MOOscenFolder	Input scenario folder path for Agriculture Optimisation	Yes
CurrSpiClassFile	Selected SPI classification file	Yes
SpiClassStandardFile	Standard SPI classification file	No
SpiClassAgnesFile	Agnew SPI classification file	No
MOOindicesFile	XML file containing Agriculture Optimisation output indices list	No
EPICoutputMapsFolder	EPIC output folder for spatial maps	Yes
EPICshp	EPIC input shapefile for spatial maps	Yes
ActiveConn	Active PostgreSQL database connection	Yes
EPICoutputFilesFolder	Folder path for exported EPIC output tables	Yes
WDscenFolder	Input scenario folder path for Water Demand Optimisation	Yes
CropTableFile	Text file containing localised crop names	No
MessageFile	XML file containing localised message texts	No
PgsqlFolder	PostgreSQL folder path	Yes

Table 1: E-Water settings list

2.3 References

In order to accomplish all its tasks, e-Water also disposes of a set of external *.dll* libraries, whose complete list is available from ***References*** through *Visual Studio* interface (Table 2). With their functions, they allow an interaction with external programs or languages, such as *PostgreSQL* (***Npgsql***), *Microsoft Office* (***Excel.4.5***) and *R* (***RDotNet***). ***Nplot*** has been implemented since it contains some useful functions and types for data plotting.

Library name	Version	Interface with
Excel.4.5	2.1.2	Microsoft Excel documents
ICSharpCode.SharpZipLib	0.85.5	Microsoft Excel documents
Mono.Security	4.0	PostgreSQL database
Npgsql	2.2.3	PostgreSQL database
Nplot	0.9.10	Plot graphics
RDotNet	1.5.4	R environment
RDotNet.NativeLibrary	1.0	R environment

Table 2: E-Water external References list

2.4 Localisation system

In E-Water, language localization is managed in two ways:

Resource files: contained in ***Resource*** folder (different from *Resources*!) of E-Water project folder, they provide all scrips used for interface labels. They can be easily edited through *Visual Studio* interface

Text files: such files contain scrips about miscellaneous information, such as indices descriptions, message texts (when not automatically provided in the system default language) and verbose versions of acronyms. They can be find into default language folders contained into application folder (e.g., ***E-Water/fr*** folder contains French scrips)

2.4.1 Resource files

Name	content
Rx.en.resx	English interface scripts
Rx.fr.resx	French interface scripts
Rx.it.resx	Italian interface scripts

Table 3: content of Resource folder (languages)

Resource files (*.resx*) are normally used to keep state variables, since all changes to their values are memorised and reloaded whenever the application restarts instead of being reinitialised. In this case, all files contain the same set of strings, but each one of them is dedicated to a single language (e.g. English). Since they all contain the same variables – but with different values -, only the one corresponding to active language will be loaded to initialise all label texts of the interface. For instance, *quit_text* string is present in all files, but its value is equal to *"Quit"* in English resource file, and *"Quitter"* in the French one.

The choice of storing all string values into distinct resource files, even if they are never modified by user (as state variables normally are), rather than simply declare conditionally internal variables, has been driven by the fact it guarantees a more comfortable managing system for the programmer. In fact, *Visual Basic* hi-level interface (Figure 8) allows him to update to make corrections in a very short time, without the need of rewriting internal code lines. The names of language files have all the same prefix in order to be correctly read, so in case a new language should be added, its relative file <u>must</u> be named *Rx.<language_initial>.resx*. The language initial is a two-letter acronym univocally associated to a distinct language (e.g. *en*, *fr*, *de*, *it*), whose value identifies an instance of *System.Globalization.CultureInfo* type.

Name	Value
phu_text	sowing PHU
poly_text	Polygons
popincr_text	Population Increasing Rate
popul_text	Population
port_text	Port
prec_text	Precipitation
preproc_text	Input Pre-Processing
pw_text	Password
q3_text	3rd Quartile
quit_text	Quit
rch_text	Reach Output File
reffields_text	Reference fields
refper_text	Reference period
region_text	Region
regions_text	Regions
relhum_text	Relative humidity
rem_text	Remove

Figure 8: string identifiers (left column) with their relative English localised values (right column), as shown on Visual Basic

2.4.2 Text files

Name	Content
crop_table.txt	Crops names
messages.xml	Messages texts

moo_indices.xml	Output indices from Agriculture Optimisation processes
swat_indices.xml	Output indices from Hydrology processes (SWAT)

Table 4: text files contained in each language folder of E-Water

For every available language in *E-Water*, its related folder in the application path (e.g. ***en***, ***fr***, ***it***) contains the same set of following files:

- ***crop_table.txt***: this text file contains all available crops used in many processes involving modelling and analysis of agricultural data. As it easy to see, every row contains a short four-letter acronym in upper case (e.g. ***CORN***) and its localised long name (e.g. *Maize*, in English), separated by a blank space. This file content is primarily used to fill drop down menus for selection purposes, during program start and for any further language change.
- ***messages.xml***: this file contains all texts of message windows not automatically generated by operative system (in that case, active language depends on device localisation). Given its XML format, the file can be seen as a table where field ***TextName*** is the reference called from application (same for all languages) while ***MessageText*** is the localised text used by the interface. In some cases, text may contain some varying terms (numbers, or other variables) and this can be easily notices by the word *'<VALUE>'* contained in the string, which will be replaced directly from the application. If the same string contains more than one variable, they are represented with an increasing integer in order to be distinguished each other (e.g. *'<VALUE_0>'*, *'<VALUE_1>'* and so on). For example, the English string of *table_import* object is *'Table '<VALUE>' imported successfully!'*, but it will be visualised by the user with the correct table name replacing the *'<VALUE>'* slot.
- ***moo_indices.xml***: this table contains all indices obtained in the single scenario report from agriculture optimisation process and displayed in their related plot interface. ***OptionName*** field contains the key identifiers (same for all languages), ***LongName*** corresponds to their localised long names and ***MeasureUnit*** has their measure units.
- ***swat_indices.xml***: this table contains all output indices obtained from agriculture processes (SWAT). ***OptionName*** field contains the key identifiers (same for all languages), ***LongName*** corresponds to their localised definitions (used for plots titles) and ***MeasureUnit*** has their measure units.

2.5 *R* environment

Despite its background structure is developed in *C#*, E-Water mainly implements functions involved in analysis and modelling algorithms in R environment. The choice has been made since the broad diffusion of this open source language, especially for scientific purposes, has guaranteed the availability of many specific libraries providing sets of statistical and mathematical functions well suited for the kind of data to be processed by the module. The *RDotNet.dll* library provides a set of functions creating a full interface bridge between the two languages, making it possible to provide directly inputs, run functions and retrieve outputs. In order to avoid compatibility issues and conflicts, E-Water has its own R environment (version *3.4.1*) with a dedicated set of packages – including newly created

RasterList package, providing optimised statistical operations on *RasterStack* objects (Cordano, 2017 [9]) -, which acts independently from possible versions already installed on user devices.

Folder ***E-Water/R/R scripts*** contains all *R* scripts, each with a set of functions specifically related to a distinct task. R scripts are listed below:

general.r: this file contains all functions which not belonging to a specific process. Their tasks includes package loading at launch (*loadPackages*), reference system conversion (*UTMconv*), import/export interactions with PostgreSQL database (*stack2pgsql*, *pgsql2stack*, *importRaster*, *exportRaster*, etc.) and other general purposes.

ClimateIndices.r: this script file contains the main processes providing climate indices from input precipitation and temperature data. This includes return periods (*MaxDailyReturnPeriod*, *MaxAnnualReturnPeriod*, *MaxMonthlyReturnPeriod*, *ThresholdReturnPeriod*, *ThresholdReturnPeriodtemp*), excess and deficit (AnnualRainDeficit, MonthlyRainDeficit) and heat waves (*HeatWaves*). In addition, it also contains data pre-processing functions (*CreateMaxDaily*, *CreateMaxMonthly*, *CreateSumMonthly*).

rt_funcs.r: this file contains statistical functions related to the evaluation of indices related to return periods, also including excess/deficit and L-moments. There is a strong redundancy, given by the fact that each function has a version for every of its related return periods (in years).

spi.r: file dedicated to SPI evaluation process. *MainSPI* function contains all the steps, taking input files and parameters, and finally creating results images.

r_swat.r: script containing all functions related to SWAT simulations. This includes input parameters editing (*edit_cio*), main process execution (*runSWAT*) and output spatial maps export (*exportSWATtable*, *exportSWATmap*).

Load_Scenario.r, Build_Model.r, Solve_Model.r, ExecutionType.r, SolutionInfo.r: set of scripts with functions related to Agriculture Optimisation. They include a preliminary test of consistency for newly edited input scenario (*ConsistencyTest*), the execution of single scenario (*Solve_Figures*) and multi optimisation report (*Report_MOO*), and sensitivity test (*SensitivityExecution*).

MainFunctionsMOWM.r, Funciones_MOWM.r: set of scripts with functions related to Water Demand Optimisation. They include the execution of single scenario (*SingleWater*) and multi optimisation report (*MOOWater*) along with solutions plotting derived from Pareto diagrams (*PlotStrategyFromPareto*).

3 File system

After installing it, it is possible to locate E-Water files in three different folders:

- **Program folder**: folder containing all application files involved during E-Water execution, plus other private elements whose access is not open to user. Deleted after uninstall.
- **Documents folder**: folder working as local default location for input and output data related to processes not strictly interacting with databases. All its contents are accessible by user and can be modified or deleted.
- **User folder**: folder containing personal user settings, which are not expected to be directly accessed by him, but only through internal processes. Deleted after uninstall.

Folder type	Folder path
Program folder	*C:\Program Files\European Commission\E-Water* (default choice)
Documents folder	*C:\Users\<user name>\Documents\E-Water*
User folder	*C:\Users\<user name>\AppData\Roaming\E-Water* (usually hidden)

Table 5: default paths of E-Water main folders

3.1 Documents folder

More specifically, this folder contains the following subfolders:

- **Configuration**: it contains the default folders for the process configurations of ***Agriculture***, ***Climate*** and ***Hydrology*** sections (they can be changed through **Configuration** menu). Configuration files are in CSV format (*<config name>.csv*) and can be read and modified, though it may result in corrupting the file and making it not selectable from E-Water menu.
- **EPIC**: default folder for files produced by EPIC simulations (**Agriculture** section). **Outputs** folder contains EPIC tables (SUM, ANN, ACY) generated at the end of the process or exported from database as CSVs, whereas images and cartographies obtained by **Spatial Maps** menu are saved into **Maps**.
- **Shapefiles**: folder with default domain shapefiles created during installation. Each distinct shapefile has its own folder (**Mekrou_AOI** and **Mekrou_Subbasin**).
- **SWAT**: default folder for files produced by SWAT simulations (**Hydrology** section). **Inputs** contains the default input folder user for the process (**TxtInOut**). **Outputs** folder contains SWAT tables (HRU, RCH, SUB) generated at the end of the process or exported from database as CSVs, and images and cartographies obtained by **Spatial Maps** menu. **Scenarios** folder contains both **Climate** and **Management** input scenarios available on the interface.

3.2 User folder

This folder is not intended to be publically accessed and its location is normally hidden by the local operating system. Other than keeping private profile information, the presence of this folder is driven by the necessity of dealing with the possibility of multiple users running E-Water on the same device, each one with his dedicated database connections and process configurations. The following subfolders are found here:

- **epic_pgsql_default**: it contains all default tables and functions needed to run EPIC simulations. After a new PostgreSQL connection is added to E-Water connections list, the user can choice to import them into the database (if not, all EPIC functionalities are disabled as long as that connection is active). Alternatively, it is possible to import them afterwards through *Reset EPIC data* function.
- **EPIC**: temporary files used during EPIC simulations are stored here. Most of them are deleted at the end of every single process.
- **SWAT**: it contains two *XML* lists serving as lookup tables defining the overall features of climate (*clm_scenarios.xml*) and management (*mgt_scenarios.xml*) scenarios folder: along each scenario name, they show its folder path and related notes.
- **SWAT_temp**: temporary folder, it is moved into documents as **SWAT** at the very first start of E-Water.
- **Shapefiles**: temporary folder, it is moved into documents as **Shapefiles** at the very first start of E-Water.
- In addition, this folder contains the following files:
- ***dblist.txt***: text file containing general information about all database connections (host, database, port, etc.)
- ***first_run.txt***: this file is present only if E-Water has not yet been executed for the first time. If found, E-Water starts all first-run procedures and then deletes it.
- ***first_run.txt***: this file is present only if E-Water has not yet been executed for the first time. If found, E-Water starts all first-run procedures and then deletes it.
- ***PostgreSQL setup file* [vers. 9.6]**: when E-Water runs for the first time, the software checks if PostgreSQL is present in program list: if not, this setup file is executed. 32bit and 64bit packages have their distinct version. Deleted after first run of E-Water, whether used or not.
- ***PostGIS setup file* [vers. 2.3]**: when E-Water runs for the first time, the software checks if PostGIS libraries are installed: if not, this setup file is executed. 32bit and 64bit packages have their distinct version. Deleted after first run of E-Water, whether used or not.

4 Database characterisation

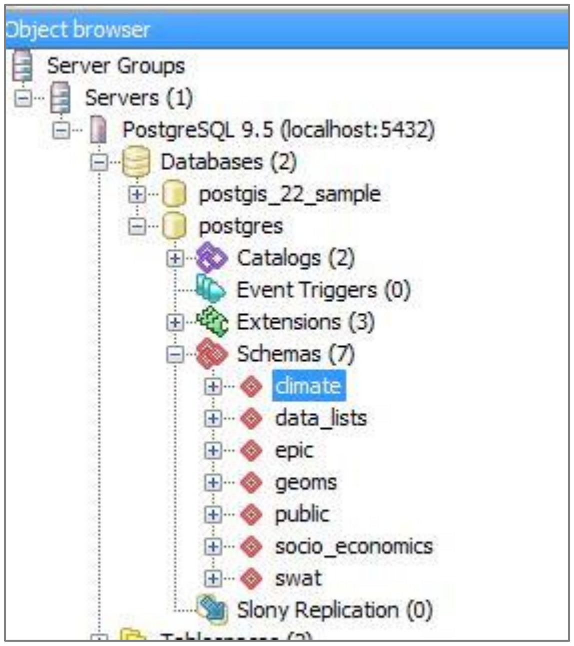


Figure 9: a sample E-Water database view, from PgAdmin interface

PostgreSQL databases are used by E-Water to store and manage a wide range of data as tables arranged in different schemas, depending on their features and utilisation. Normally, tables contain imported data used for process inputs, or data created as results and ready to be displayed and/or exported. Here are listed all schemas involved into E-Water processes.

4.1 Schema "data_lists"

Though it does not contain actual data, it retains a high importance for the external visualisation of the tables into lists. It has only eight tables, which are all created by default when E-Water sets a new connection. Each of them is related to a distinct E-Water interface list and acts as a lookup table between external names (visualised by user) and real tables containing data - in some cases, a group of multiple tables is visualised as s single object.

ewater_climate: this table contains all data imported as input for climate processes (**Climate** menu) and stored into ***"climate"*** schema. Each element in a row includes the name visualised into external list along with the data view and all other features defined by user during import process (Table 6).

Field Name	Description
view_name	Name of the view associated to a single element
label	Name of the element visualised by user
var_type	Variable type (T: temperature, P: precipitation)
var_name	Name of the field inside view containing climate variables
date_start	Data time period, starting date
date_end	Data time period, ending date
time_step	Time step of time period (daily, monthly, annual)
meas_unit	Measure unit of climate variable

Table 6: list of tables from ewater_climate schema

ewater_crop_management: all **crop management** tables used for EPIC simulations (**Agriculture** menu) are listed here. *Current* table is created by default, but others can be furtherly added by user (they are all stored into ***"epic"*** schema). Each element in a row includes the name visualised into external list along with the data view and all other features defined by user during import process (Table 8).

Field Name	Description
table_name	Name of the table associated to a single element
label	Name of the element visualised by user

Table 7: list of tables from ewater_crop_management schema

ewater_epic: this table contains all output tables created at the end of EPIC process (**Agriculture** menu) and stored into ***"epic"*** schema. Since every output is composed by up to four different tables (*LOG*, *SUM*, *ANN*, *ACY*), this list acts as a link between them and the common ID given by user (*label*) (Table 8).

Field Name	Description
label	Name of the element visualised by user
log_name	Name of log table (LOG)
sum_name	Name of SUM table
ann_name	Name of annual hydro table (ANN)
acy_name	Name annual crop yield table (ACY)
mfs_name	Name of monthly table (MFS) [deprecated]
start_year	Starting year
end_year	Ending year

Table 8: list of tables from ewater_epic schema

ewater_epic_meteo: all **daily meteo** tables used for EPIC simulations (**Agriculture** menu) are listed here (they are all stored into ***"epic"*** schema). *Current* table is created by default, but others can be furtherly added by user. Each element in a row includes the name visualised into external list along with the data table and all other features defined by user during import process (Table 9).

Field Name	Description
table_name	Name of the table associated to a single element
label	Name of the element visualised by user

Table 9: list of tables from ewater_epic_meteo schema

ewater_geoms: here are all geometry tables imported from shapefiles (.*shp*) and used as inputs for processes involving geographical domains user (they are all stored into ***"geom"*** schema).

Field Name	Description
label	Name of the element visualised by user
notes	Brief description of the domain (optional)
n_poly	Number of polygons forming the whole domain
table_name	Name of the table associated to a single element

Table 10: list of tables from ewater_geoms schema

ewater_site_crops: all **site crop** tables used for EPIC simulations (**Agriculture** menu) are listed here. *Current* table is created by default, but others can be furtherly added by user (they are all stored into ***"epic"*** schema) (Table 10).

Field Name	Description
table_name	Name of the table associated to a single element
label	Name of the element visualised by user

Table 11: list of tables from ewater_site_crops schema

ewater_socioecon: this table contains all tables imported from input files of socio-economic data (**Socio-Economics** menu) and stored into ***"socio_economics"*** schema (Table 12).

Field Name	Description
table_name	Name of the table associated to a single element
label	Name of the element visualised by user
n_countries	Number of countries in the table
n_districts	Number of districts (regions) in the table
n_villages	Number of villages in the table
n_samples	Number of total samples

Table 12: list of tables from ewater_socioecon schema

ewater_swat: this table contains all output tables created at the end of SWAT process (**Hydrology** menu) and stored into ***"swat"*** schema. Since every output is composed by three different tables (*RCH*, *SUB*, *HRU*), this list acts as a link between them and the common ID given by user (*label*) (Table 13).

Field Name	Description
label	Name of the element visualised by user
rch_name	Name of reach table (RCH)
sub_name	Name of subbasin table (SUB)
hru_name	Name of HRU table (HRU)
start_year	Starting year of simulation
n_years	Length of simulation (years)
n_hru	Number of HRU in the table
n_sub	Number of subbasins in the table
time_steps	Simulation time step (daily, monthly, annual)

Table 13: list of tables from ewater_swat schema

4.2 Schema "climate"

This schema contains all climate input tables imported through **Import** menu from **Climate** section. Since climate data consist on time series distributed on a point grid, each distinct value needs to be identified by a distinct time and position, the final view is a result obtained by joining three kind of tables:

Data table (*<view_name>_data*): it contains all the climate values for every geographical point and time instant, each one associated to a string containing a unique time-space identification (*postime*)[n° rows = n° dates * n° points]

Time table (*<view_name>_time*): it contains all instants covered by the whole interval, from start to end date, each one associated to a distinct ID. Regardless of time step, the format is always *'yyyy-mm-dd hh:MM:ss'*. [n° rows = n° dates]

Position table (*<view_name>_pos*): it contains all grid points occupied by climate data, which are represented as geometry strings. Also in this case there is an ID field associated to each position. [n° rows = n° points]

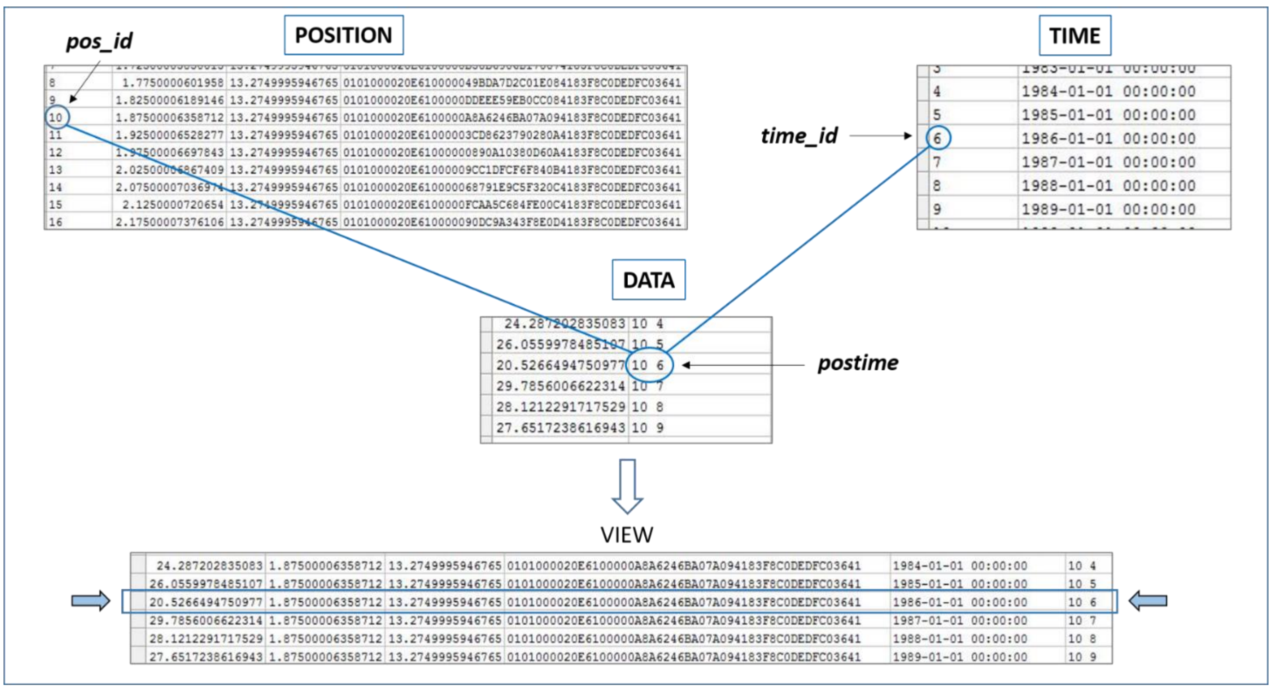


Figure 10: joining of climate tables and displaying of final view

During the join phase (Figure 10), each piece of data is associated to a *postime* index "*<pos ID>_<time ID>*" which is directly related to the corresponding position and date, respecty marked with "*<pos ID>*" and "*<time ID>*". In this way, the obtained view shows all information and is ready to be processed by E-Water – or checked by user, if needed. This configuration is used in order to cut redundancies and save disk space, rather than simply have one single climate table.

4.3 Schema "epic"

This schema contains all elements involved in the processes of **Agriculture** section. First, it is possible to find here all tables used as inputs during the execution of EPIC model, which comprehends:

- **Crop management** tables
- **Site crop** tables
- **Daily meteo** tables
- Default tables, used for every process

Except for default tables (Table 14), which are always used no matter the configuration (they contain information such as regions, Site IDs, crops and climate features), there is more than one table available for all the other kinds, and user can select which one to use before any process run, or even import and delete them from database.

In addition, all output tables created at the end of every process are saved here. More specifically, each output is associated to a set of tables, differing by their suffix: log tables (*<output_name>epiclog*), average annual summary tables (*<output_name>epicsum*), annual water summary tables (*<output_name>epicann*) and, optionally, annual crop yield tables (*<output_name>epicacy*). It is important to notice that all tables composing a distinct output set are intended as a single entity and should be treated – and deleted - as a whole.

Finally, this schema contains two functions, used in the process:

- ***delete_output***
- ***spmanageouttablesv0509***

Table Name	Description
crop_management	Current crop management table
meteo_daily	Current daily meteo table
meteo_monthly	Default table
meteo_stations	Default table
site	Default table
site_crops	Current site crop table
site_fertilization	Default table
site_id_nuts_names	Default table, containing point positions and names (regions)
site_irrigation	Default table
site_soil	Default table

Table 14: EPIC environment tables from "epic" schema

4.4 Schema "geoms"

In this schema, it is possible to import all shapefiles (*.shp* or *.tif* files) related to any region of interest where to focus on **Hydrology** and **Agriculture** displaying processes (spatial maps). The structure of obtained tables is strictly related to the original features of each source files, so they may differ each other in terms of columns and data types. However, they all share the *geom* field, which contains all needed information about geometry and reference system (SRID).

4.5 Schema "socio_economics"

This schema contains all tables imported through **Import** menu from **Socio-Economics** section. Every list contains a set of custom indices whose values are related to a single respondent per row identified by a unique ID and generic addressing info (*country*, *district*, *village*), as displayed on a view obtained by joining the two following tables:

Data table (*<view_name>_data*): it contains all the indices values for every respondent, in addition to his identifier (*id*)

Reference table (*<view_name>_ref*): it contains all geographic information (*country*, *district*, *village* fields) about respondents, plus their *id*.

As it is possible to see, the field *id* is in both tables and is used to join them into the view, as seen in Figure 11.

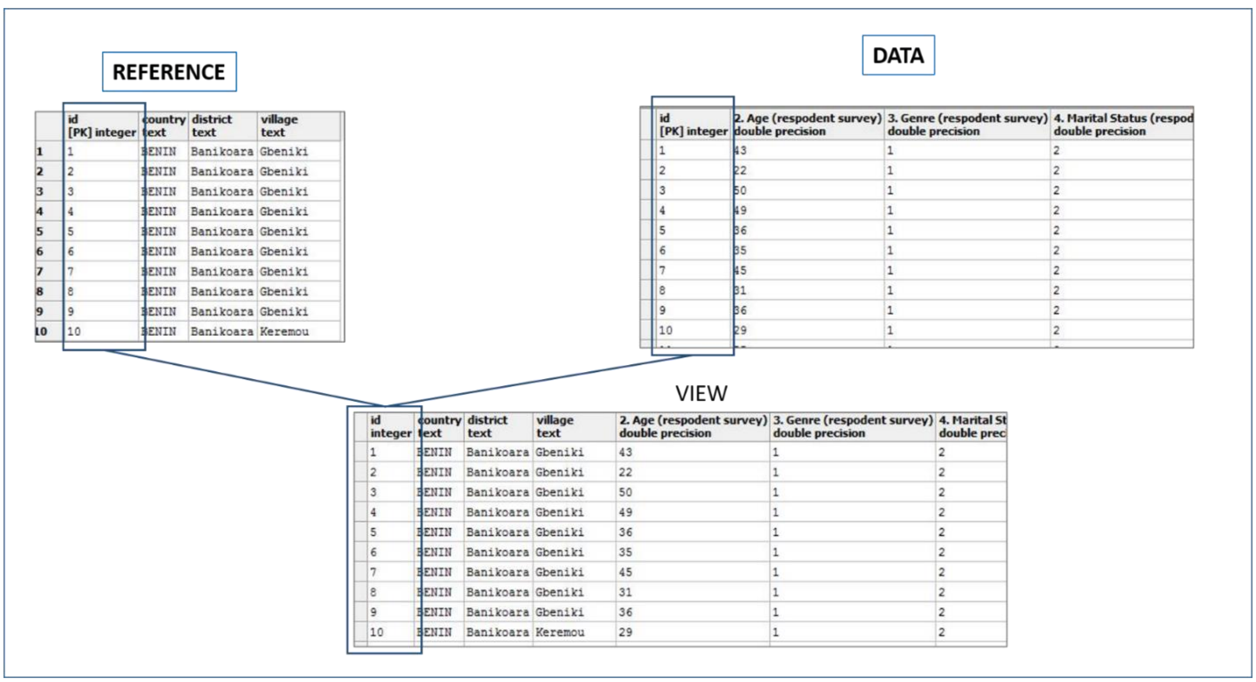


Figure 11: joining of socio economics tables and displaying of final view

4.6 Schema “swat”

This schema contains all outputs obtained by **Hydrology** processes. After running SWAT simulation, the process generates three data files (HRU, RCH, SUB) which are directly imported into PorstgreSQL database as the following tables:

HRU table (*<view_name>_hru*): it contains all indices associated to every HRU (Hydrologic Response Unit) of the basin. Imported from file *output.hru*.

Reach table (*<view_name>_rch*): it contains all indices associated to every reach of the basin. Imported from file *output.rch*. [reaches are as many as the HRUs]

Subbasin table (*<view_name>_sub*): it contains all indices associated to every subbasin of the area. Imported from file *output.sub*.

Since HRU and reach tables also show the subbasins they belong to (*SUB* field), it is easy to join them all into a single view to be visualised on the interface.

5 Installation and general settings

E-Water has two distinct setup packages for 32 and 64-bit operative systems in order to avoid compatibility issues that may rise when using particular libraries, but interface and processes featured in the software are always the same. Here are the basic requirements for running it properly:

- **CPU**: 2.2 GHz recommended
- **RAM**: 4GB recommended
- **Operative system**: ***Windows 7***, or later
- ***Microsoft Office 2007***, or later
- ***.NET Framework 4.5***, or later

By default, installation folder is *C:\Program Files\European Commission\E-Water*. At the end of setup process, a desktop shortcut will also be created.

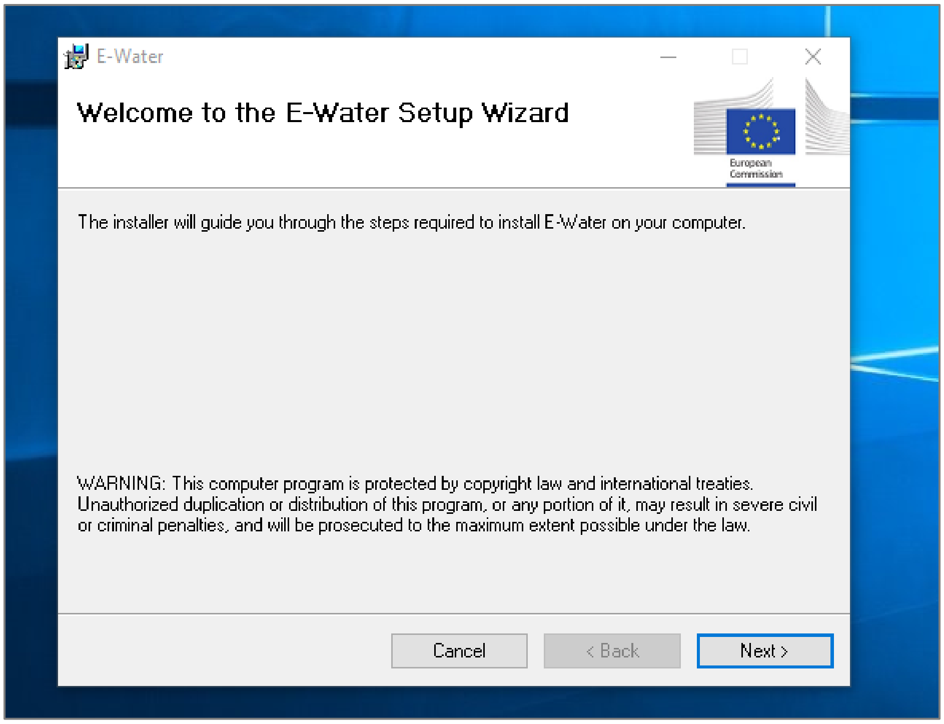


Figure 12: E-Water installation menu

During the first run, E-Water will check if all needed software are already present on the device and will propose to install the missing ones. Table 15 shows all requested applications with the version potentially used by this module.

Software	Description
PostgreSQL (v9.6.3)*	SQL database system
PostGIS (v2.3.2)	PostgreSQL plugin for geographic data

Table 15: list of required software installed by E-Water during its first execution

***IMPORTANT**: a new database connection will be created during *PostgreSQL* installation, so user has to take note of its features (port number, password, etc.) since they will be required for further operations on E-Water. The installation of *Stack Builder* is not required, so it can be skipped.

E-Water interface is currently localised in three different languages: English, French, and Italian. Other than selecting it at this stage, it is always possible to change language with ***Settings-> Language*** option on main menu.

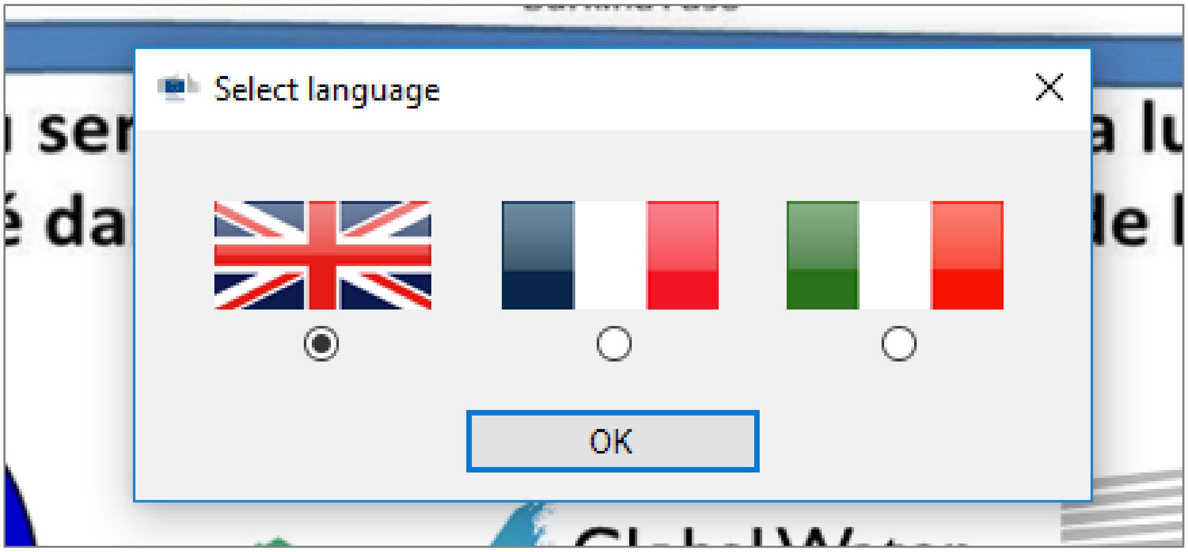


Figure 13: language selection menu (first run)

5.1 Connection

E-Water generally works with a set of *PostgreSQL* connections, so it will check at every start if there is at least one available. If not, it will automatically show the connection settings menu (Figure 14), where it is possible to add any number of existing connections.

First, it takes to insert all parameters related to desired connection:

- **Server**: localhost or remote machine (IP address needed)
- **Database** (default: *postgres*)
- **User ID** (default: *postgres*)
- **Password** (chosen during installation)
- **Port** (usually 5432)

At this point, it is possible to ***Add*** the connection to E-Water connection list (***Test*** it to check if it is correctly found) and visualise it on the left tree view panel, where servers are displayed as upper nodes and lower ones correspond to single databases. During this process, E-Water will also create a set of schemas, tables and functions needed to execute properly internal processes. However, since the tables related to ***Agriculture*** menu (EPIC) have significant space requirements, user will be asked explicitly whether to create them, since they are not needed aside of EPIC processes.

Through the tree view, it is possible to check all connections E-Water will effectively take into account (unchecked databases will be completely hidden for further operations). In addition, selected connection could be removed from list with ***Remove*** button (tables are not deleted since real database is not affected), or have its entire EPIC environment reset to original state with ***Reset EPIC data*** (this will also add default EPIC tables if not present). After quitting this section (OK), the main menu will appear automatically.

From the ***Connection*** drop-down list from the main menu (lower left corner), it is possible to select desired connection (represented as: *<host name>:<database>*) among all the one registered from this section. When a new connection is selected, all data lists related to database will be updated.

User may always manage connection list by accessing to this menu through ***Settings-> Connection*** from main interface.

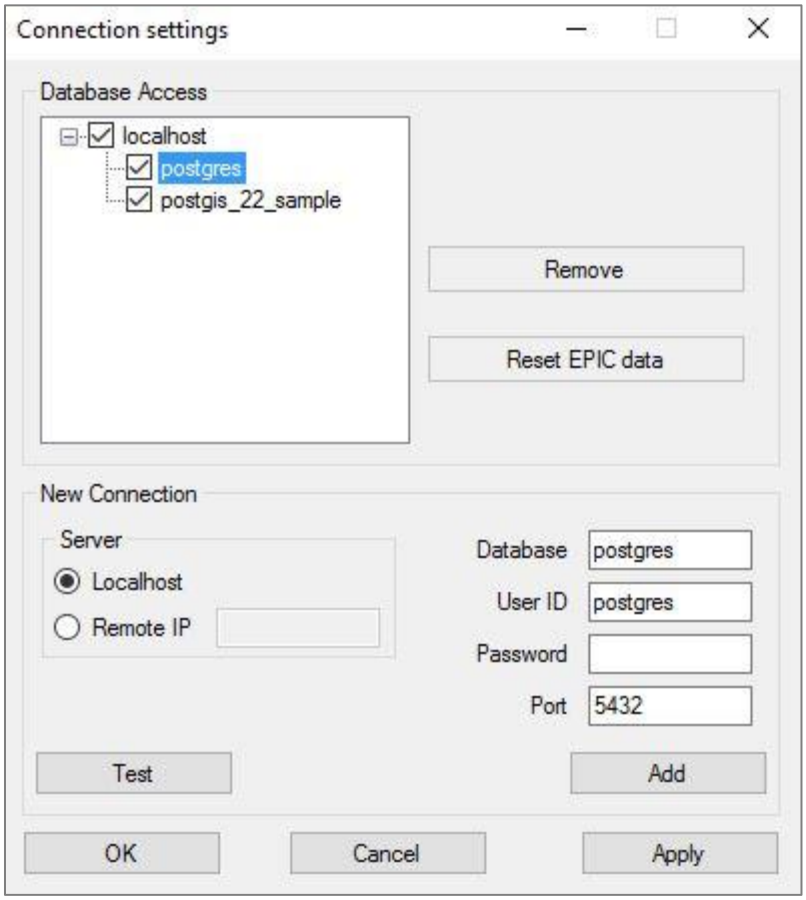


Figure 14: connection settings menu

5.2 Edit tables

When uninstalling E-Water, all PostgreSQL objects (schemas, tables, etc.) are not involved and remain available for further operations, even in case of reinstallation of this module. Rather than using *PgAdmin* interface, user can delete tables through ***Settings-> Edit Tables*** menu (Figure 15). In fact, it is possible to visualise easily any *E-Water* table on a checklist by selecting desired connection (represented as: *<host name>:<database>*) and schema, also using keyword search and multi checking to make selection easier and faster. With ***Drop table(s)*** command, all selected tables will be permanently removed from database, so it is recommended to proceed with caution since deleting objects related to internal processes may compromise a correct execution of the module. In order to limit this feature exclusively to internal data, schemas not related to E-Water environment (e.g. ***public*** schema) are not visualised in this menu.

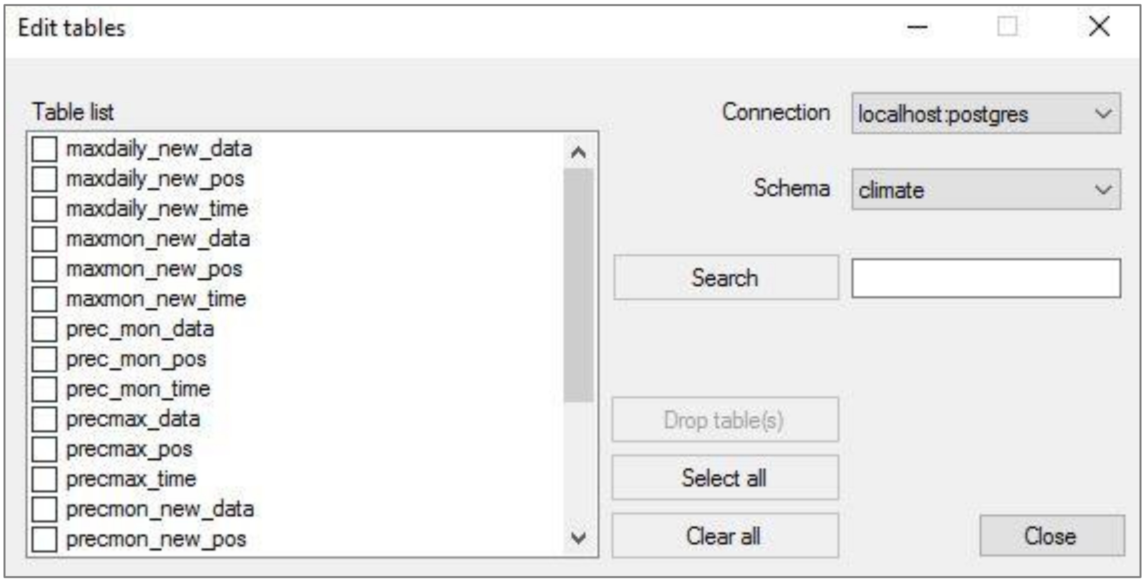


Figure 15: edit tables menu

Anyway, it is normally suggested to delete tables through the list panels on the main interface, since it allows removing them properly from databases without corrupting inner functionalities.

5.3 Configuration folders

Hydrology, ***Agriculture*** and ***Climate*** sections allow to manage all parameters composing process settings, by saving them in a configuration list with an identification name in order to load them for future operations. There is a distinct folder for each section containing all single configurations in form of *.xml* files: if desired, user may change the folder path through ***Settings-> Configuration folders*** menu (Figure 16).

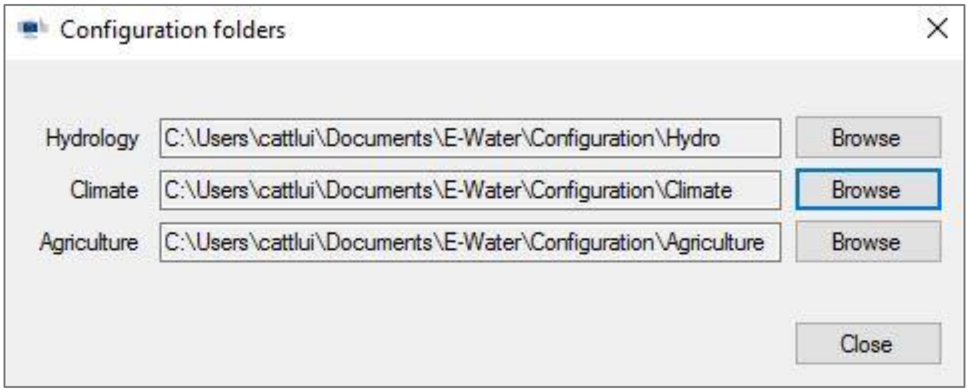


Figure 16: configuration folders menu

The interface for configuration managing is the same for all the three sections.

Configuration selection: from this drop-down menu, user can select all configurations already created. Selecting a new object instantly updates process configuration interface.

Save changes: updates current configuration by simply overwriting existing parameters.

Restore: reloads current configuration. All unsaved changes to configuration parameters are discarded.

Delete: delete current configuration from the list. It will be permanently removed from list.

Save as: creates a new configuration containing all current parameters in process configuration interface. Before saving, it takes to enter a name into the textbox aside which will be used to identify the configuration in the list.

6 Hydrology

In order to evaluate the impact of production activities on the basin area, land and crop management are key parameters for such analysis. Through this section, it is possible to execute Soil and Water Assessment Tool (SWAT) (Neitsch et al., 2011 [1]) model over hydrological data belonging to different subbasins – each one containing a set of one or more Hydrological Responsive Units (HRUs) - and display its results through plot charts and maps. All inputs are stored by default in Documents subfolder ***SWAT***, while model outputs are mainly saved as tables into PostgreSQL schema ***swat***. During the simulation, E-Water runs the process with SWAT executable file (release 2012) depending on software version installed (32 or 64 bits).

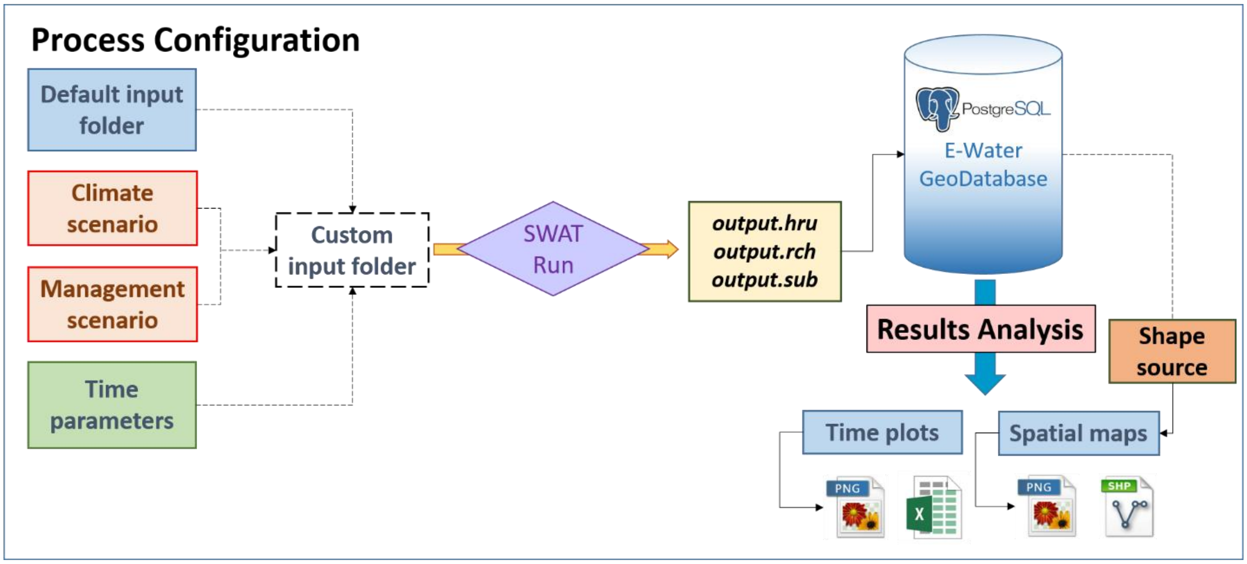


Figure 17: process diagram of Hydrology menu

As is possible to see on Figure, ***Hydrology*** menu (Figure 18) is divided in two distinct sections: ***Process configuration*** and ***Results analysis***.

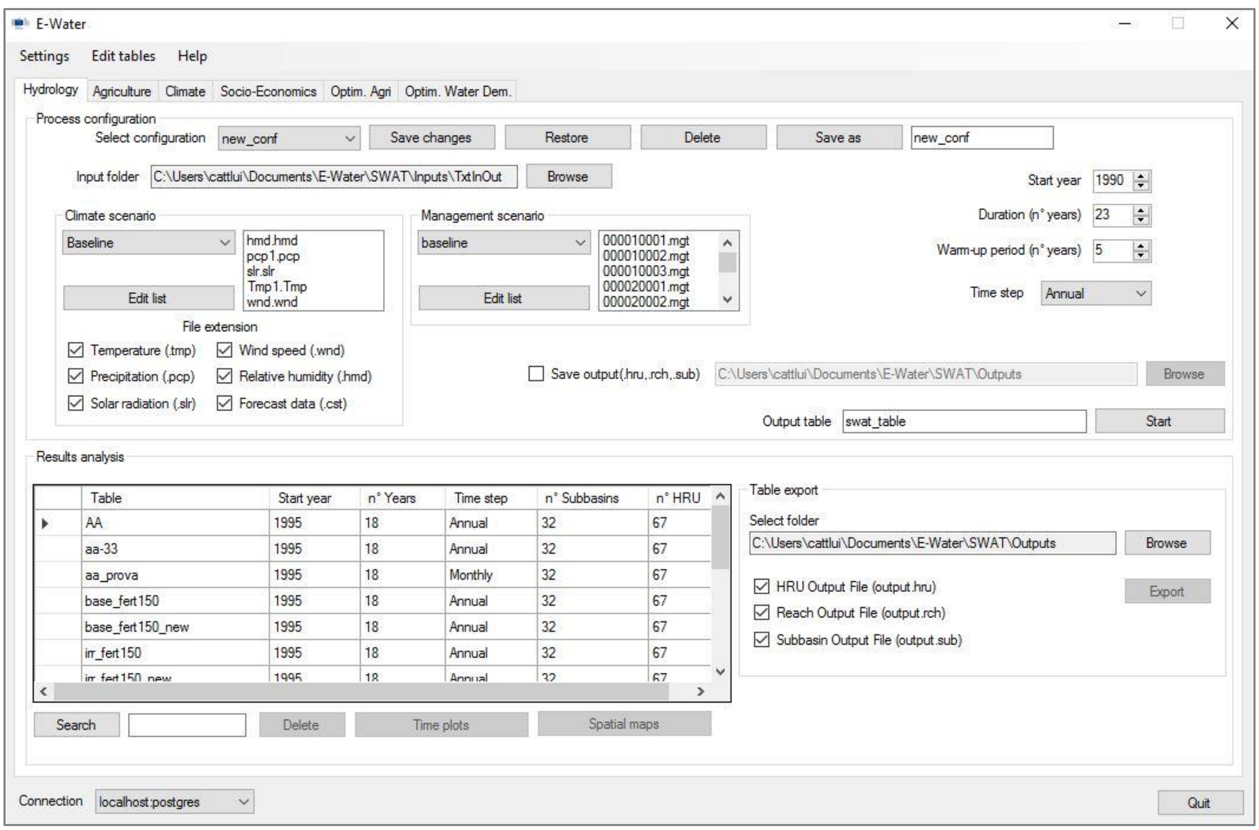


Figure 18: Hydrology menu

6.1 Process configuration

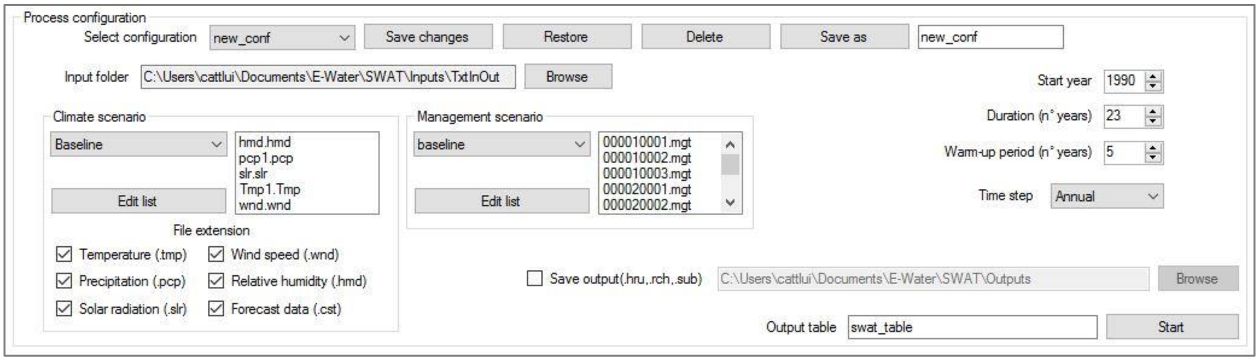


Figure 19: process configuration section, from Hydrology menu

The upper part of Hydrology menu (Figure 19) is used as a configuration tool for SWAT processes, where user can choose input data and scenarios along with time parameters, and define output names.

Save/load configuration: this feature (upper part of the menu) allows saving the whole process configuration (all data explained in this paragraph) with a name and reloading it afterwards by selecting it from a drop-down menu.

Input folder: this folder contains a large set of local data interested in the process. Many files are identified with a number suggesting the subbasin, watershed or HRU they describe, while others contain scenario information affecting the whole area. During the process, E-Water creates a temporary copy of this folder, partially subscribing it – selected scenario files are copied here and *file.cio* is filled with all parameter values for that specific run – so the original one is kept unchanged. By default, all input data are located in folder *Documents\E-Water\SWAT\Inputs\TxtInOut*.

Climate/management scenario: it is possible to choose among different climate and management scenarios (two separated lists), each one consisting in a set of files contained in a dedicated folder (all of them are displayed on the tables aside). For climate scenarios, user can decide whether considering or not a specific kind of file by checking on its extension name (temperature, precipitation, wind, radiation, humidity, forecast). Both list are editable by adding or removing existing scenarios (removed scenarios are no longer visualised, but they are not deleted from hard disk) from menus accessible through ***Edit list*** buttons which have the same structure (Figure 20). The panel on the upper part of the menu contains a list of existing scenarios available on the upper part of the form, where it is possible to see the main features of selected element (name, folder path and notes) and to remove it. From ***New scenario*** section, scenarios can be added by specifying their folder path, choosing a name and add notes (optional) before importing it into the list (***Add*** button).

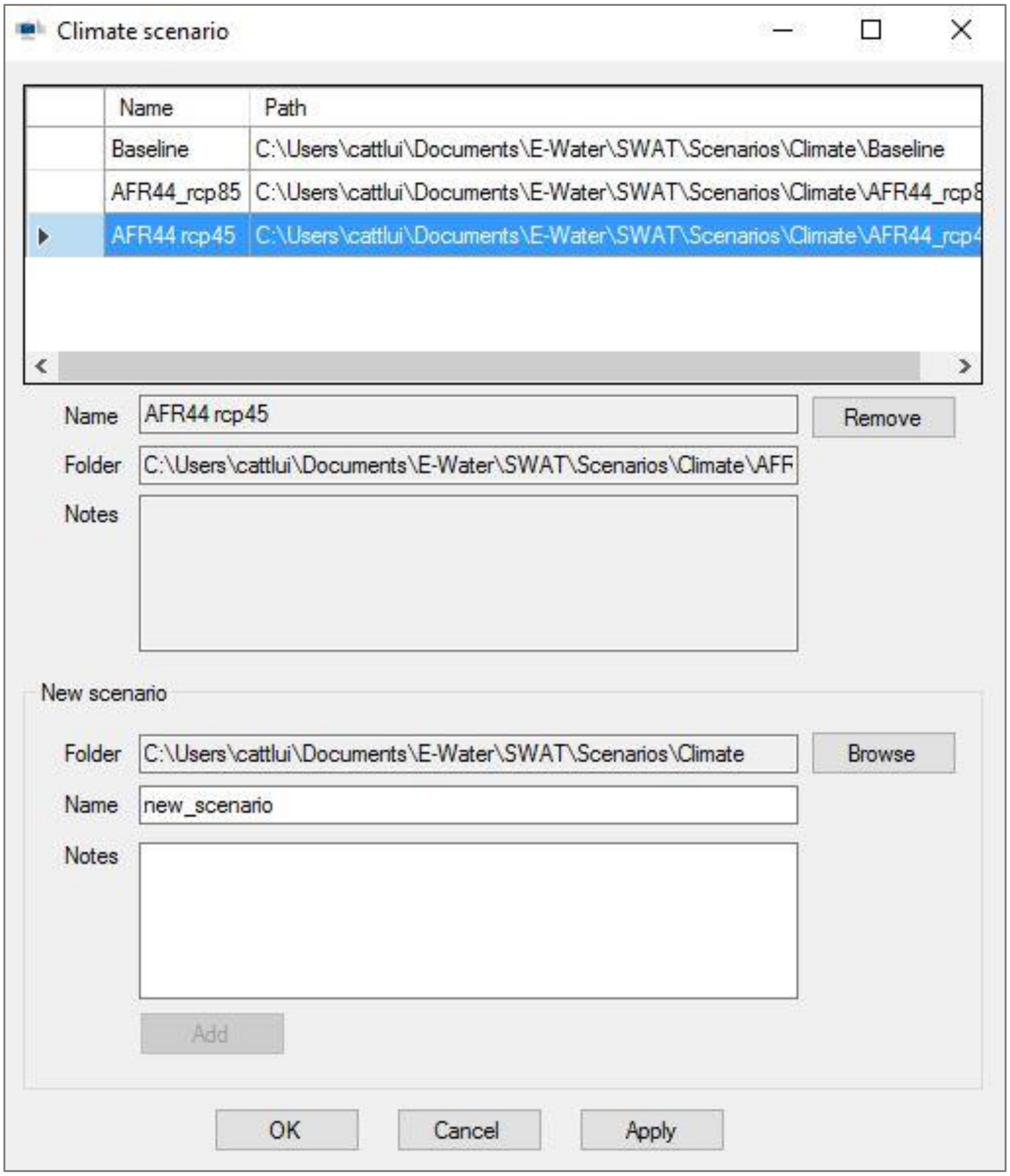


Figure 20: E-Water interface for climate scenario editing

Time: user can define time interval of output data by setting these parameters:

- **Start year**: when process starts (the initial date is always 1st January of this year).
- **Duration**: number of years of process total duration.
- **Warm-up period**: number of initial years to exclude from total duration. Since the earlier values of modelled data are affected by an initial transitory trend (warm-up) before reaching a stable behaviour, it is a common choice to simply remove them in order to avoid evaluation errors (default value: **5 years**).

- **Time step**: time interval between two consecutive steps of output time series. It could be annual, monthly or daily. Shorter steps mean higher amounts of data to evaluate, and a subsequent increase of process duration.

> Example: selecting 1990 as start year, a duration of 23 years, a warm-up period of 5 years and an annual step will produce data from 1995 to 2012 (18 years) with 18 total time steps.

Output: output data produced by every SWAT run is stored into a set of PostgreSQL tables identified with a single name chosen by user (***Output table***) and visible through a list on **Results analysis** menu. In addition, it is possible to save results as text files (*.rch*, *.hru*, *.sub*) into SWAT output folder by checking ***Save output*** option and selecting the desired folder path.

6.2 Results analysis

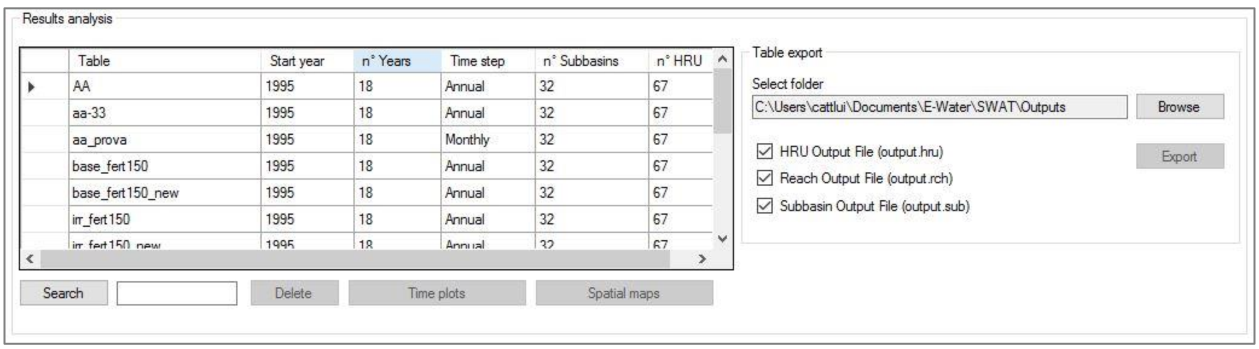


Figure 21: process configuration section, from Hydrology menu

This section (Figure 21) shows a panel with all outputs created by previously executed SWAT simulations. Every output is displayed on a row along with main information about it (effective start year and duration, time step, number of subbasins and HRUs). After selecting the row corresponding to desired output (***Search*** + keyword for a better selection), user can decide to analyse its indices as time charts (***Time plots***), or with georeferenced values on a geographical domain (***Spatial maps***). Other than that, it is possible to permanently remove it from database (***Delete***), or export it (***Export***) as a set of text files (*output.rch*, *output.hru*, *output.sub*, only checked ones) into selected folder.

6.2.1 Time plots

By selecting this display option, a form will appear (Figure 22), showing all indices belonging to selected output table (see Table 16). Once the form is shown, it is possible to choose desired index (by ***Index*** drop-down menu) and its whole period by setting ***Start date***, ***End date*** and all available ***Time step*** options (e.g. in case of monthly output data, it will not be displayable on daily scale). Then, on the ***SUB/HRU*** subpanels set, user may select all subbasins or HRUs to be plotted, each one as a distinct chart: they are all listed with their original subbasin (or HRU) identification number (***Show subbasins*** to see them on the map). In case of HRUs, its related subbasin is also shown (***HRU: 5 (S2)*** means that HRU #5 belongs to subbasin #2). Once they are checked, it is possible to edit their ***Colour***/***Width***/***Dash*** style on the plot panel from the above list; if checked, a ***Legend*** listing all displayed elements will also appear: it is possible to move it within the panel through arrow buttons ***up***/***down***/***right***/***left***, while diagonal arrow buttons directly put it on the corners.

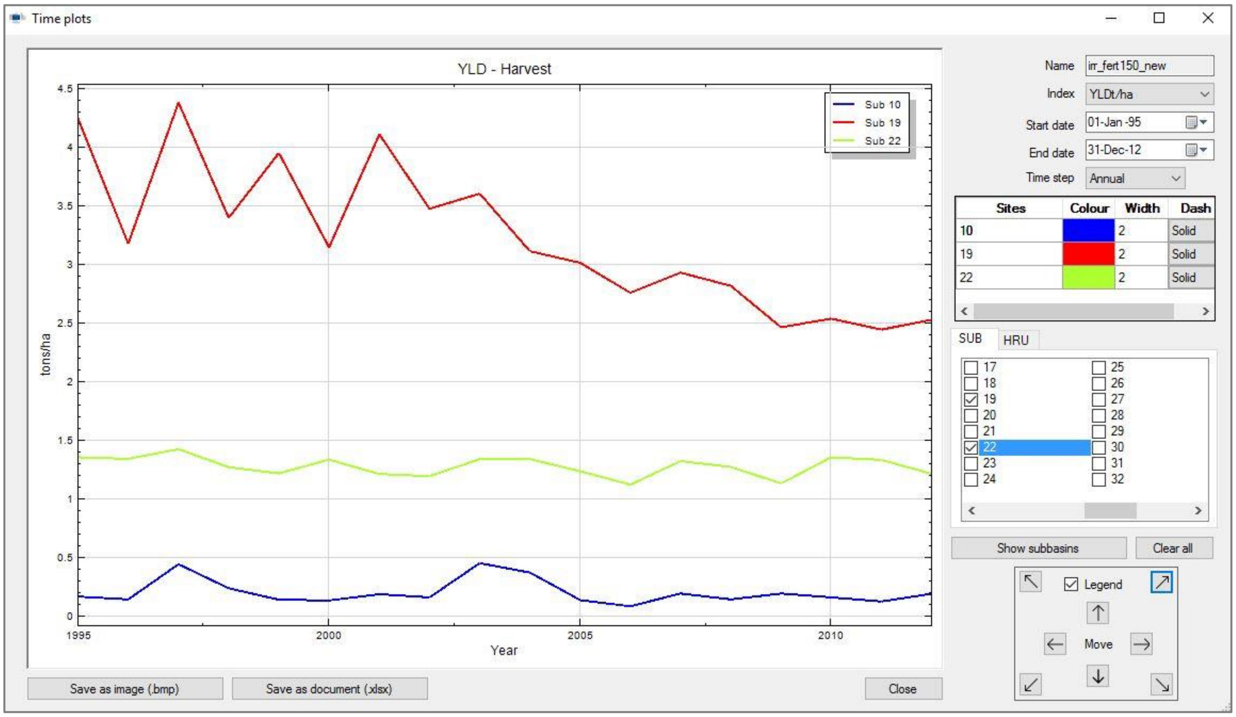


Figure 22: time plot interface for Hydrology outputs

Finally, displayed results may be saved as a .bmp images at 1:1 scale (remember that the whole window is resizable, so image size may vary at user will), or .xlsx tables, where the sheet is named after the index, and each column corresponds to a distinct subbasin/HRU.

Index name	Definition	Unit
AREA_RCHkm2	Area drained by reach	km2
FLOW_INcms	Average daily flow in	m3/sec
FLOW_OUTcms	Average daily flow out	m3/sec
NO3_INkg	NO3 in	kg (N)
NO3_OUTkg	NO3 out	kg (N)
NH4_INkg	NH4 in	kg (N)
NH4_OUTkg	NH4 out	kg (N)
NO2_INkg	NO2 in	kg (N)
NO2_OUTkg	NO2 out	kg (N)
AREA_SUBkm2	Subbasin area	km2
PRECIPmm	Rain	mm
PETmm	Potential evapotranspiration	mm
ET_SUBmm	Subbasin evapotranspiration	mm
SWmm	Soil water content	mm
PERC_SUBmm	Subbasin percolation	mm
SURQmm	Surface runoff	mm
WYLD_SUBmm	Subbasin water yield	mm
NSURQkg_ha	N-NO3 in surface runoff	kg/ha
LAT_Qmm	Later flow	mm
LAT_Q_NO3kg_ha	N-NO3 in lateral flow	kg/ha
GWNO3kg_ha	N-NO3 in groundwater flow	kg/ha
AREA_HRUkm2	HRU area	km2
IRRmm	Irrigation	mm
GW_RCHGmm	Aquifer recharge	mm
DA_RCHGmm	Deep aquifer recharge	mm
REVAPmm	Water returning to Root zone	mm
TMP_AVdgC	Mean temperature	°C
TMP_MXdgC	Maximum temperature	°C
TMP_MNdgC	Minimum temperature	°C
N_APPkg/ha	Fertilizer Applied (manual)	kg(N)/ha
NAUTOkg/ha	Fertilizer Applied (auto)	kg(N)/ha
NUPkg/ha	N Plant uptake	kg(N)/ha
W_STRS	Water stress days	n° days
TMP_STRS	Temperature stress days	n° days
N_STRS	N stress days	n° days
P_STRS	P stress days	n° days
BIOMt/ha	Biomass	tons/ha
LAI	Leaf area index	n
YLDt/ha	Harvest	tons/ha

Table 16: list of Hydrology output indices

6.2.2 Spatial maps

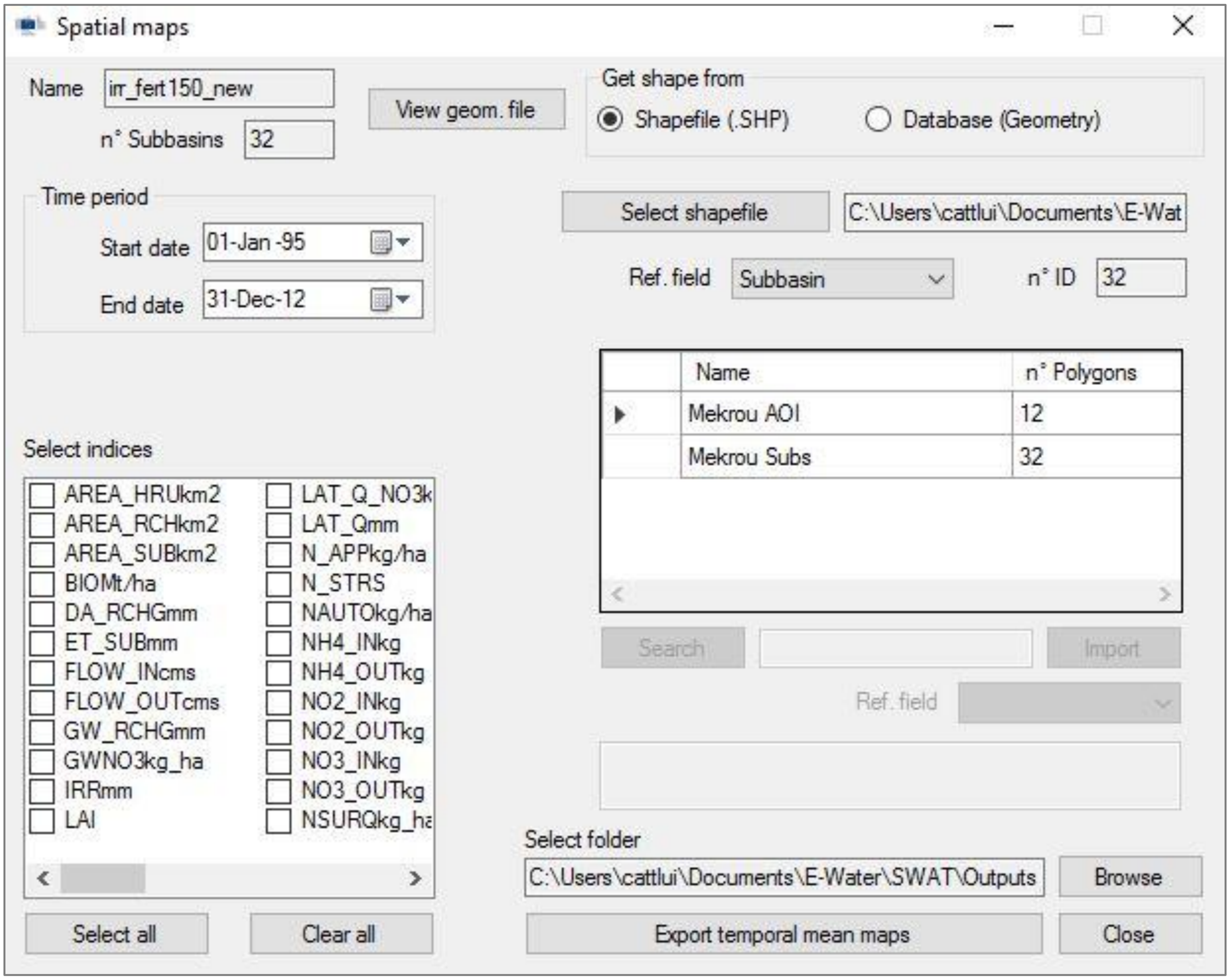


Figure 23: spatial maps interface for Hydrology outputs

Through this menu (Figure 23), E-Water is able to produce and export SWAT output indices on all the subbasins of the area in form of spatial maps: for each index selected in the checklist on the lower left corner, it will evaluate its temporal means between selected ***Start date*** and ***End date*** for every subbasin belonging to output table.

Then, it also takes to choose a spatial reference object in order to properly localise every subbasin, either as file (*.shp*) or *PostgreSQL* table with *geometry* field (***Get shape from*** selection menu). In order to have a proper spatial representation, it is necessary to have a correspondence of subbasin identifications between data output and space reference. For this reason, when shape source is chosen, it is necessary to select the ***Ref. field*** containing subbasin identification values (they must be in form of distinct integers, starting from **1**) which are expected to exactly match with the ones contained in output table. User can have an overall look through ***View geom. File*** (Figure 24), showing all featured fields and values of selected space reference, before choosing the right subbasin field. The best possible condition is when the number of subbasins from output table (***n° Subbasins***) equals the ones contained in spatial table (***n° ID*** or ***n° Polygons***, depending on selected source type);

if not, a warning message will appear before final evaluation, informing user of the incomplete match.

mekrou_subs

gid	objectid	gridcode	subbasin	area	slo1
1	1	1	1	16933.8600000	1.52556288242
2	2	2	2	77425.4700000	1.69211292267
3	3	3	3	56459.4300000	2.06690406799
4	4	4	4	23461.6500000	1.30085992813
5	5	5	5	54312.1200000	1.41976404190
6	6	6	6	13446.8100000	1.40117681026
7	7	7	7	3324.24000000	1.43629360199
8	8	8	8	44277.0300000	1.35219991207
9	9	9	9	42038.1900000	1.38378036022
19	19	19	19	41898.0600000	2.34030032158
20	20	20	20	48855.1500000	1.78063189983
21	21	21	21	41821.1100000	1.79284334183
22	22	22	22	32525.5500000	3.12729215622
23	23	23	23	23800.2300000	3.65587711334
24	24	24	24	22421.6100000	1.91000235081
25	25	25	25	26041.5000000	3.01695990562
26	26	26	26	15276.6000000	1.89820897579
27	27	27	27	[illegible]	[illegible]

Figure 24: geometry table view from database

With ***Export temporal mean maps*** button, results will be saved into selected folder path as image files (**Images** folder, one for each index, as seen in Figure 25), and their corresponding shapefile (**Shapefile** folder) containing all indices to be visualised on GIS interface.

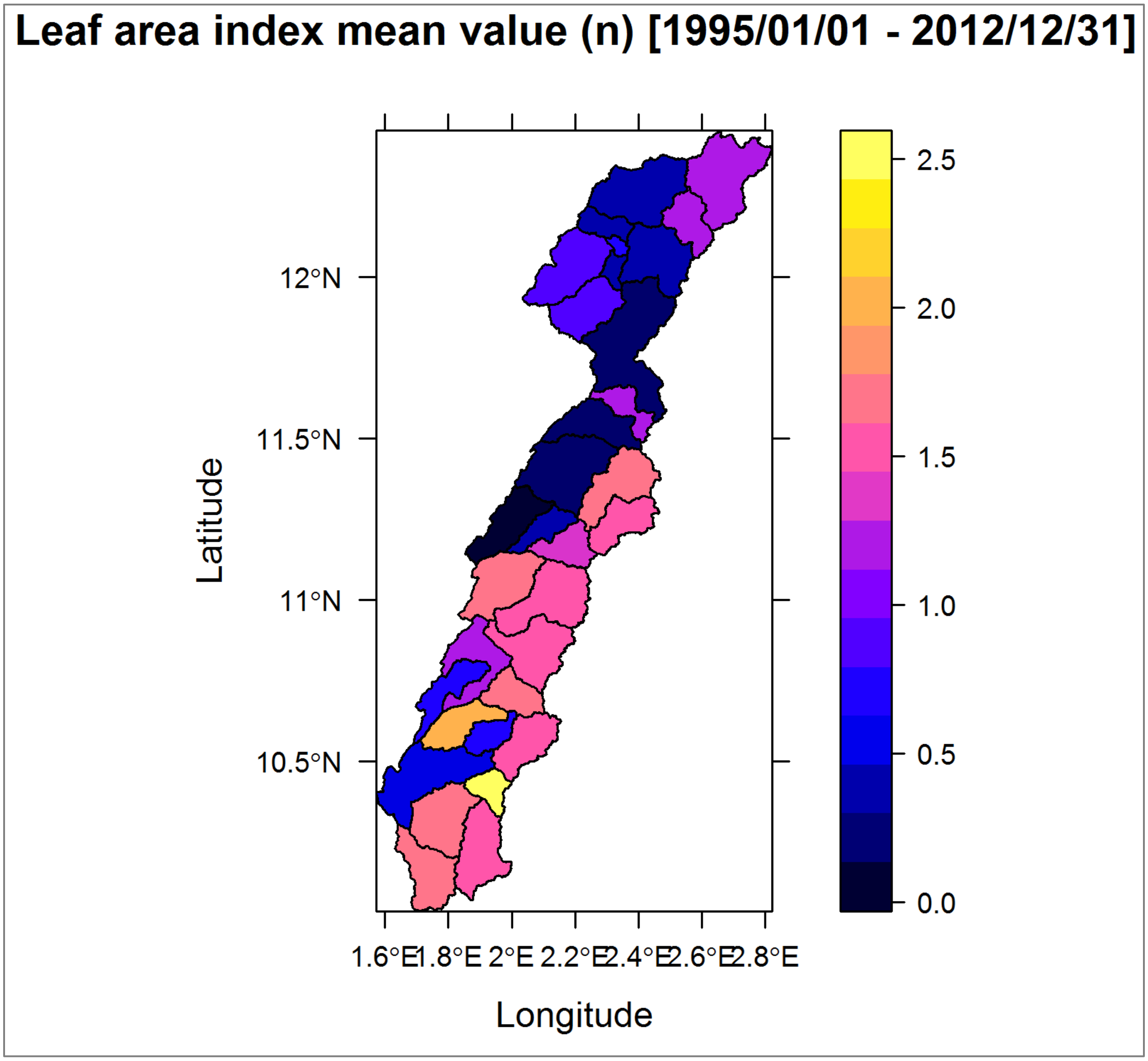


Figure 25: output spatial map showing the mean values of Leaf Area Index (LAI) on Mékrou subbasins between years 1995 and 2012.

7 Agriculture

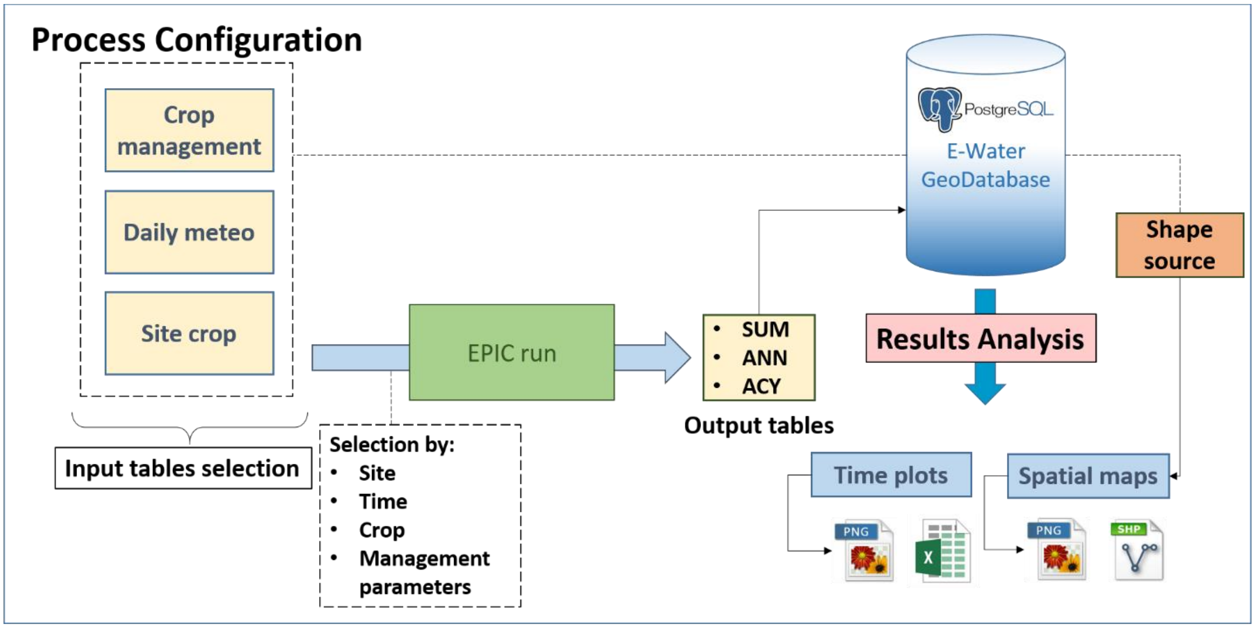


Figure 26: process diagram of Agriculture menu

Environmental Policy Integrated Climate (EPIC) is a biophysical model for cultivated soil management which can be applied to agricultural purposes (Wang et al., 2012 [2]). By evaluating the effects of management decisions on soil, water, nutrient and pesticide movements, it is able to simulate crop development, even mixed crops, and their replacement, depending on all agronomic parameters (climate, soil, agricultural activities) with a daily time step modelling. Its processes mostly involve weather simulation, hydrology, erosion-sedimentation, nutrient cycling, pesticide fate, crop growth, soil temperature, tillage, economics, and plant environment control.

As shown on Figure 27, EPIC management interface is available from E-Water through ***Agriculture*** menu. After user sets up overall parameters (time, sites, crops, management) and selects input tables (*daily meteo*, *crop management*, *site crop*) from database, the module interface runs the EPIC model executable (version 0810, released in 2013) and finally produces a set of output tables which are first stored inside database, making them available for further visualisation or export (Figure 26).

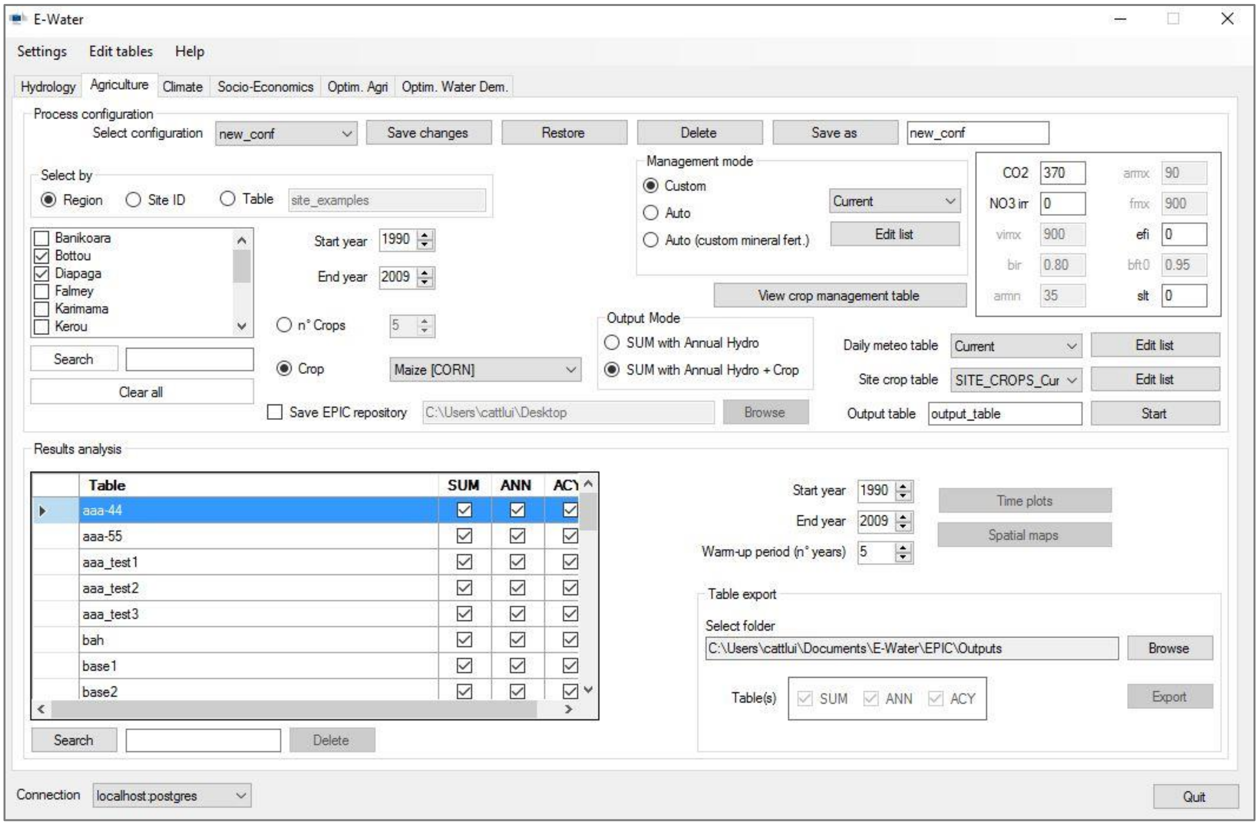


Figure 27: Agriculture menu

7.1 Process configuration

The upper part of Agriculture menu (Figure 28) is used as a configuration tool for EPIC processes, where user can choose input data tables and parameters, and define output mode and table name.

Save/load configuration: this feature (upper part of the menu) allows saving the whole process configuration (all data explained in this paragraph) with a name and reload it afterwards by selecting it from a drop-down menu.

Spatial selection: the process is limited within the points defines by ***site_id_nuts_names*** table, but it is possible to define where to focus the process in three different ways. Selecting by ***Region***, the below checklist shows all regions available, so only checked ones will be processed. ***Site ID*** selection, rather than previous option criterion, allows user to focus on single station points, each one identified by an ID number. Finally, it is also possible to directly use a custom ***Table*** from **epic** schema (*PosgreSQL*) containing all points to process (*site_example* is included by default and may be used for test runs). It takes to notice that selecting a large number of points/regions often leads to long process. Anyway, a popup window will show the progress percentage (points processed / total) and process can be aborted anytime.

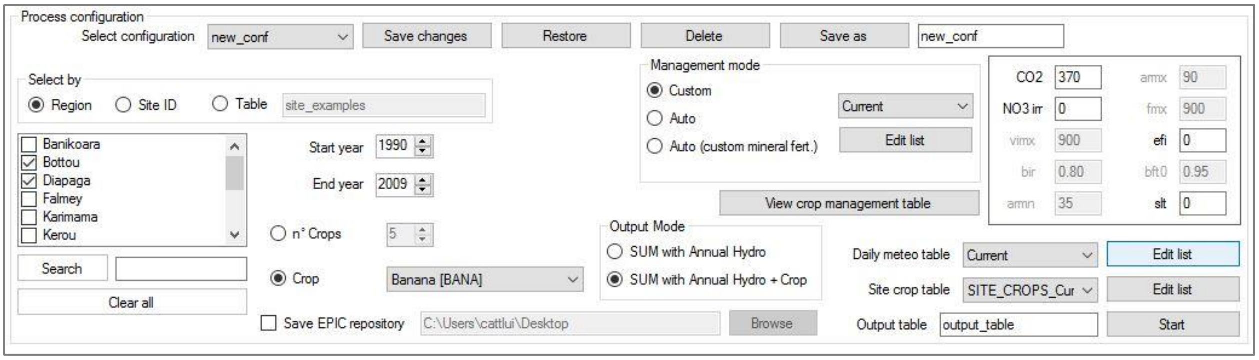


Figure 28: process configuration section, from Agriculture menu

Year selection: user can choose the time interval (***Start year*** / ***End year***) for the simulation to carry out. **IMPORTANT**: since default data start from **1990**, there is no point to select earlier starting years.

Crop selection: there are two ways to choose the crops to be focused in the ongoing process. Selecting a number (***n° Crops*** option), the algorithm will take into account the *n* most relevant crops for each processed point, meaning that if multiple points are involved there could be more than *n* total crops in the output. Otherwise, it is possible to force the selection on a single crop (***Crop*** option) specifically determined by user.

Management mode: this selection involves the crop management table to be selected as input (containing daily parameters about irrigation, fertilisation, sowing, etc.), the mode it is used during the process, plus ten numeric parameters in the aside panel to set manually. ***Custom*** mode will take into account all fertilisation parameters (Table 17) contained in selected table (disabling them from the parameters panel), while ***Auto*** will reject them all, using instead the parameters values aside (*vimx*, *vir*, *armn*, *armx*, *fmx*, *bft0*). ***Auto (custom mineral feat.)*** option is an intermediate stage, taking from table only mineral fertilisation parameters. Through ***View crop management table***, it is possible to have a glimpse of all available crop management tables, and have a correct visualisation depending on selected mode (custom, auto, etc.), crop and region (Figure 29).

Parameter	Definition	Unit
CO2	Starting CO2 concentration in atmosphere	ppm
NO3 IRR	Starting NO3 concentration in irrigation water	ppm
VIMX	Max. annual irrigation	mm
BIR	Plant water stress trigger	-
ARMN	Min. single irrigation	mm
ARMX	Max. single irrigation	mm
FMX	Max. annual N application	kg/ha
EFI	Runoff/volume of irrigation	mm
BFT0	Fertilization stress trigger, plant N stress	-
SLT	Starting concentration of salt	ppm

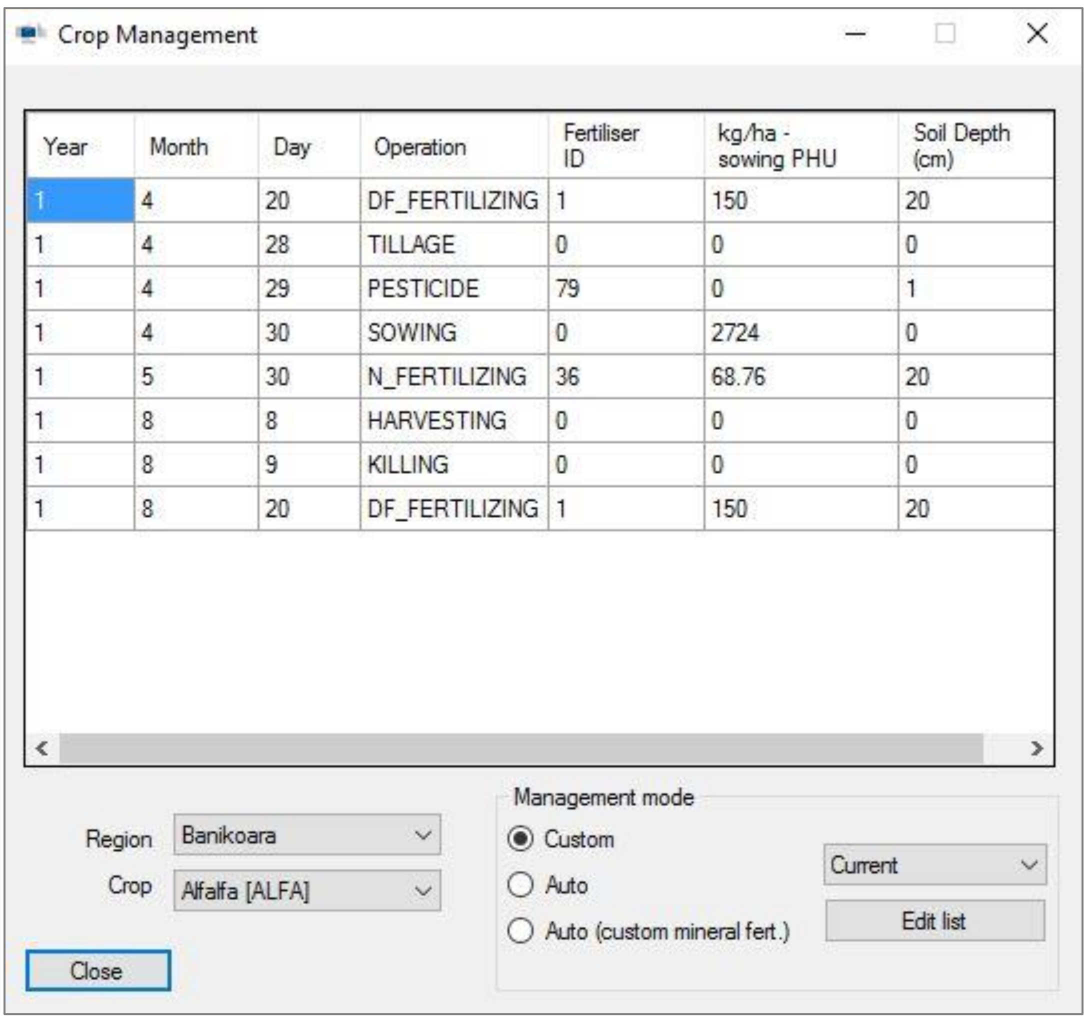
Table 17: crop management parameters list

Daily meteo / site crop table: the selection of these two kinds of tables is carried out in the same way. *Daily meteo* tables provide daily meteorological data for every station point, and user can select which one to use as input scenario. Each table contains the following fields:

- Temperature (max. / min.)
- Rainfall
- Wind speed
- Vapour pressure
- Solar radiation

Site crop tables show the usage rate of every crop for all station points (one column per crop, plus station ID field).

Both table lists can be edited (***Edit list***, see Figure 30) by adding or removing existing tables: new tables are imported from *.csv* files into ***epic*** schema with name chosen by user (***Table name***), but are actually visualised on the interface with selected ***Short name (label)***.



Year	Month	Day	Operation	Fertiliser ID	kg/ha - sowing PHU	Soil Depth (cm)
1	4	20	DF_FERTILIZING	1	150	20
1	4	28	TILLAGE	0	0	0
1	4	29	PESTICIDE	79	0	1
1	4	30	SOWING	0	2724	0
1	5	30	N_FERTILIZING	36	68.76	20
1	8	8	HARVESTING	0	0	0
1	8	9	KILLING	0	0	0
1	8	20	DF_FERTILIZING	1	150	20

Figure 29: crop management table view

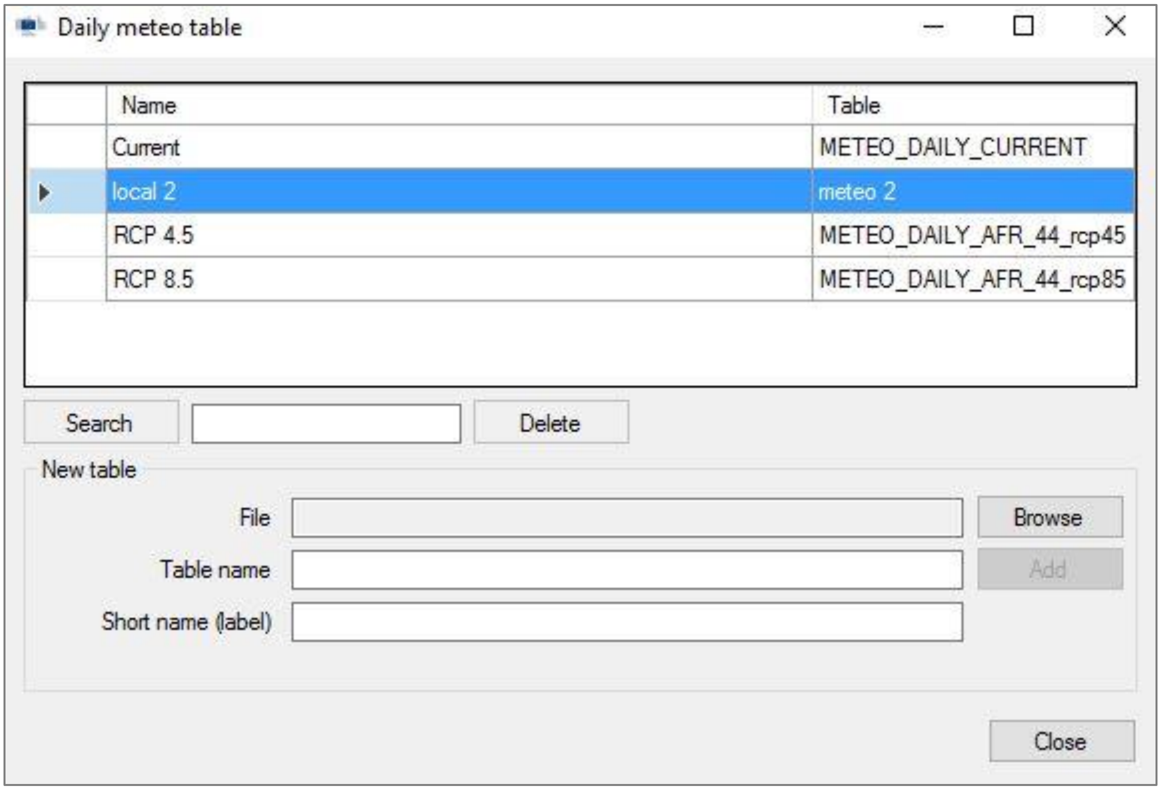


Figure 30: editing interface for daily meteo tables list (same for other input tables types)

Output mode: through this option, it is possible to decide which kind of output tables to produce at the end of the process. With ***SUM with Annual Hydro***, just annual summary (SUM) and annual water summary (ANN) tables will be created, while choosing ***SUM with Annual Hydro + Crop*** will also produce crop yield table (ACY).

Output: finally, user can choose how to name output table will be visualised on output list. In addition, by checking ***Save EPIC repository*** it is possible to save a copy of the repository folder at the end of process, containing all data used and produced by that EPIC simulation.

7.2 Results analysis

After EPIC simulation is completed, it is possible to see results saved into output table list on the left side of this section (Figure 31). Each row represents a distinct output, showing which data tables it contains exactly (SUM, ANN, ACY). When one of them is selected, ***Start year*** and ***End year*** fields are automatically set according to the total period used for that simulation; such values can be modified, but it is not possible to exceed the original period. Also, ***Warm-up period*** (5 years minimum) indicates how many starting years will be ignored in order to avoid evaluation errors. For instance, an output table starting from 1990 with a 5 years warm-up will actually count as if it begins on 1995.

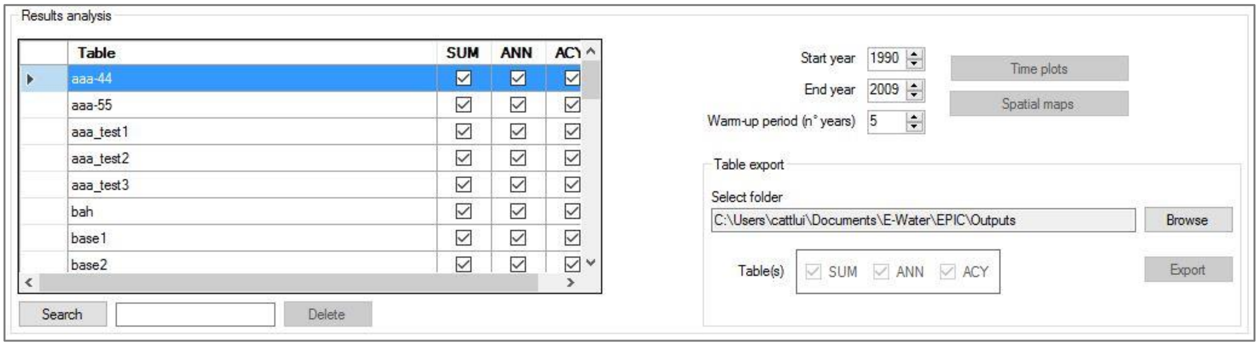


Figure 31: results analysis section, from Agriculture menu

Table export feature enables user to export output tables as *.csv* files into selected folder. They are exact replicas of the tables stored into database, one for each type (only checked ones are exported).

7.2.1 Time plots

By selecting this display option, a form will appear, showing all indices belonging to selected output table. Once the form is shown, it is possible to choose which index of a given crop to display (using ***Index*** and ***Crop*** drop-down menus), and its whole period by setting ***Start date***, ***End date***. User must also select between two sets of indices, provided respectively by ***Annual Hydro*** (ANN) (Table 18) and ***Annual Crop Yield*** (ACY) (Table 19) tables.

Each displayed plot corresponds to the time series of selected index value for a single region (in case of more station points belonging to a region, the value is a spatial mean). All regions interested by output data appear in the ***Sites*** list, but only checked ones are plotted (it is also possible to select their ***Colour***/***Width***/***Dash*** style). If checked, a ***Legend*** listing all displayed elements will also appear: it is possible to move it within the panel through arrow buttons ***up***/***down***/***right***/***left***, while diagonal arrow buttons directly put it on the corners.

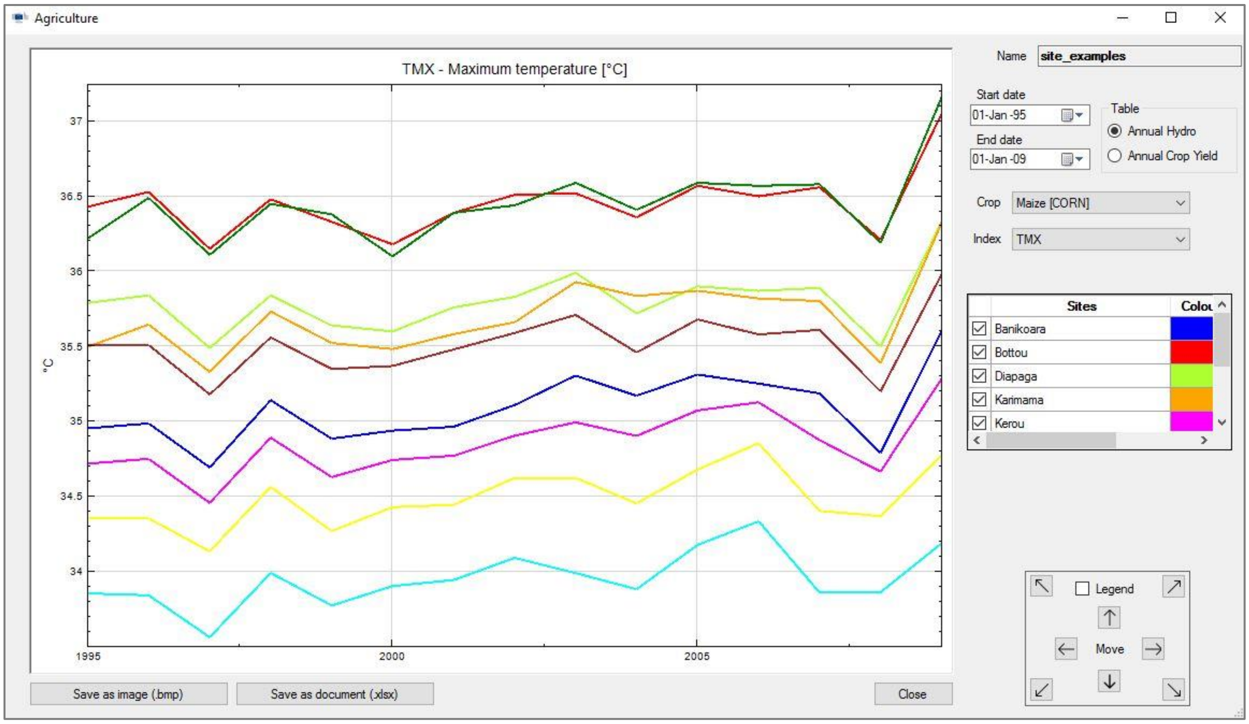


Figure 32: time plot interface for Agriculture outputs

Finally, displayed results may be saved as a .bmp images at 1:1 scale (remember that the whole window is resizable, so image size may vary at user will), or *.xlsx* tables, where the sheet is named after the index, and each column corresponds to a distinct region.

Index name	Definition	Unit
TMX	Maximum temperature	°C
TMN	Minimum temperature	°C
RAD	Solar radiation	mJ/m^2
PRCP	Precipitation	mm
PET	Potential evaporation	mm
EP	Transpiration	mm
Q	Annual Surface Runoff	mm
SSF	Lateral subsurface flow	m
PRK	Percolation below the root zone	mm
IRGA	Irrigation water applied	mm
USLE	Soil loss from water erosion using USLE	T/ha
PRKN	Mineral N loss in percolate	kg/ha
QNO3	Nitrate loss in surface runoff	kg/ha
GMN	Nitrogen mineralized	kg/ha
DN	N loss by denitrification	kg/ha

NFIX	N fixed by leguminous crops	kg/ha
NITR	Nitrification	kg/ha
AVOL	Nitrogen volatilization	kg/ha
QAP	Labile phosphorus loss in runoff	kg/ha
PRKP	Phosphorus loss in percolate	kg/ha
FNO	Nitrogen fertilizer (animal waste)	kg/ha
FNO3	N fertilizer nitrate	kg/ha
FNH3	N fertilizer ammonia	kg/ha

Table 18:output Annual Hydro (ANN) indices

Index name	Definition	Unit
YLDG	Grain yield	T/ha
YLDF	Forage yield	T/ha
BIOM	Biomass	T/ha
YLN	Nitrogen used by crop	kg/ha
YLP	Phosphorus used by crop	kg/ha
FTN	Nitrogen applied	kg/ha
FTP	Phosphorus applied	kg/ha
IRGA	Irrigation water applied	mm
WS	Water stress days	dd/yr
NS	Nitrogen stress days	dd/yr
PS	Phosphorus stress days	dd/yr
KS	Potassium stress days	dd/yr
TS	Temperature stress days	dd/yr

Table 19: output Annual Crop Yield (ACY) indices

7.2.2 Spatial maps

Through this menu (Figure 33), E-Water is able to produce and export EPIC output indices on all the regions of the area in form of spatial maps: for each index selected in the checklist on the lower left corner, it will evaluate its temporal means between selected ***Start date*** and ***End date*** for every region belonging to output table.

Then, it also takes to choose a spatial reference object in order to properly localise every site (region), either as file (*.shp*) or PostgreSQL table with *geometry* field (***Get shape from*** selection menu). In order to have a proper spatial representation, it is necessary to have a correspondence of region names between data output and space reference. For this reason, when shape source is chosen, it is necessary to select the ***Ref. field*** containing the region names (e.g. *Banikoara*, *Diapaga* for Mékrou Basin) which are expected to exactly match with the ones contained in output table. User can have an overall look through ***View geom. file***, showing all featured fields and values of selected space reference, before choosing the right name field. The best possible condition is when the number of sites from output table (***n° Regions***) equals the ones contained in spatial table (***n° ID*** or ***n° Polygons***, depending on selected source type).

With ***Export temporal mean maps*** button, results will be saved into selected folder path as image files (**Images** folder, one for each index), and their corresponding shapefile (**Shapefile** folder) containing all indices to be visualised on GIS interface.

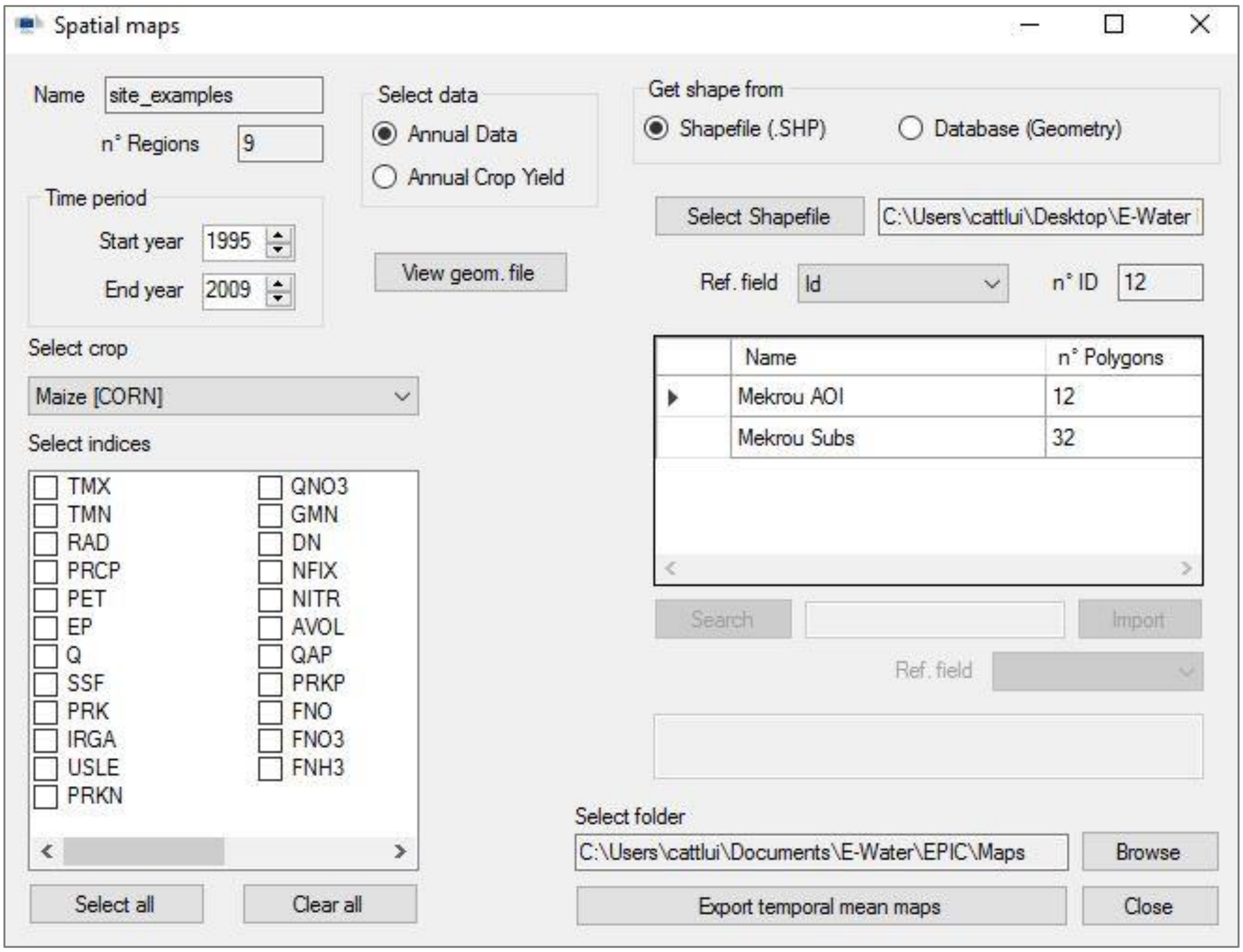


Figure 33: spatial maps interface for Agriculture outputs

8 Climate

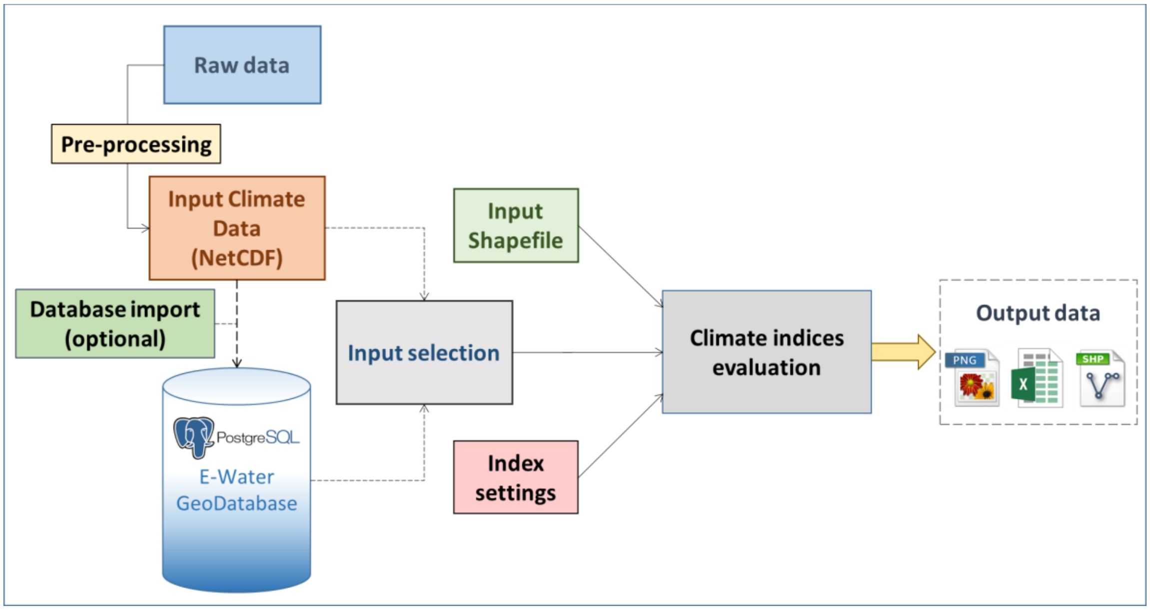


Figure 34: process diagram of Climate menu

As a part of the project, it is crucial for both policy makers and researchers to understand climate variability at local-regional-continental scales. In this context, the software allows to gather and process climate data available in river basin, in order to produce concise and clear information about the variability of key climatic variables, such as precipitation and temperature.

More specifically***, Climate*** feature of e-Water module allows user to process climate data in order to evaluate indices like return periods, deficit rates and heat waves. First, raw data (usually a set of files in NetCDF format) is arranged into a suitable input for further processes. Once desired input is available, it can be used to obtain desired indices, which can be displayed as images (.*png*) or through GIS interface (.*tif*) (Figure 34).

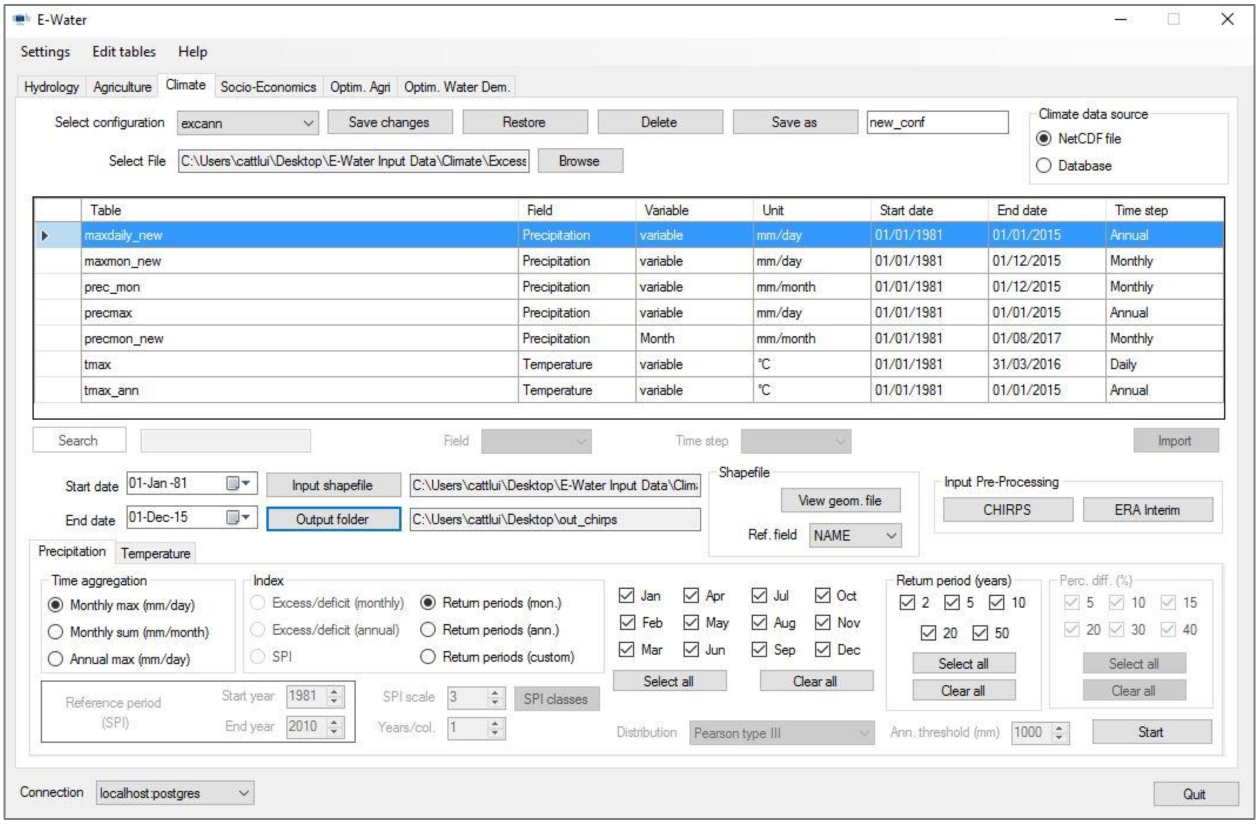


Figure 35: Climate menu

8.1 Data pre-processing

From ***Data Pre-processing*** box, it is possible to modify imported ***CHIRPS*** and ***ERA-Interim*** datasets in order to have a total compatibility with E-Water processes. Both temperature and precipitation datasets are provided in the Hierarchical Data Format (HDF). HDF is a set of file formats (i.e. HDF4, HDF5) designed to store and organize large amounts of data. Originally developed at the National Centre for Supercomputing Applications, it is supported by the HDF Group, a non-profit corporation whose mission is to ensure continued development of HDF5 technologies and the continued accessibility of data stored in HDF. In keeping with this goal, HDF libraries and associated tools are available under a liberal license for general use. HDF is supported by many commercial and non-commercial software platforms, including Java, MATLAB, IDL, Python and R. The freely available HDF distribution consists in the library, command-line utilities, test suite source, Java interface, and the Java-based HDF Viewer (i.e. HDFView).

Given their different format, there is a dedicated menu for each dataset that allows choosing time aggregation and, by checking ***Clip with Shapefile*** option, clipping data into the rectangular boundaries of a selected domain shapefile. When setting is complete, ***Create NetCDF*** will produce processed file, ready for index analysis (output path is defined by user).

CHIRPS: precipitation datasets are available from the Climate Hazards Group Infrared Precipitation with Station data (CHIRPS) product (http://chg.geog.ucsb.edu/data/chirps/). Starting from 1981, CHIRPS is a 30-year quasi-global rainfall dataset. Spanning 50° S to 50° N (and all longitudes) it incorporates 0.05° resolution satellite imagery with in-situ station data to create gridded monthly time series for seasonal drought monitoring and water resource management (Funk et al., 2015 [3]).

Downloaded data belong to global daily dataset (version 2) with 0.05° resolution and are composed by yearly daily series, each one named *chirps-v2.0.<year>.days_p05.nc* (e.g. *chirps-v2.0.2001.days_p05.nc* for year 2001). Before starting the process, it takes to specify the folder path containing all yearly files, the whole period covered (***Start year*** / ***End year***) – all years must be present within this interval – and aggregation type (***File to create***), which may be annual (maximum of daily precipitation) or monthly (total precipitation OR maximum of daily precipitation).

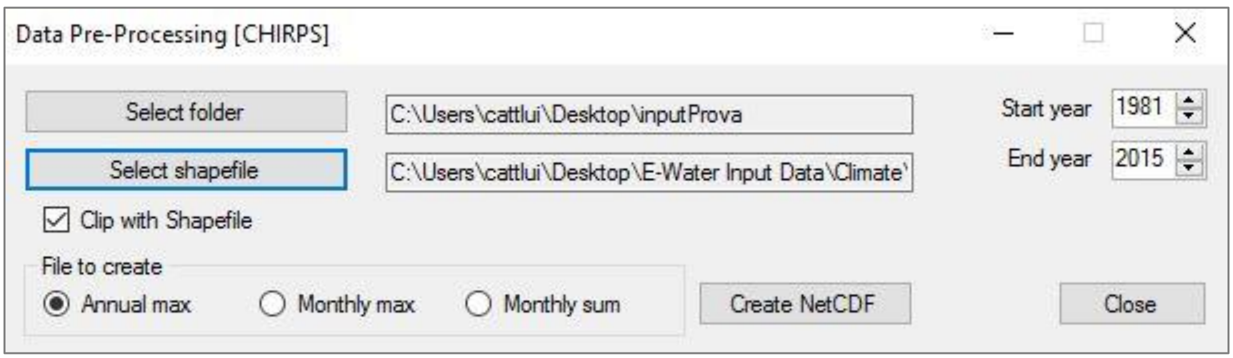


Figure 36: pre-processing interface for CHIRPS datasets

ERA Interim: temperature datasets are available from the ERA-INTERIM (http://apps.ecmwf.int/datasets/data/interim-full-daily/levtype=sfc/), the dataset of reanalysis providing hydro meteorological variables such as maximum and minimum temperature across land at various temporal scales (Dee et al., 2011 [4]). Reanalysis has been increasingly used to address a variety of climate-change issues and has by now become an important method in climate change research. ERA-Interim is a reanalysis product of the European Centre for Medium-Range Weather Forecasts (ECMWF) available from 1979 and continuously updated in real time. The data assimilation system used to produce ERA-INTERIM is based on a 4-dimensional variation scheme (4D-Var) with a 12-hour analysis window.

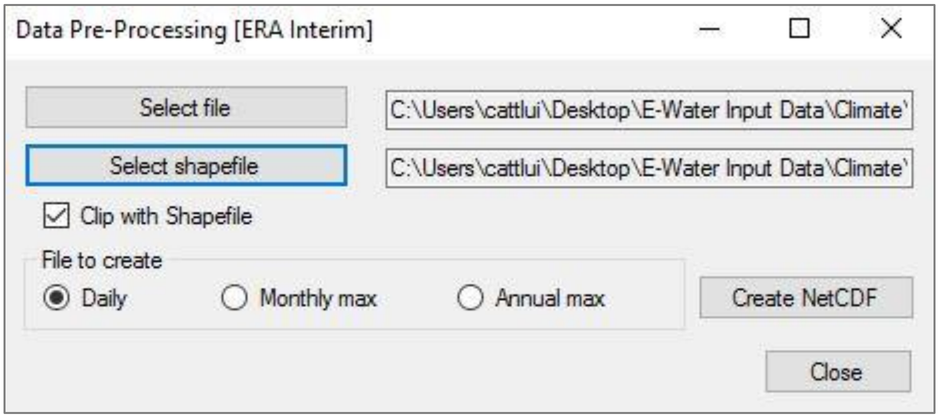


Figure 37: pre-processing interface for ERA Interim datasets

The reanalysis dataset used in this case has a spatial resolution of 0.75° (i.e., approximately 80 km at the equator), and a temporal resolution of one day (i.e., a time step of 24 hours). ERA-Interim allows a consistent spatial and temporal resolution by incorporating millions of observations into a stable data assimilation system that would be nearly impossible for an individual to collect and analyse separately.

Expected input is a single file containing daily temperatures (maximum or minimum) over the whole period, and use must define output aggregation from ***File to create*** box (daily maximum, monthly maximum or annual maximum).

8.2 General settings

Save/load configuration: this feature (upper part of the menu) allows saving the whole process configuration (all data explained in this paragraph) with a name and reloading it afterwards by selecting it from a drop-down menu.

Climate data source: there are two possible input formats for climate processes: ***NetCDF*** local files or ***Database*** tables.

It is possible to use *NetCDF* files (*.nc*) providing climate data with time and space (longitude/latitude) references. The whole time series must be contained into a single file, so usually it takes to execute ***Input pre-processing*** first. When this input type is selected, user can load desired file through ***Select file*** browser.

Else, if desired input is a database table, the list panel below becomes active input selection is performed by row selection (each one corresponds to a distinct climate dataset). Other than table name, there are information about climate field (temperature or precipitation), measure unit, start/end dates and time step (annual, monthly or daily). In case of many tables available, user can narrow his search by giving a keyword, or specifying ***Field*** and ***Time step*** of input data. In order to be available from this list, climate data have to be imported into database first, through ***Import*** menu.

Once input data have been chosen, it is possible to set other parameters in order to evaluate properly desired climate index.

Time selection: by setting ***Start date*** and ***End date*** of input time series, user defines the interval used for evaluating the index. If data are taken from database, these fields are automatically set (time interval has been already evaluated during import phase) but user can change them to focus the ongoing process on a smaller interval within the original boundaries. Else, if *NetCDF* files are used, user has to provide manually such information and the interval must be the *whole* period covered by the dataset.

Input shapefile: the shapefile (*.shp*) provided through this file path determines the geographic area where the index has to be evaluated and focus the analysis within its boundaries. In addition, its polygons are printed on maps for spatial representations. A set of shapefiles (Mékrou Basin area) is available in *Shapefiles* folder (from E-Water documents folders).

Output folder: user can define where to save output data produced at the end of process. All created files are stored into a folder named after evaluated index and the when it has been performed (e.g. ***heat_waves_23_10_2017_00_00_00***).

Shapefile menu: it needs to be set <u>only</u> for **SPI** index evaluation (check its description for further information).

8.3 Data Import (climate database)

By opening this menu (Figure 38), it is possible to import *NetCDF* files into database as *PostgreSQL* tables (schema ***climate***), and manage existing ones. The list panel on the top shows already available tables (it is actually the same shown on ***Climate*** menu), which can be selected by user (table search can be filtered with a keyword, or using ***Time step*** and ***Field*** filters) and permanently removed from database (***Delete***).

Input file: first, a local *NetCDF* file needs to be selected by providing its path through the folder browser. Then, E-Water automatically sets other fields in order to fit further settings the best way possible. The interface is designed to work with files created with ***Input pre-processing*** algorithm, otherwise import success is not guaranteed.

The ***Coordinate Reference System*** (***CRS***) describes how geographical information are provided from input file and is automatically found and displayed with its related Spatial Reference Identifier (***SRID***). If not properly found, E-Water assumes by default a *WGS 84* coordinate system (SRID 4326).

Time selection: user has to specify exactly the period covered by input dataset, by filling ***Start date*** and ***Time step*** fields, whereas the end date will be calculated out of these values. Dates are <u>always</u> represented at daily scale, even if belonging to monthly or annual datasets.

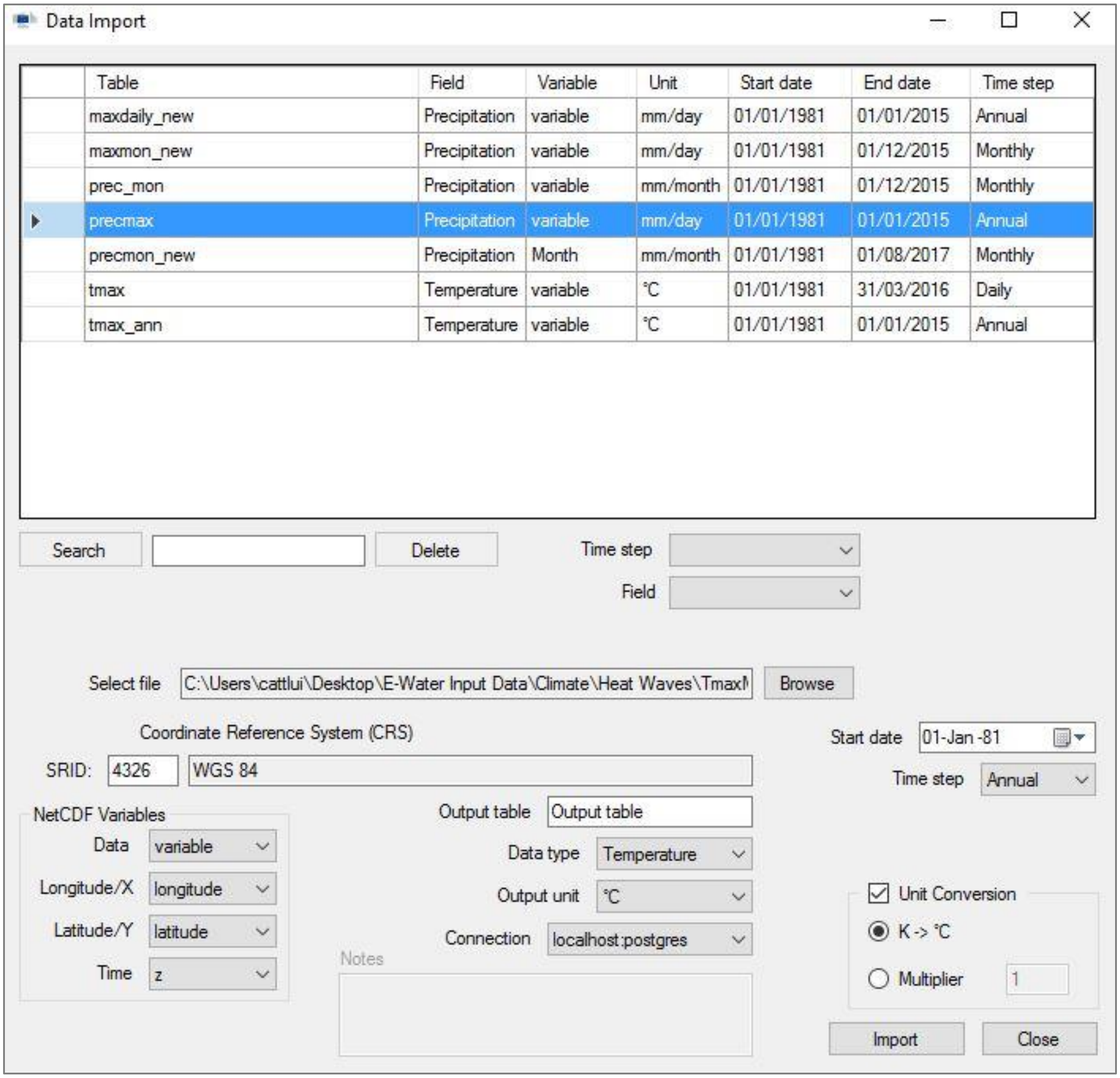


Figure 38: menu for climate data import

NetCDF Variables: in this section, it is possible to define the kind of information provided by variable fields found in input file. When a file is selected, E-Water performs an automatic match by reading fields' name, but user can modify these settings if they are wrong. Fields can be set through the following drop-down menus:

- ***Data***: field assumed to contained the actual climate values (temperature, precipitation)
- ***Longitude/X***: field with horizontal spatial values (longitude or X, depending on reference system)
- ***Latitude/Y***: field with vertical spatial values (latitude or Y, depending on reference system)
- ***Time***: field with temporal values (dates)

In addition, it takes to fill these fields:

- ***Output table***: the name used for imported data when displayed in table list. If selected name already belongs to an existing table, it will not be accepted.
- ***Data type***: the type of climate data to import (Temperature or Precipitation). It defines the processes available for this specific input data.
- ***Output unit***: measure unit used to classify imported data. In case of conversion, it refers to final converted value (e.g. for Kelvin to Celsius conversions, it takes to select *°C*)
- ***Connection***: database connection for data to import (*host: database name*). From the drop-down menu, it is possible to choose among all connections registered from ***Connection settings*** menu.

Unit conversion: by checking this option, user can apply a conversion to input climate values during import process. ***K->°C*** conversion applies a *-273.15* modifier to each value, while ***Multiplier*** operates a multiplication by the value contained in the box aside (useful for measure scale changes).

After setting up all these parameters, import can be executed (***Import***). At the end of process, a message will declare whether data has been successfully imported or not.

8.4 Specific settings (indices)

Once input has been chosen, along with time and space selection (see **General settings**), user can select the index to evaluate. Depending on input data type (***Precipitation*** or ***Temperature***), there are two different sections, each with their distinctive settings.

If data are taken from database tables, they also provide information about type, so E-Water automatically selects the right menu, while the other is completely disabled until another input selection.

Else, if data come from *NetCDF* file, such selection is completely up to the user.

8.4.1 Precipitation parameters

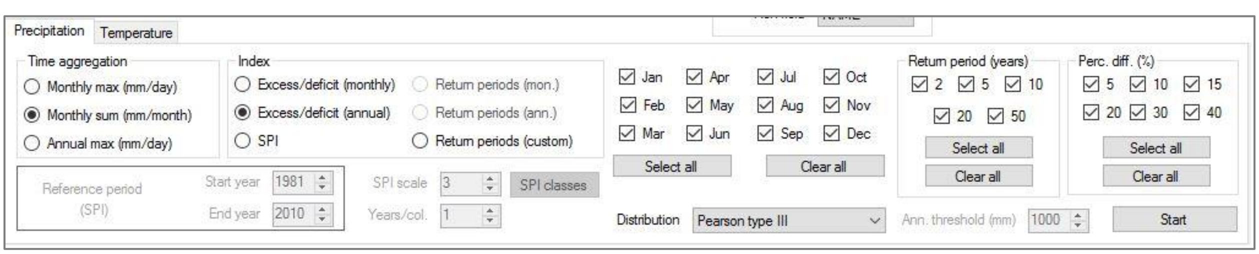


Figure 39: specific parameters for input precipitation data

This section allows evaluating climate indices out of precipitation data (Figure 39). Since processes need a particular type of input, once the time aggregation of input data is specified, the choice of possible indices (Table 20) is narrowed in order to avoid wrong results. After setting all parameters, the process will be executed with ***Start***.

Aggregation	Index
Monthly max. (mm/day)	Return periods (monthly/annual/custom)
Monthly sum (mm/month)	• Excess/deficit (monthly/annual) • SPI • Return periods (custom)
Annual max. (mm/day)	Return periods (annual)

Table 20: precipitation indices, arranged by input aggregations

Time aggregation: user has to specify the aggregation of input data, which may consist in maximum values of daily precipitation (on monthly of annual scale), or monthly total precipitations. This selection is carried out automatically if data source is a database table.

Index: every available index has its own parameters to set, resulting in enabling determined menus in this section. Basically, indices are ***Return periods*** (monthly, annual or custom), ***Excess/deficit*** (monthly or annual) and ***SPI***.

Months: in case of monthly data, index evaluation is performed exclusively for the months selected in this check panel. For instance, by checking only January it is possible to obtain a result where data coming from other months are totally rejected.

Return period: if return periods are involved during an index evaluation, user has to specify one or more return periods (expressed in years) to focus process on. Used for ***Return period*** (monthly and annual) and ***Excess/deficit***.

Percentage Difference: during ***Excess/deficit*** evaluation, user has to choose one or more excess/deficit deviation rates (expressed in percentage) from average value to determine related return periods.

Distribution: this parameter indicates the statistical distribution used for evaluating indices. Default method is *Pearson Type III*.

SPI parameters: these parameters are enabled only if selected index is ***SPI***:

- ***Reference period*** indicates the total period (from ***Start year*** to ***End year***, usually around 30 years) which has to be taken as reference from input data. For a proper evaluation of the index, it should not coincide with the input period chosen from general settings.
- ***SPI scale*** is the length (number of months) of the moving window within the year where SPI is effectively evaluated.
- Through ***SPI classes*** menu (Figure 40), user can select its desired value characterisation of every SPI class by choosing default sets (Standard or Agnew) or even entering fully customised values. Classes must always be seven, from ***Extremely Dry*** (lowest) to ***Extremely Wet*** (highest). In addition, it is possible to import and export such values through ***Save***/***Load*** functions in form of *.csv* tables.
- ***Years/column*** simply indicates the number of years contained in a single row of the output resume table. This parameter does not affect the effective algorithm but only its graphic representation.

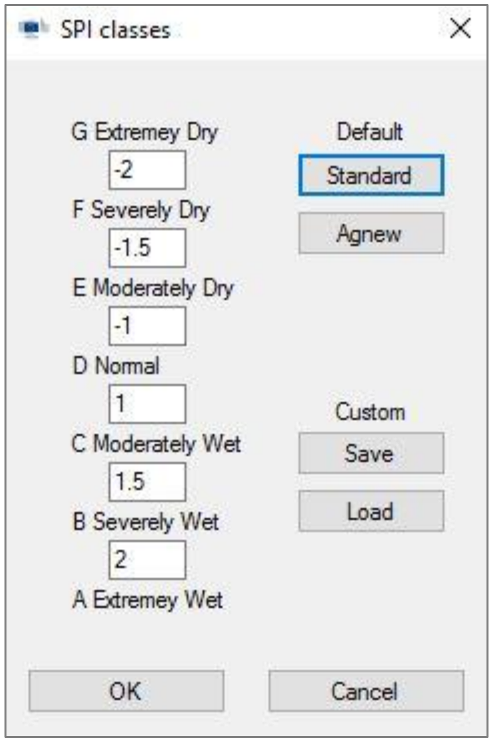


Figure 40: SPI classification menu (set to Standard values)

Annual threshold: this value indicates for which annual precipitation threshold (expressed as mm per year) custom ***Return periods*** are to be evaluated.

8.4.2 Temperature parameters

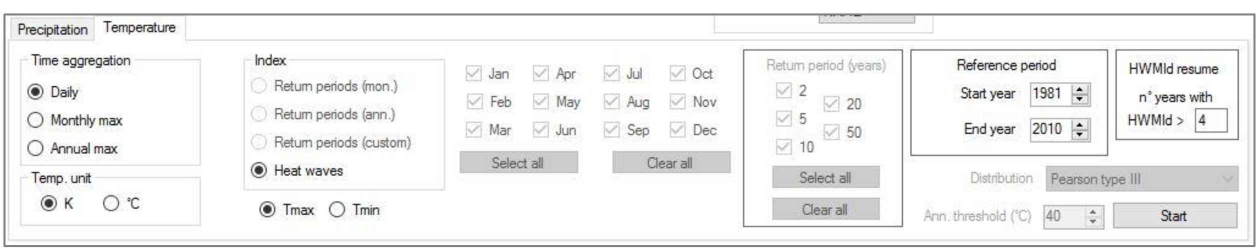


Figure 41: specific parameters for input temperature data

This section (Figure 41) allows evaluating climate indices out of temperature data. Since processes need a particular type of input, once the time aggregation of input data is specified, the choice of possible indices (Table 21) is narrowed in order to avoid wrong results. After setting all parameters, the process will be executed with ***Start***.

Aggregation	Index
Daily	Heat waves
Monthly max.	Return periods (monthly/annual/custom)
Annual max.	Return periods (annual)

Table 21: temperature indices, arranged by input aggregations

Time aggregation: user has to specify the aggregation of input data, which may consist in daily values, or monthly or annual maximums. This selection is carried out automatically if data source is a database table.

Temperature unit: user has to specify whether input temperature values are in Kelvin (***K***) or Celsius degrees (***°C***). This selection is carried out automatically if data source is a database table. Output data is always represented in Celsius degrees.

Maximum/**minimum**: the specification of input temperature type – maximum (***Tmax***) or minimum (***Tmin***) - is needed only for labelling purposes, without affecting index evaluation process.

Index: every available index has its own parameters to set, resulting in enabling determined menus in this section. Basically, indices are ***Return periods*** (monthly, annual or custom) and ***Heat waves***.

Months: in case of monthly data, index evaluation is performed exclusively for the months selected in this check panel. For instance, by checking only January it is possible to obtain a result where data coming from other months are totally rejected.

Return period: if return periods are involved during an index evaluation, user has to specify one or more return periods (expressed in years) to focus process on. Used for ***Return period*** (monthly and annual) and ***Excess/deficit***.

Heat waves parameters: these parameters are enabled only if selected index is ***Heat waves***:

- ***Reference period*** indicates the total period (from ***Start year*** to ***End year***, usually around 30 years) which has to be taken as reference from input data. For a proper evaluation of the index, it should not coincide with the input period chosen from general settings.
- In order to obtain an ***HWMId resume***, it is necessary to define a minimum threshold to evaluate the number of years characterised by an HWMId above this value.

Distribution: this parameter indicates the statistical distribution used for evaluating indices. Default method is *Pearson Type III*.

Annual threshold: this value indicates for which annual temperature threshold (in Celsius degrees) custom ***Return periods*** are to be evaluated.

8.5 Indices description

8.5.1 L-Moments

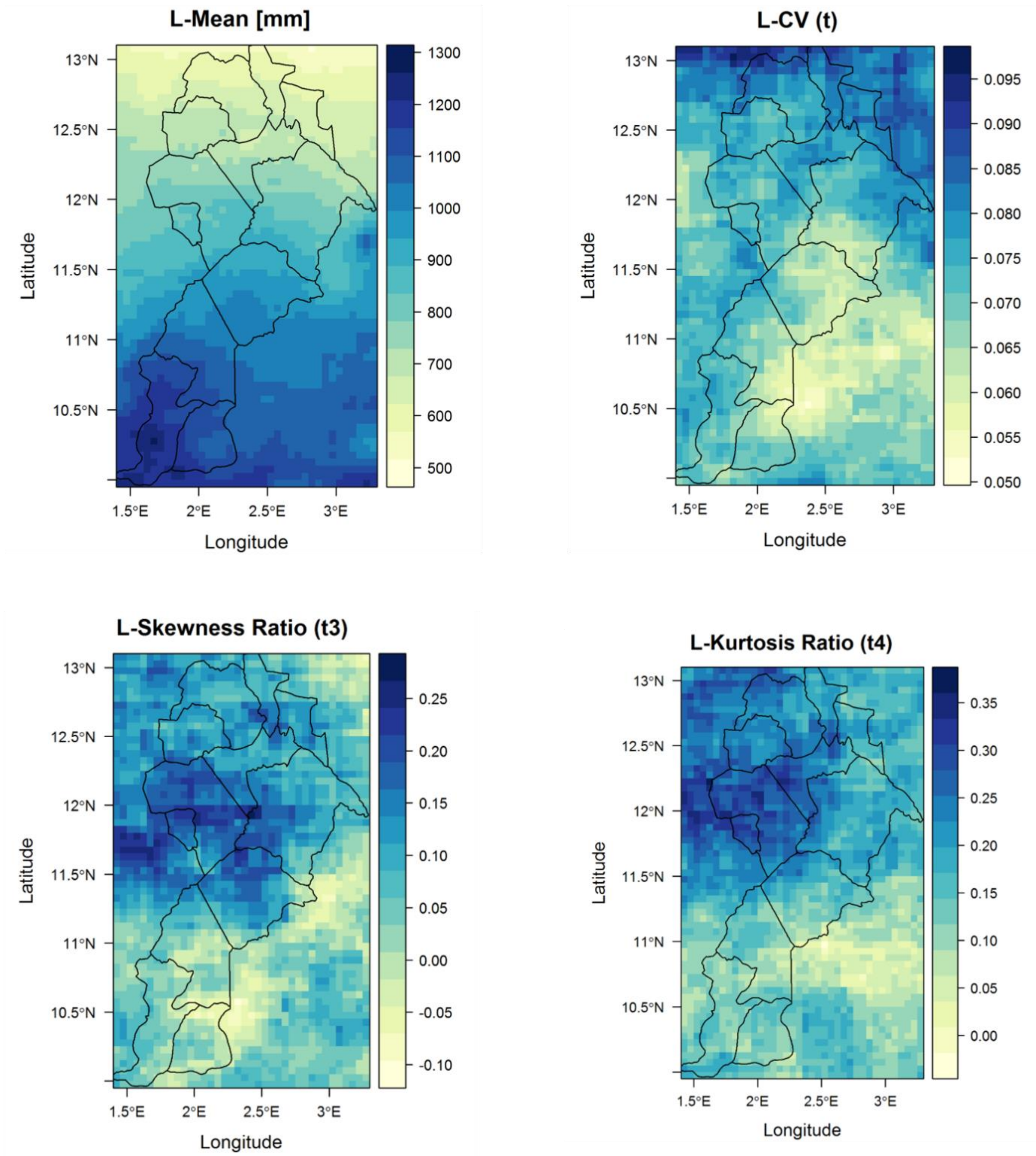


Figure 42: output maps of precipitation L-Moments

Given the statistical nature of climate indices, it is very useful to evaluate factors who synthetically characterise the behaviour of the distributions, better described as L-Moments. More specifically, during most processes (all but ***SPI***) E-Water also evaluates the four main L-Moments of input variable (mean, scaling, skewness, kurtosis) (Figure 42). All output temperature data are expressed in Celsius degrees, but it takes to specify whether input temperature unit is in *K* or *°C* degrees (***Temp. unit*** entry).

8.5.2 Return periods (annual, monthly)

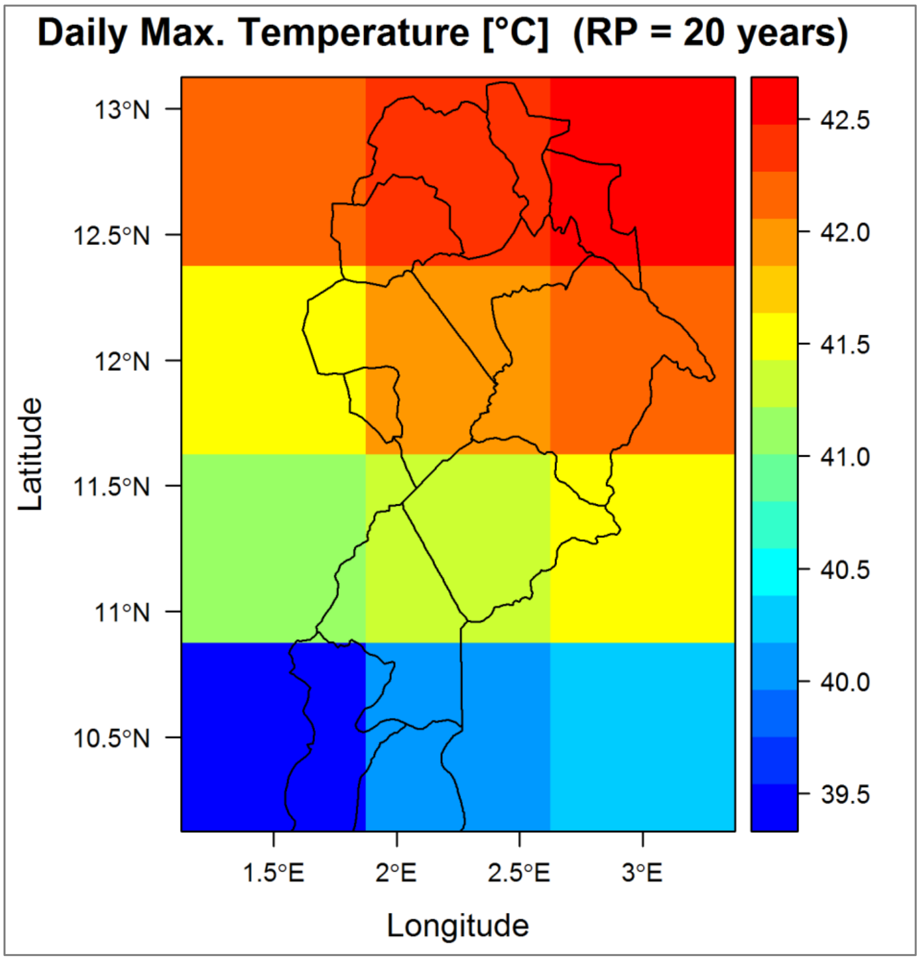


Figure 43: daily max temperatures registered with a return period of 20 years

The representation of the likelihood of receiving a specific amount based on a given time series is best accomplished by fitting a parametric statistical distribution, using a continuous function allowing for a comprehensive analysis of the climate value based on the acquired sample. With the available modelled history of climate data, it is possible to fit parameterized statistical distributions to the data, but first an appropriate distribution must be selected. Much research has been carried out related to fitting and evaluating statistical distributions for rainfall. All maxima have been modelled using the generalized extreme value (GEV) distribution. GEV is a family of continuous probability distributions developed within extreme value theory to combine the Gumbel, Fréchet and Weibull families also known as type I, II and III extreme value distributions. By the extreme value theorem, the GEV distribution is the only possible limit distribution of properly normalized maxima of a sequence of independent and identically distributed random variables. Note that a limit distribution need not exist: this requires regularity conditions on the tail of the distribution. Despite this, the GEV distribution is often used as an approximation to model the maxima of long (finite) sequences of random variables.

Return periods are available for both precipitation and temperature data: they evaluate annual/monthly maxima of input values, related to all checked return periods in ***Return Period (years)*** menu (from 2 to 50 years) (Figure 43). In case of ***custom*** index, the process is inverse: starting with a fixed threshold value (mm for precipitation, °C for temperature), E-Water will evaluate two return period maps, relative to values falling over and under this threshold. When this index is evaluated for monthly precipitations (mm/month), it will be applied to the total amount over all checked months, so if all 12 of them are selected the return periods will be related to the total amount of annual precipitation.

8.5.4 Excess/deficit (annual, monthly)

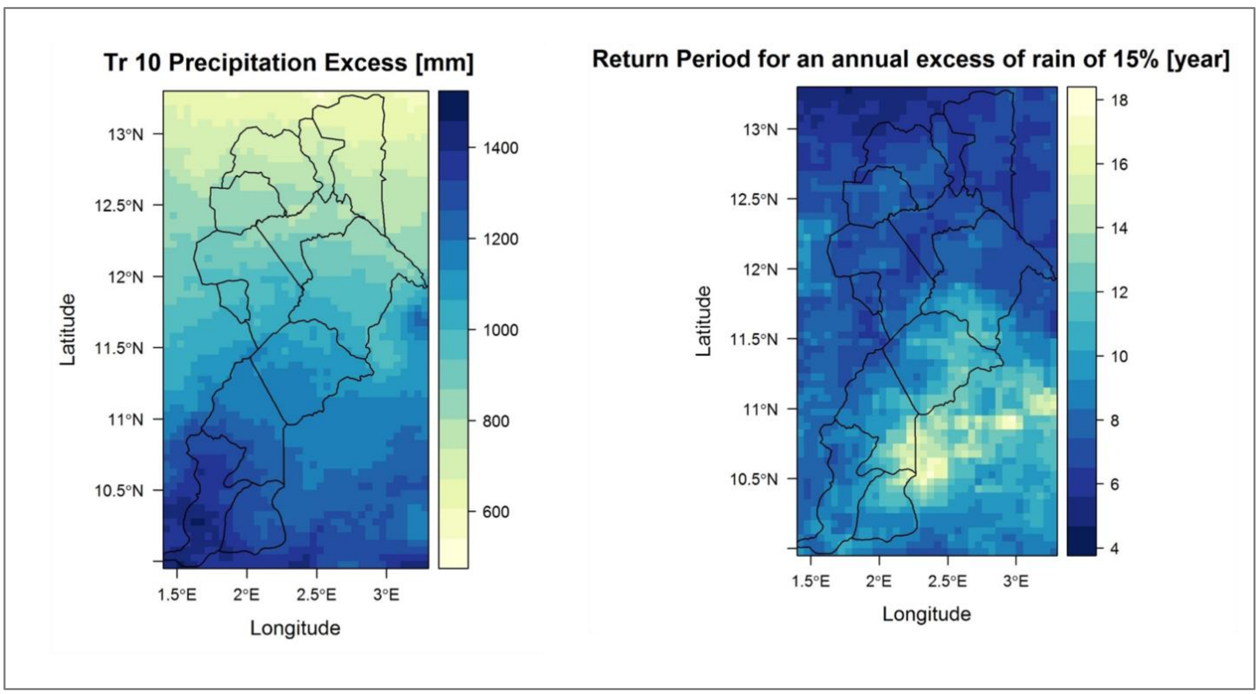


Figure 44: annual precipitation Excess with a return perios of 10 years (left) and returnperiods corresponding to annual precipitation excess of 15%

During the ***Excess/Deficit*** evaluation process, E-Water evaluates relationship between the deviation rate (expressed in percentages and absolute values) of precipitation values over their average, both positive (excess) and negative (deficit), and the return periods corresponding to such variations. In fact, the output contains excess and deficit of monthly or annual precipitation at selected return periods (from 2 to 50 years), along with return periods registered with each selected percentage of excess/deficit (from 5% to 40%). ***Monthly*** index is evaluated for every selected month, so there is a distinct result for each one of them (excess/deficit of January precipitation), whereas ***annual*** index consists in a single analysis on annual sums of precipitation collected from checked months (Figure 44). By default, statistical analysis of precipitation is made using the Pearson-III distribution, a good choice for describing cumulated values because turning to be more flexible and allowing a better fit to any number of rainfall regimes with reasonable accuracy.

8.5.6 SPI

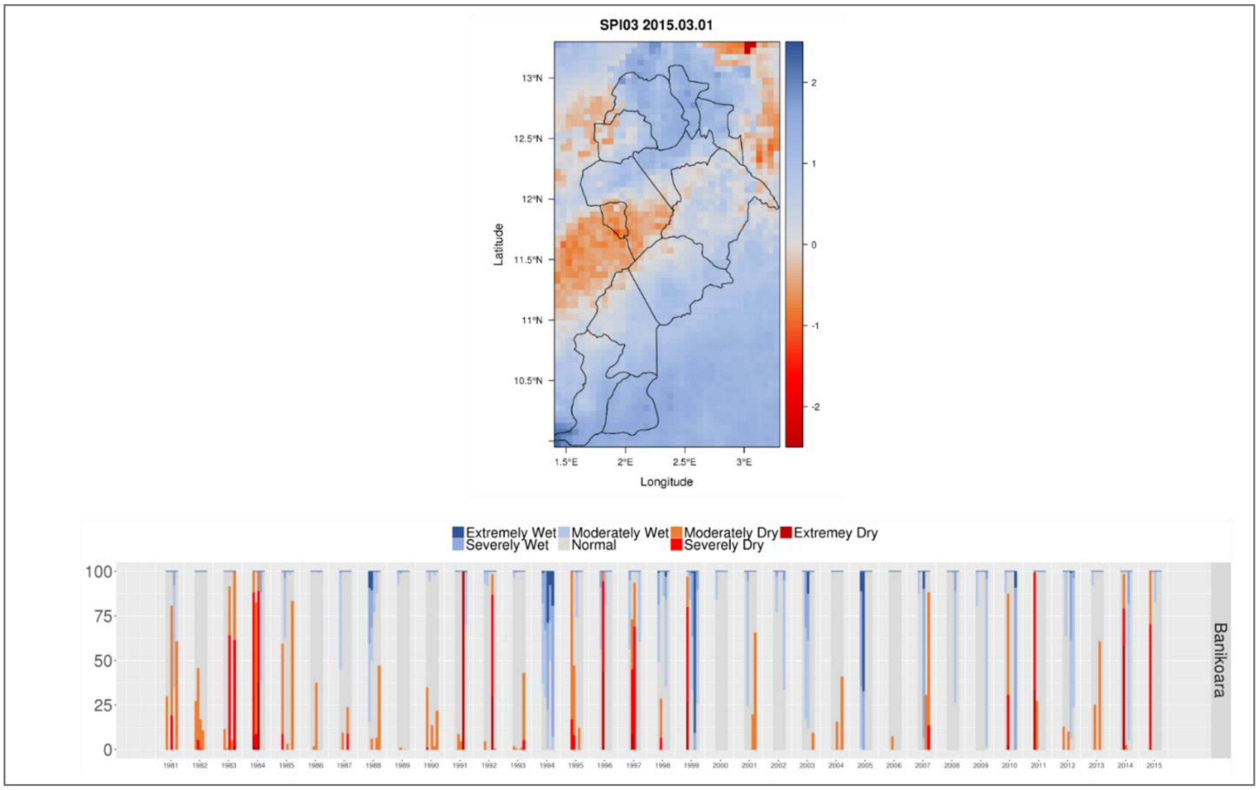


Figure 45: SPI map with to 3-months window in March 2015 (top) and yearly distributions of SPI classes in Banikoara region from Mékrou area (bottom)

The Standardised Precipitation Index (***SPI***) allows to evaluate phenomena of precipitation anomalies and droughts affecting geographical domains. It can be used along with other indices in order to estimate dry spells for hydrological and agricultural studies. More specifically, SPI is a probability index based on monthly precipitations intended to provide information about anomalies of humidity and/or dryness. Since it is useful to compare it through multiple regions characterised by different climate conditions, by defining the region field name (***Ref. Field***) of input shapefile it is possible to obtain a distinct SPI analysis for each one of them. If normalised, it allows to evaluate a drought occurrence rate. SPI can be calculated for different periods (generally, 1 to 36 months) taken from monthly input data (e.g. SPI-3 from March covers the period January - March) by setting the ***SPI scale*** parameter to the desired number of months (e.g. 3 for SPI-3).

SPI evaluation is carried out through the following three steps:

1. Using L-Moments, a probability density function (PDF) is set up according to the frequency distribution of monthly precipitations and the time scale of interest.
2. Reference period (usually 30 years) is divided into 12 parts, each per month. Every data fraction, containing one monthly value per year, is then adjusted with a custom statistical distribution.
3. Cumulated probability undergoes an inverse transformation of its standardised Gaussian distribution, thus giving SPI a description in terms of difference from temporal average.

Since its time-space invariance, different SPI instances can be directly compared regardless of the climate zones they come from and their related annual period.

Wet/dry class	**SPI values (McKee et al., 1993 [7])**	**SPI values (Agnew, 2000 [8]).**
Extremely Dry	<-2	<-1.64
Severely Dry	[-2,-1.5)	[-1.64,-1.28)
Moderately Dry	[-1.5,-1)	[-1.28,-0.84)
Moderate	[-1,1)	[-0.84,0.84)
Moderately Wet	[1,1.5)	[0.84,1.28)
Severely Wet	[1.5,2]	[1.28,1.64]
Extremely Wet	>2	>1.64

Table 22: default SPI classifications

As shown on Table 22, SPI values can be grouped into different SPI classifications (see ***SPI classes*** menu), depending on climate conditions characterizing selected area. If not modified, E-Water uses by default the Standard classification (McKee).

At the end of process, SPI values are divided by region and month (only the ones checked in month panel) then grouped into selected classification. Finally, they are displayed in a diagram showing each monthly SPI class percentage (%) for every region: the image is arranged as a grid where each row contains all monthly analyses for a single region. Other than that, E-Water produces a set of maps displaying SPI values over the whole domain for every single month within selected period (Figure 45).

8.5.5 Heat waves

The *Heat Wave Magnitude Index daily* (HWMId), recently defined by Russo et al. ([5]), is a simple numerical indicator that takes both the duration and the intensity of the heat wave into account. Basically, the magnitude index sums excess temperatures beyond a certain normalized threshold and merges durations and temperature anomalies of intense heat wave events into a single indicator, according to the methodology described in [5] and [6]. The HWMId is an improvement on the previous Heat Wave Magnitude Index (i.e., HWMI, Russo et al. [2]) and it is able to overcome its limitations. More precisely, HWMI has some problems in assigning magnitude to very high temperatures in a changing climate, thus resulting in an underestimation of extreme events.

The HWMId is defined as the maximum magnitude of the heat waves in a year. Specifically, a heat wave is a period of three or more consecutive days with maximum temperature above a daily threshold calculated over a reference period (usually 30 years). The threshold is defined as the 90th percentile of daily maxima temperature, centred on a 31-day window. The inter quartile range (IQR, the difference between the 25th and 75th percentiles of the daily maxima temperature) is used as the heat wave magnitude unit, since it represents a non-parametric measure of the variability. If a day of a heat wave has a temperature value equal to IQR, its corresponding magnitude value will be equal to one. According to this definition, if the magnitude value on a given day is *3*, it corresponds to an anomaly of 3 times the IQR. Starting with daily maximum temperatures, E-Water will produce a set of spatial maps containing HWMId yearly values (one for year) and a resume map with the number of years characterised by a HWMId value exceeding a threshold selected by user (Figure 47).

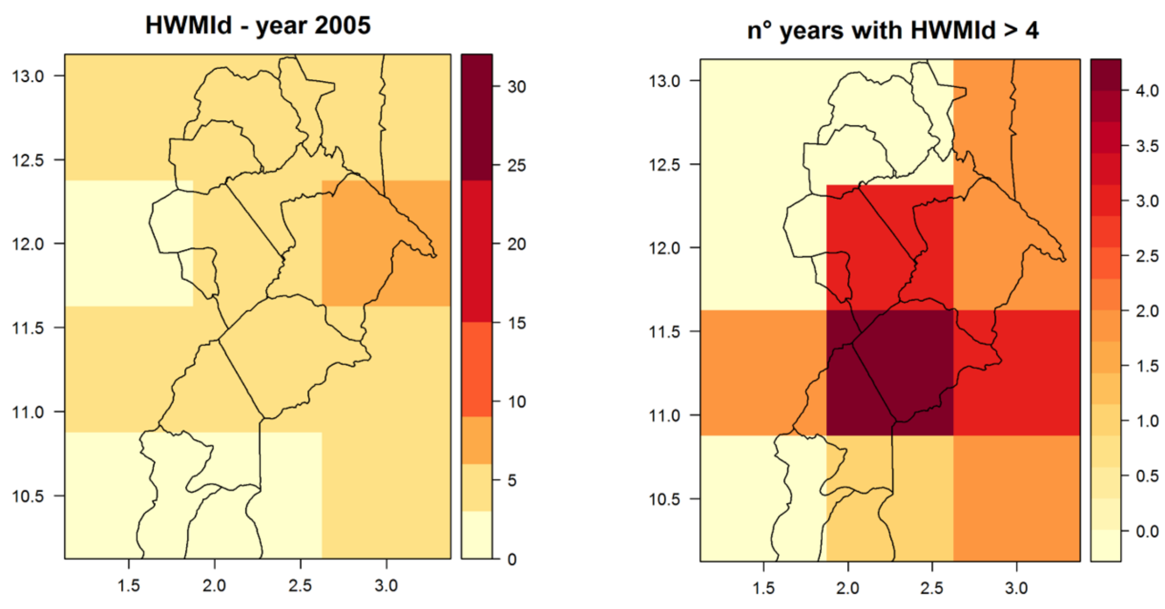


Figure 46: HWMId map on year 2005 (left) and resume map (1981-2015) with the number of years with HWMId values higher than 4 (right)

9 Socio–Economics

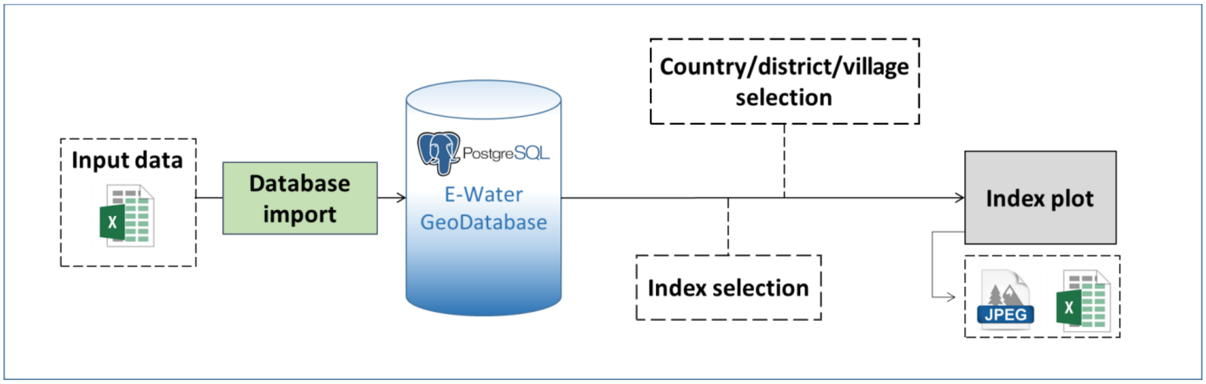


Figure 47: process diagram of Socio-Economics menu

This part of the module interface is focused on the socio-economic information (water sources for the local population and their location, habits of the people and an overall view about their perceptions on water related issues) gathered among the villages within the area. Data have been collected through interviewing the local populations and each sample is related to a single household composed by a variable number of people. Each information index is broken down to country level, but there are also values available for the total area. Differences between different countries are more obvious in some surveys, and this helps identify different issues or perspectives among the population of the river basin. From the input menu, it takes to select which index to display and the geographical areas to be included in the analysis. Results are shown as histograms with absolute or percentage values, and can be saved as images or documents (Figure 47).

9.1 Data import

First, from the Import menu it is possible to select an *.xlsx* file containing all data to analyse, where all indices are arranged per column and each survey sample is identified by a unique number and a set of addressing information (country/district/village), and to import it into the database.

Reference fields: once an input file is chosen, it takes to properly identify all his reference fields (columns) for a correct import of all its data. Selection is made through drop-down menus, all filled with the column headers found into selected ***Sheet*** field (this one containing all sheet names from input file). Rather than a specific index, these fields provide information about the source of every piece of data (identification, address). ***ID field*** is the main identifier, since it contains a set of unique integer values, each related to a single sample household. ***Country*** , ***District*** and ***Village*** give instead a more accurate description of the samples with a three-level localisation starting from the most general information (country) and then proceeding with more specific ones (district, and finally village). The latter fields appear only in one sheet that must be indicated as ***Sheet*** field, while the ***ID field*** is instead contained in all of them. After the right fields are selected, ***Load*** will save the current configuration.

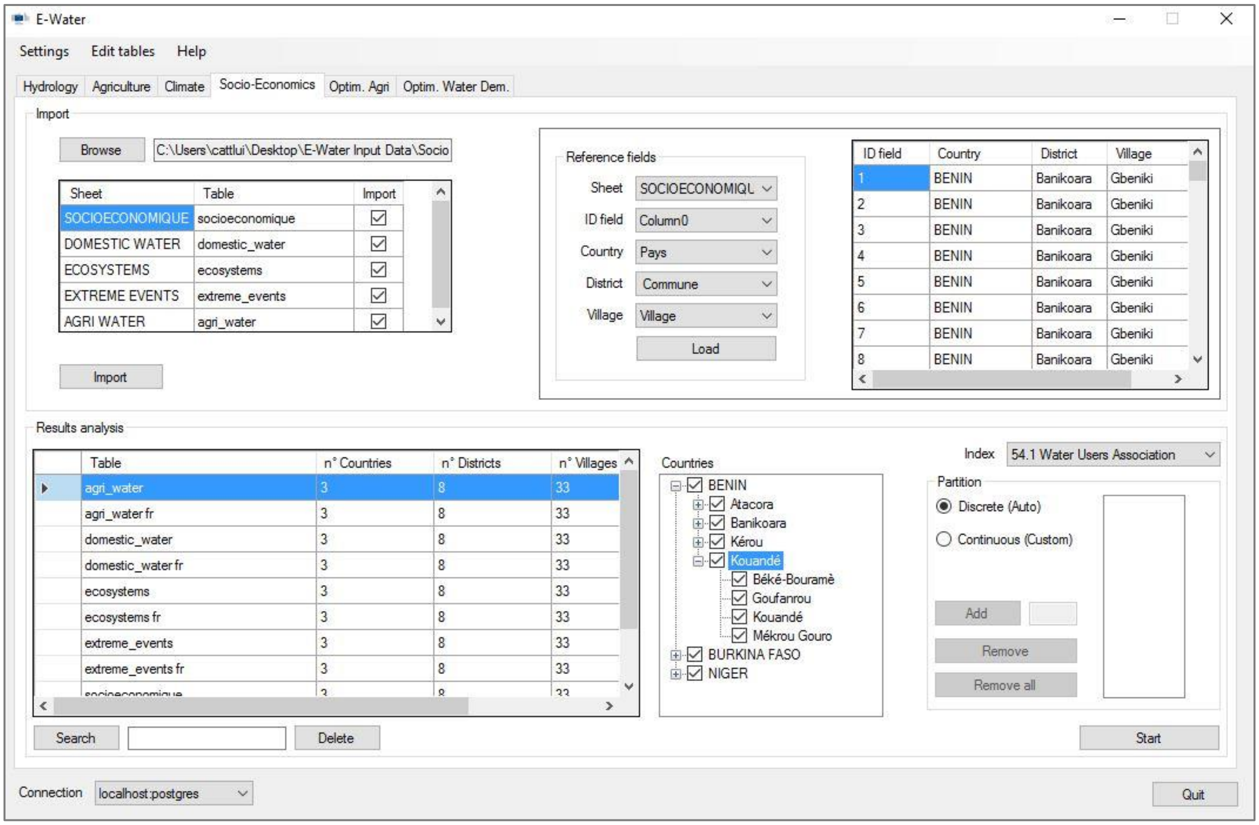


Figure 48: Socio-Economics menu

Import list: the panel on the left side contains a list of all names of the sheets contained in input file under ***Sheet*** column. Through this list, by checking ***Import*** field, user can select which sheet to import with all his indices, each into a database table, which name can be edited in ***Table*** field. Once finished, it is possible to start ***Import*** process.

9.2 Results analysis

At this point, user can select the imported table of input values from the top list (number of sample values are also shown) and decide which geographical area (village, district, country) to include in the histogram evaluation, other than the index of interest: in this way, results are broken down to country level, but there are also values available for the total area. Countries and indices involved in the process are not predetermined, but they are simply read from the input file, so the analysis can be carried out regardless of the geographic area and the nature of results to evaluate. Differences among the countries are more obvious in some surveys, and this helps identify different issues or perspectives among the overall population.

Country: on the tree view menu, it is possible to see all countries (represented as upper nodes) with all their related districts and villages (lower nodes). Only index values associated to checked elements of this list will be displayed in the histogram plot.

Index: from this drop-down menu, it is possible to select one index among the whole set provided by selected table. It will be displayed on histogram plot with ***Start***.

Partition: socio-economics indices are expected to have two kinds of possible values. In the first case, index value can be fully described as a finite set of responses (e.g. gender of interviewed persons), each identified with an integer for practical reasons. In the second one, index has a real numerical significance (e.g. household income), so potential values are nearly infinite and may differ one each other. Depending this feature, user can decide to operate an automatic classification (***Discrete***) consisting on simply grouping all responses with the same identical values, or determining a set of threshold steps (first check ***Continuous***, then ***Add*** all desired values,) in order to form a set of value intervals, each one gathering results falling within its boundaries.

Finally, histogram plot is presented in a dedicated interface, allowing to edit it in the best way possible before exporting it. Its results are grouped by each different country (represented with a colour), but also total amount is shown. On the upper panel, it is possible to modify histogram colours related to countries, as also displayed on a legend (optional, editable on ***Legend*** menu on the lower right).

Results are shown as histograms (Figure 49) with ***Absolute*** or ***Percentage (%)*** values, whereas the analysed index may require either a continuous scale of values, which is divided into custom intervals (e.g. age) to be inserted from user, or a discrete set automatically determined by the interface (e.g. educational level). If displayed index is contained in internal index description files of E-Water, its values will be described with a definition of one or more words rather than a simple number (only for ***Discrete*** partition). Partition menu is also here, so further modifications can be made in case original configuration has not been approved.

Outputs can be exported as images or documents (***Save as image/document***).

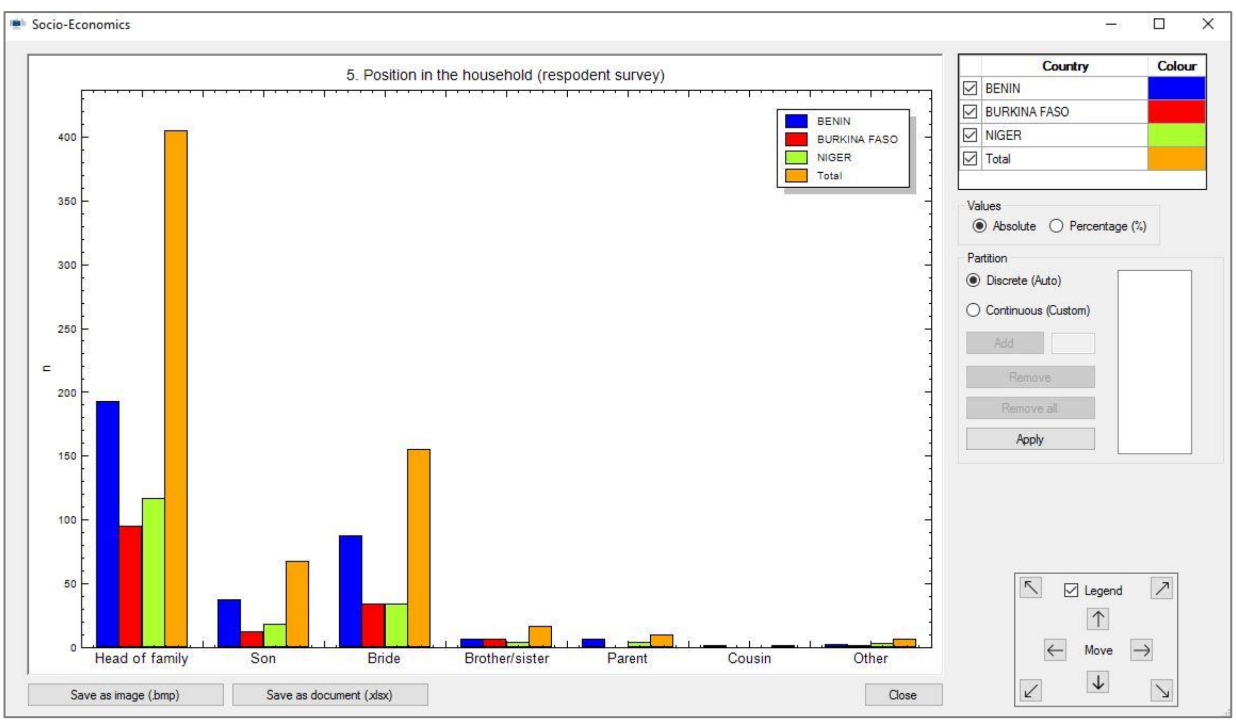


Figure 49: histogram interface for Socio-Economics indices

10 Optimisation: Agriculture

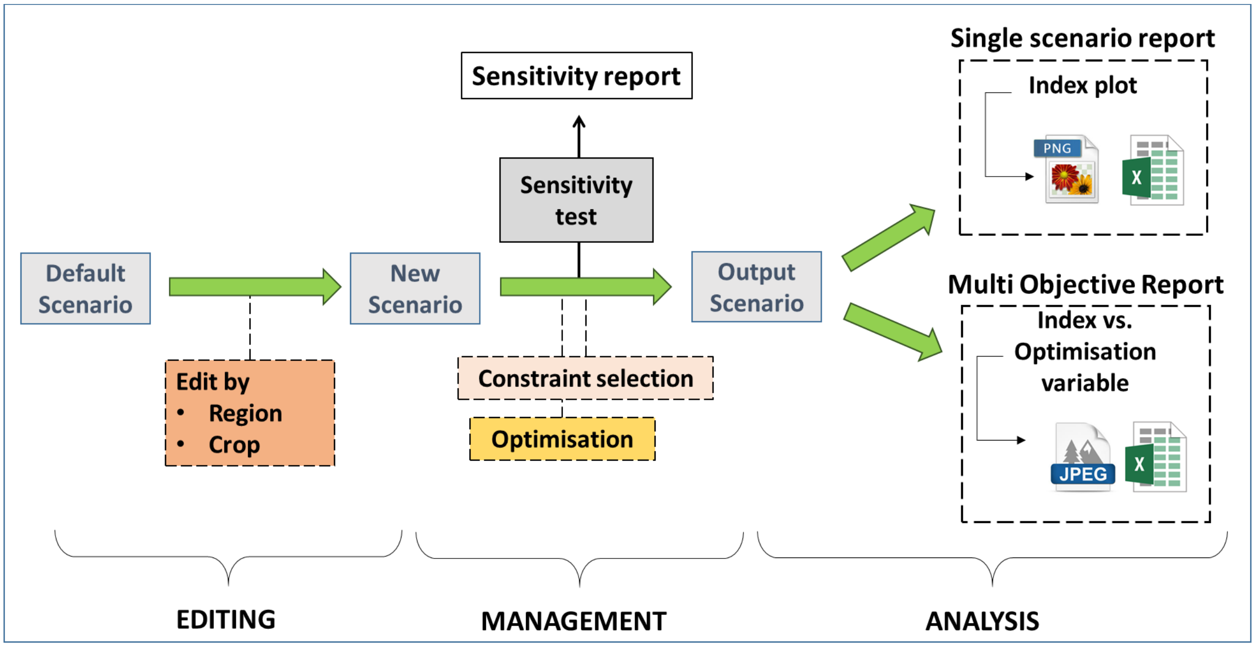


Figure 50: process diagram of Agriculture Optimisation menu

Through ***Optim.–Agri.*** Menu (Figure 51), E-Water allows the execution of an optimization utility where user has a large choice of settings, starting from available data taken from the region of interest (agriculture, population, diet requirements, etc.). Optimization model needs input data from an analysis scenario. In fact, user has to define a set of inputs values and scenario constraints, plus other specific parameters about ongoing process. Currently, there are only 10 available crop varieties for this process (Table 23).

Crop ID	Crop name
CASS	Cassava
CORN	Maize
COTS	Cotton
COWP	Cowpea
PMIL	Millet
PNUT	Peanut
RICE	Rice
SGHY	Sorghum
TOMA	Tomato
YAMS	Yams

Table 23: available crops for agriculture optimisation

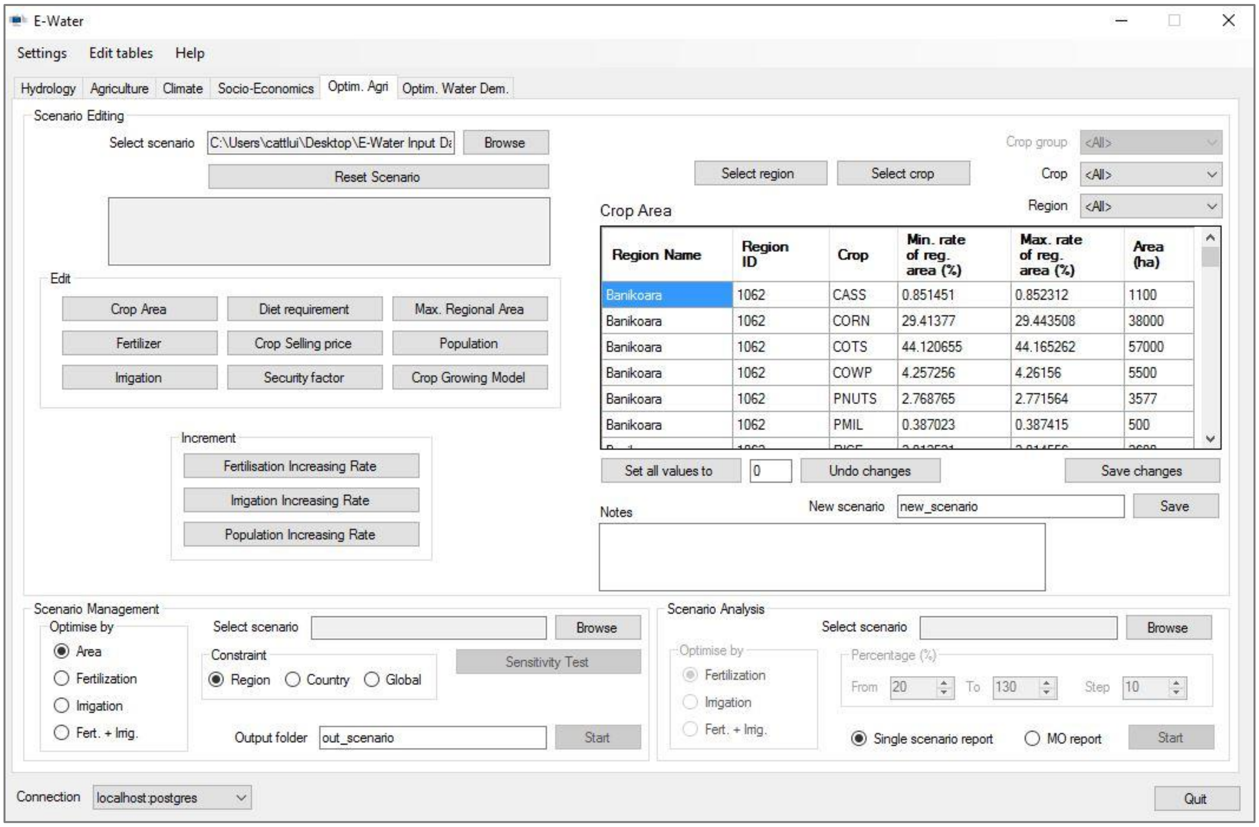


Figure 51: Agriculture Optimisation menu

10.1 Scenario editing

This section is specifically aimed to modify the values of an existing scenario in order to create a customised input for a further optimisation process. In fact, all changes are not applied to the original scenario, but they are instead used to build a new one.

Input scenario selection*: first, it takes to load an available scenario folder containing all parameters needed to run the process. All files are .csv tables containing all available regions and crops along with different kinds of information characterizing them all (*

Table 24). With ***Reset Scenario***, input scenario data will be totally reloaded from its folder, de facto undoing all further changes to its parameters and allowing user to start over the whole process.

Region selection: by entering this tree view panel showing countries (upper nodes) with their regions (lower nodes), it is possible to select which region(s) to involve in the process of optimization by (un)checking them. All parameters belonging to unselected regions will be totally hidden from edit panel of the interface.

Crop selection: similar to previous menu, this tree view panel allows to select which crop to analyse in optimization process. Singular crops are grouped into more general crop groups, which appear as upper nodes.

File name	Related parameter
Area.csv	Crop Area
AreaRegMax.csv	Max. Regional Area
AreaMaxCropRegRate.csv	Crop Area
AreaMinCropRegRate.csv	Crop Area
FertilizerAvailable.csv	Fertilizer
FertilizerMax.csv	Fertilizer
FertilizerMin.csv	Fertilizer
IrrigAvailable.csv	Irrigation
IrrigMax.csv	Irrigation
IrrigMin.csv	Irrigation
Diet.csv	Diet Requirement
CropSalePrices.csv	Crop Selling Prices
Regions.csv	Region list
CropsGroup.csv	Crop list
FertilizerIncreasingRate.csv	Fertilizer Increasing Rate
IrrigIncreasingRate.csv	Irrigation Increasing Rate
DemandIncreaseByCropGroup.csv	-
ExportRatesLimits.csv	-
PopulationByRegion.csv	Population
PopulationIncreased.csv	Population Increasing Rate
SecurityFactor.csv	Security Factor
CropEquations.csv	Crop Growing Model

Table 24: input files from scenario folder, along with their related editable factor

Input data edit: this section allows visualizing and editing all parameters loaded from input scenario files. Clicking on a button with the name of a parameter, its related table will be visualized in on the panel grid on the right. Through related drop down menus (***Crop***, ***Crop group*** and ***Region***), it is possible to visualize all regions and crops, or to focus on one of them. User can manually edit values one by one, or select multiple cells and set them to a single value (***Set all values*** to + value). All changes made must be validated with ***Save changes***, otherwise they will be totally ignored; on the other hand, it is possible to reset the table with ***Undo changes***. It should be remarked that each modification does not affect original files, but is taken into account during the creation of the new scenario folder.

The main data to provide for Optimisation analysis are:

- ***Crop Area***: land use of cultivated surfaces, arranged by village. It is possible to set the effective surface (***Area***) and its maximum and minimum margins of variation percentage (***Min./Max. rate of reg. area***)
- ***Fertilizer***: fertilization rate by crop and village (kg/ha). Other than its base amount (***Current***), it is possible to set its constraints for the optimization (***Min./Max.***)

- ***Irrigation***: irrigation performance by crop and village (mm/ha). Other than its base amount (***Current***), it is possible to set its constraints for the optimization (***Min./Max.***)
- ***Diet requirement***: annual individual food requirements (kg), arranged by crop groups (not singular crops!)
- ***Crop selling price***: crop-selling price (€/kg) for every region are evaluated starting from survey results and FAO statistics. By default, such prices are provided at country scale, but user can focus more in detail and retrieve values for every distinct village.
- ***Security factor***: index to evaluate the real production capabilities of crop groups for fulfilling the population food requirements. The security factor is a percentage value, which verifies the effectiveness of management of vegetal products. Its suggested purpose is turning real production to an effective products availability for a reliable analysis of food security. The percentage indicates effective production by considering potential losses. By default, proposed values for this index are provided by FAO reports.
- ***Max. Regional Area***: maximum area available for every region
- ***Population***: total population of every region
- ***Crop Growing Model***: base crop yield (***Intercept***) and dependence to fertilization and irrigation represented as slope coefficients (***Fertil_coef*** and ***Irrigation_coef***)
- ***(Fertilizer/Irrigation/Population) Increasing Rate***: on these three tables, it is possible to set a percentage increment to fertilization and irrigation for distinct regions and crops, or to population by region.

New scenario: at last, user can save all changes made so far by creating a new scenario containing the updated version of all input scenario files (***Save***). Optionally, he can write some comment or descriptions on the ***Notes*** panel that will be visualized new scenario is used as input for further processes.

10.2 Scenario management

At this point, user can choose which factor to optimise in order to evaluate the impact on food security at a selected scale. Results are saved into output folder as *.csv* tables and can be directly read by E-Water interface.

Scenario selection: if a new scenario has just been created in ***Scenario Editing*** section, its path folder is automatically set as input for the optimization process. Otherwise, user can select any other one.

Optimisation selection: through this menu, user has to choose which scenario factor he wants optimise. Available choices are ***Area***, ***Fertilizer***, ***Irrigation*** or both (***Fert. + Irrig.***)

Constraint selection: another choice is to decide to limit mutual variations of parameters induced by optimisation within ***Region***, ***Country***, or to extend them to ***Global*** scale.

Output selection: before starting optimisation, it takes to define a folder path for optimisation results. E-Water will create a new subfolder (*PlotSol*) where to store all obtained parameters.

Sensitivity test: user can perform a sensitivity test on its scenario with considering the influence of crop prices and their impact on distribution. By comparing percentages of price variations to specific culture surface changes, it is possible to verify how much a certain crop is affected to its selling price. Such report is saved in form of *.csv* tables into a dedicated subfolder (*Sensitivity*).

Optimisation will be executed by pressing ***Start***. After creating output files, E-Water will show an interface to have them displayed in form of plots (see ***Single scenario report*** on next paragraph).

10.3 Scenario analysis

This final section is used to produce analysis reports on optimized scenarios previously created and save them as files (tables and images) into chosen folder. Output interface of ***Single scenario report*** is the same produced at the end of optimization process of previous section.

Scenario selection: if a new scenario has just been created in ***Scenario Management*** section, its path folder is automatically set as input for the optimization process (only main scenario folder should be specified in the path, since the process will directly point to *Plotsol*). Otherwise, user can select any other one. This path is also user as root folder for saving data produced after ***MO report***.

Analysis selection: before starting analysis process (***Start***), user can choose between two different approaches: single or multi scenario report.

Optimisation index	Meas. unit	Grouped by
Area distribution	Ha	Crop
Crop group demand	Kg	Crop group
Optimal fertilizer applied	Kg/Ha	Crop
Infactivilities	Kg	Crop group
Irrigation	mm	Crop
Surplus production (product. - demand)	Kg	Crop group
Total income	€	Crop
Total production	Kg	Crop
Average yield	Tons/Ha	Crop
Total fertiliser	Kg	Crop

Table 25: optimisation indices from single scenario report

Single scenario report: this analysis report synthetises the behaviour of optimised scenario by displaying a set of representative indices (Table 25). All of them are represented as histogram plots (Figure 52) grouped by each distinct crop (or crop group) and with colours associated to different regions. On the upper panel, it is possible to modify histogram colours related to countries, as also displayed on a legend (optional, editable on ***Legend*** menu on the lower right). Outputs can be exported as images or documents (***Save as image/document***).

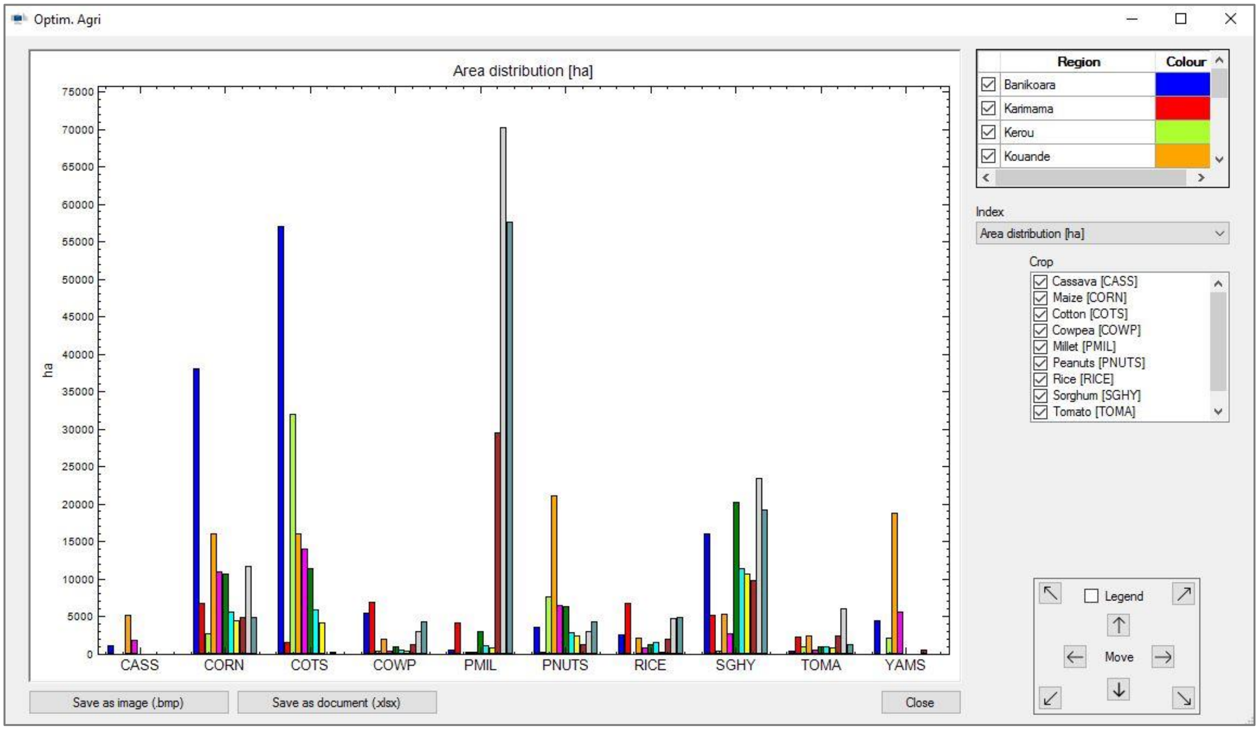


Figure 52: histogram interface for Agriculture Optimisation indices

MO report: the multiple optimisation (MO) process evaluates another set of plots for the comparison of several indicator of the optimisation (for example by comparing change of benefit to the change of fertilization rate and other examples as showed in Figure 53, Figure 54 and Figure 55).

This process consists in comparing obtained scenario with another one with different optimisation criteria. On the ***Optimise by*** panel, user can decide the factor to optimise in the second set, choosing among ***Fertilization***, ***Irrigation***, or both (***Fert. + Irrig.***). Moreover, it takes to define also the percentage range of this optimisation in ***Percentage (%)*** panel, defining extremes (***From*** and ***To***) and ***Step***. In this case, the report is produced in form of plot images and saved into *PlotMOOSol* folder as image *.jpeg* files and *.csv* tables containing plotted values.

Here are plotted indices:

- Benefit increment, indicating production variation in percentage and total values, depending on optimisation factor
- Infactivility diagrams, showing economic impact on optimisation factor variations
- Food demand, verifying if the overall demand is fulfilled (100%) by increasing the percentage of optimisation factor
- Yield increment rate depending on the percentage of optimisation factor

It is easy to notice that images show that the optimised factor chosen by user is represented along the abscissa of output plots.

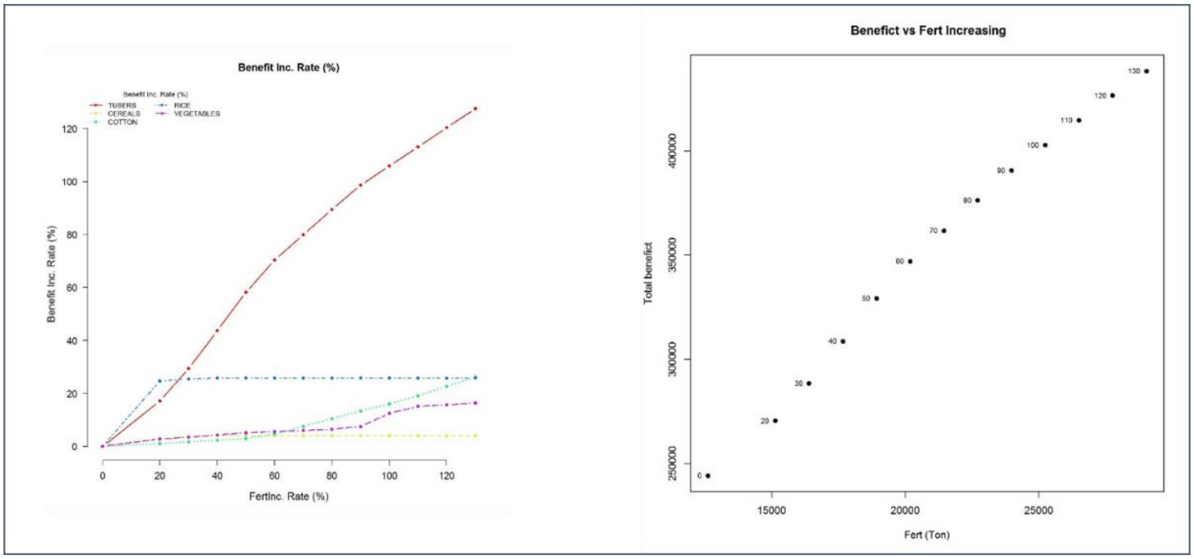


Figure 53: benefit increment index represented with total values (left) end percentages (right), optimised by fertilisation factor

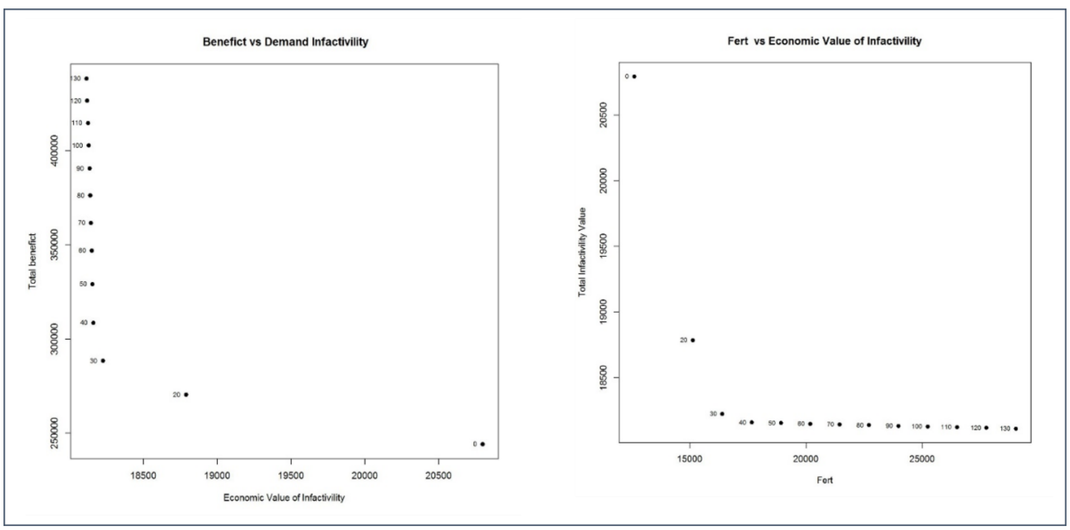


Figure 54: benefit vs Infeasibility diagram (left) with its inverse function (right), optimised by fertilisation factor

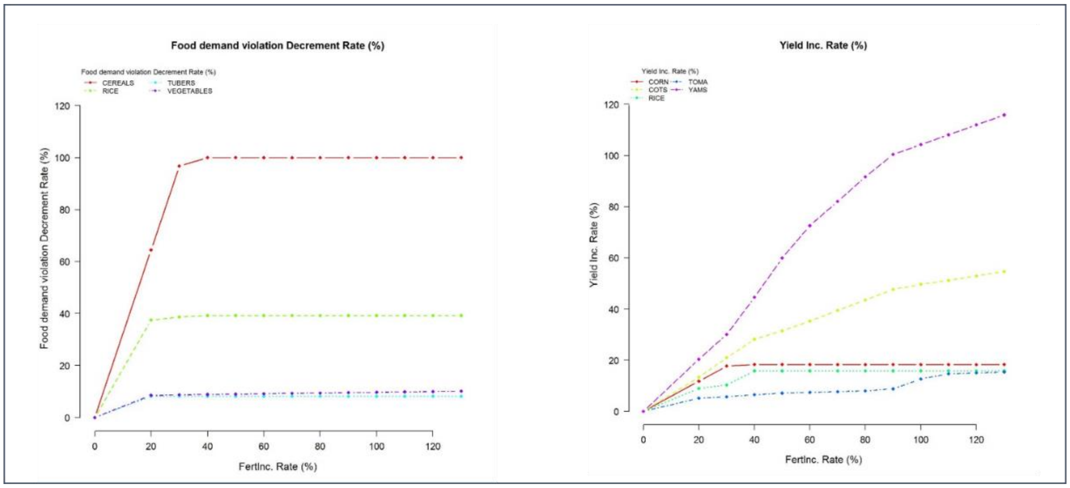


Figure 55: missing quantity for food demand satisfaction indicator (left) and crop yield increment rate (right) diagrams, both optimised by fertilisation factor

11 Optimisation: Water Demand

Through ***Optim. Water Dem.*** Menu (Figure 57), E-Water allows to execute an optimisation tool which evaluates the overall water demand resulting from different uses and its impact on an environmental indicator (WEI). Input data consist on setting default files from original scenario providing general information about available freshwater consumption arranged by population, livestock and agriculture needs (water used for irrigation).

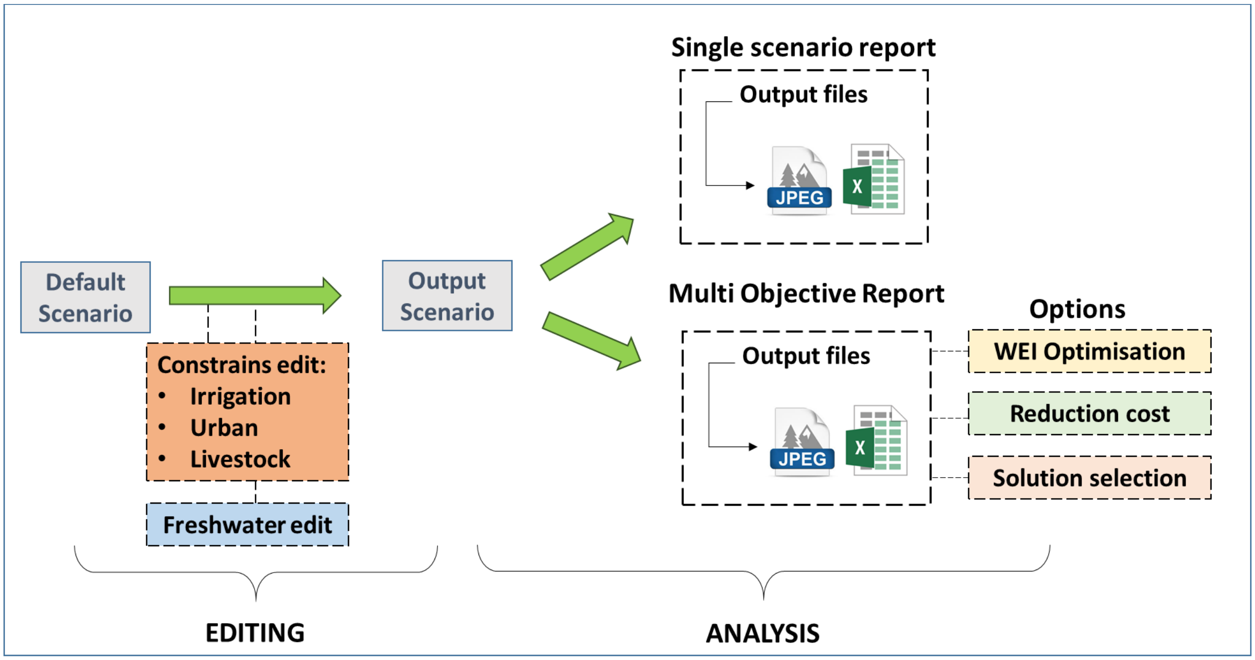


Figure 56: process diagram of Water Demand Optimisation menu

11.1 Scenario editing

This section is specifically aimed to modify the values of an existing scenario in order to create a customised input for a further optimisation process. In fact, all changes are not applied to the original scenario, but they are instead used to build a new one.

Input scenario selection: first, it takes to load an available scenario folder containing all parameters needed to run the process. All files are *.csv* tables containing all available regions and crops along with different kinds of information characterizing them all (Table 26). With ***Reset Scenario***, input scenario data will be totally reloaded from its folder, de facto undoing all further changes to its parameters and allowing user to start over the whole process.

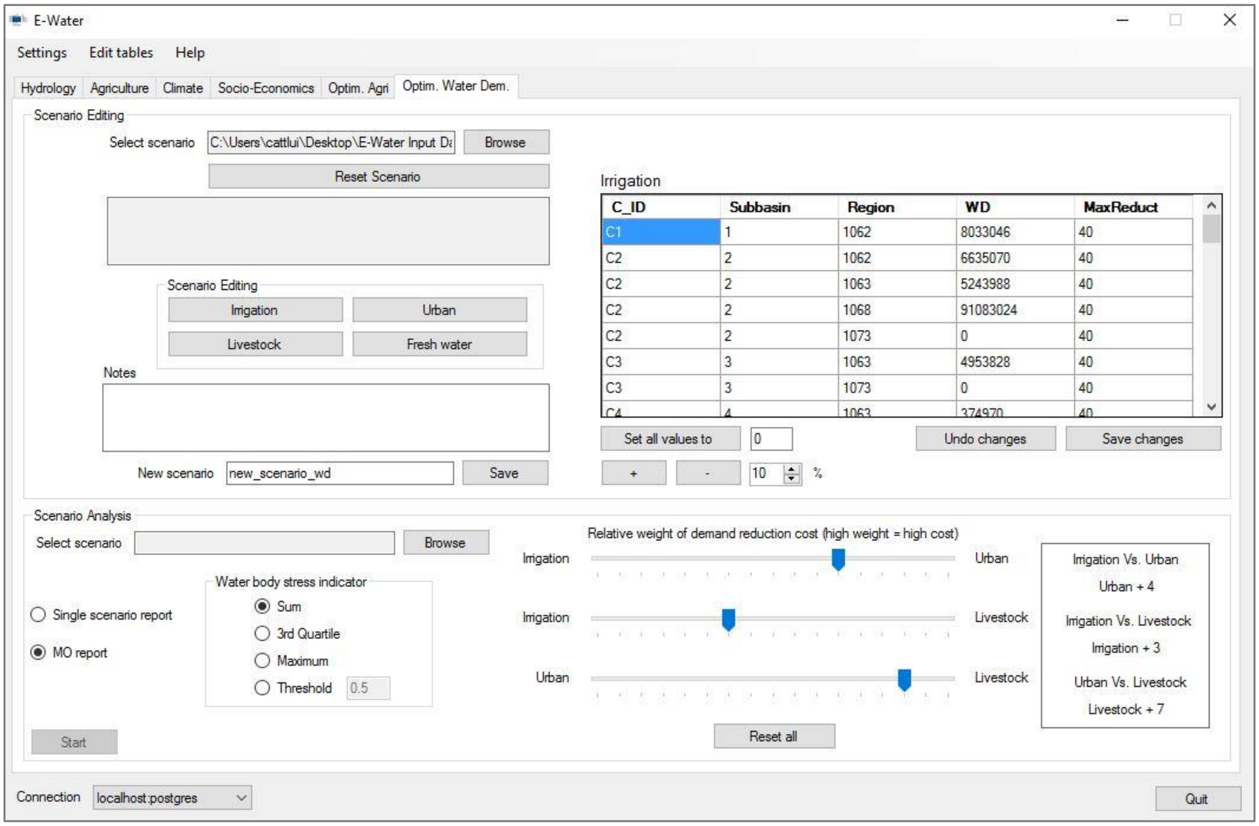


Figure 57: Water Demand Optimisation menu

File name	Related parameter
FreshWater.csv	Fresh Water
IrrigationWD.csv	Irrigation
LivestockWD.csv	Livestock
RoutingM.csv	-
UrbanWD.csv	Urban

Table 26: input files from scenario folder, along with their related editable factor

Input data edit: this section allows visualizing and editing all parameters loaded from input scenario files. Clicking on a button with the name of a parameter, its related table will be visualized in on the panel grid on the right. Through related drop down menus, it is possible to visualize all regions and crops, or to focus on one of them. User can manually edit values one by one, or editing multiple cells, setting their content to a single value (***Set all values*** to + value) or varying each one of them by a percentage (**+** or **–** with a percentage value). All changes made must be validated with ***Save changes***, otherwise they will be totally ignored; on the other hand, it is possible to reset the table with ***Undo changes***.

The main data to provide for optimisation analysis, each arranged by subbasin (***C_ID***), are:

- ***Irrigation***: water demand (***WD***) from irrigated surface and its maximum reduction percentage amount (***MaxReduct***)
- ***Livestock***: water demand (***WD***) from the cattle number and its maximum reduction percentage amount (***MaxReduct***)
- ***Urban***: water demand (***WD***) for each person and its maximum reduction percentage amount (***MaxReduct***)
- ***Fresh Water***: Total availability of fresh water (***freshW***), derived from SWAT model

Water demand values are expressed in volume (m^3). After reading input table *RoutingM.csv*, containing all information about connections between subbasins, E-Water automatically sets the routing configuration.

New scenario: at last, user can save all changes made so far by creating a new scenario containing the updated version of all input scenario files (***Save***). Optionally, he can write some comment or descriptions on the ***Notes*** panel that will be visualized new scenario is used as input for further processes.

11.2 Scenario analysis

This section is used to produce analysis reports on optimized water demand scenarios previously created and save them as files (tables and images) into chosen folder.

For each water consumption usage, it is possible to set a weight to measure the importance (or the cost) to reduce this usage compared with the others (this measure can be easily setup by user by using the interface bar indicator). All saved changes are taken into account during the creation of a new scenario folder, containing updated tables.

For every subbasin, the numeric value to optimize is identified as a ratio between annual water volume and its available water volume, also defined as Water Exploitation Index (WEI).

Scenario selection: if a new scenario has just been created in ***Scenario Editing*** section, its path folder is automatically set as input for the optimization process. Otherwise, user can select any other one.

Analysis selection: before starting analysis process (***Start***), user can choose between two different approaches: single or multi scenario report.

Single scenario report: this report mode produces image plots comparing original and modified values of the following indices, provided for every subbasin by catchment ID:

- Available water (*Fig1_fresW_st.jpeg*, left image of Figure 58)
- Water demand (*Fig2_Demands_st.jpeg*, right image of Figure 58)
- Cumulated water demand (*Fig3_Demands_routed_st.jpeg*)
- Water Exploitation Index (WEI), arranged by subbasins (*Fig4_WEI_st.jpeg,* left image of Figure 59) and in statistic distribution diagrams (*Fig5_WEI_boxplot_st.jpeg*, right image of Figure 59)

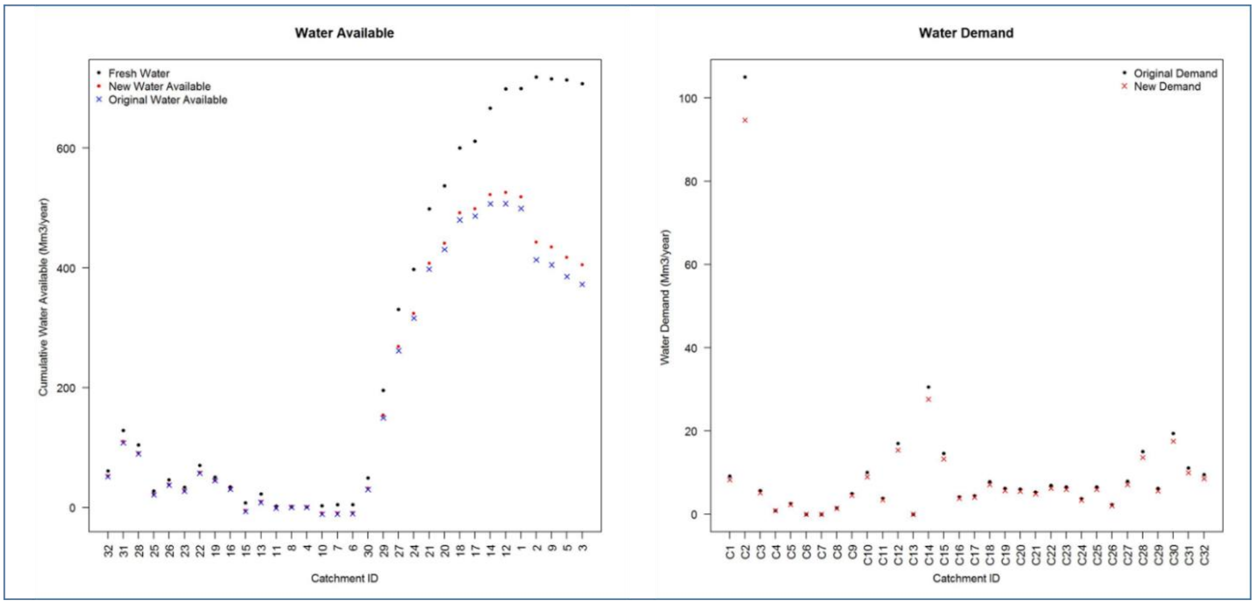


Figure 58: available water (left) and water demand (right) diagrams, arranged by catchment IDs. Original values (dots) are compared with new ones (crosses)

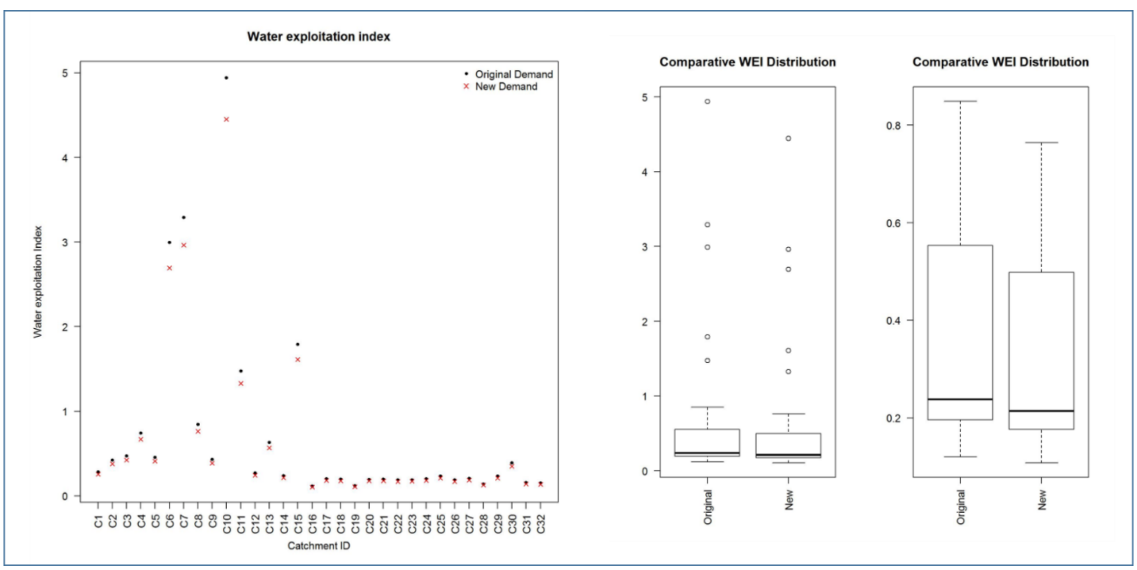


Figure 59: WEI diagram, arranged by catchment IDs (left) and in statistic distributions (right). Original values (dots) are compared with new ones (crosses)

MO report: when evaluating a multi optimization report (MO), it is possible to optimize the WEI indicator by using different aggregation (see ***Water body stress indicator***). In this case, E-Water produces a Pareto scatter plot showing all WEI values obtained using different reduction hypotheses.

Water body stress indicator: in case of ***MO report***, the optimization process is performed by evaluating the ***Maximum*** WEI possible (worst case) for every subbasin, in order to reduce water stress over resources (decrease WEI). Optionally, it is possible to optimize WEI through the ***Sum*** of all indices for every subbasin, or finding a proper ***3^rd^ Quartile***, or even setting a custom ***Threshold*** value.

Relative weight of demand reduction cost: user also has to decide which distinct water demand factor to be prioritized with respect to the others, all represented by three mutual trade-off relationships (***Urban*** Vs. ***Livestock*** Vs. ***Irrigation***). Each of them are expressed as numerical values varying from 1 to 9, representing the ratio of reduction cost of a unit of water volume (1 m^3) between the two factors (the *1* in the middle states that both of them are given the same weight).

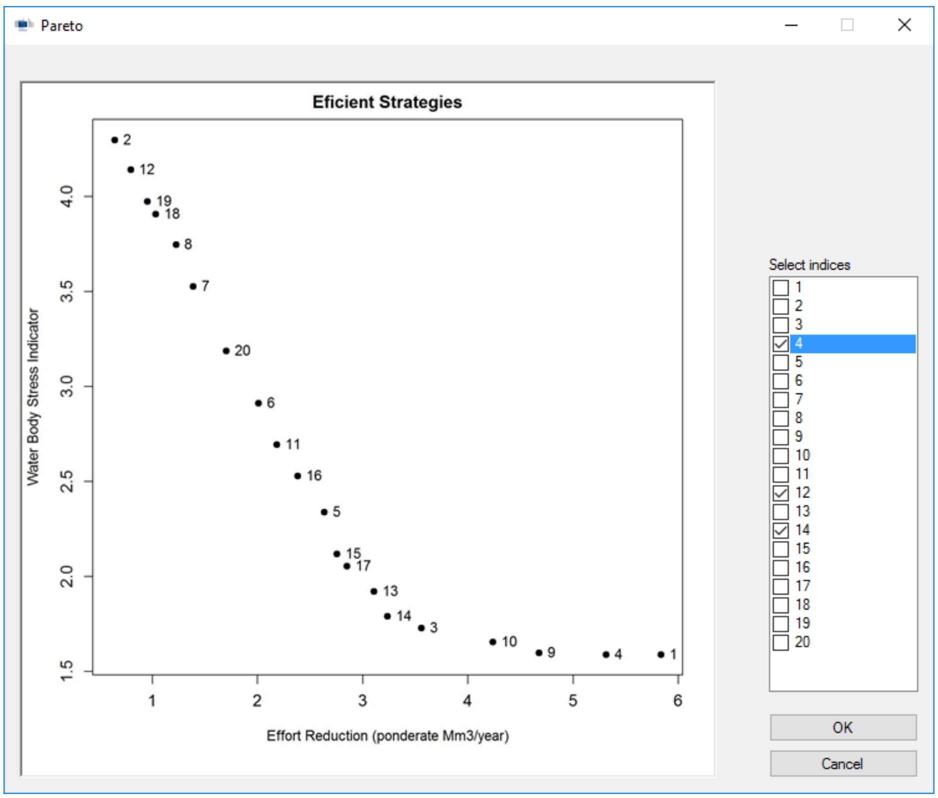


Figure 60: Pareto scatter plot. Each obtained WEI value is relative to a specific solution index

At the end of process, *E-Water* produces a Pareto scatter plot (Figure 60) showing all points corresponding to WEI values obtained using different reduction hypotheses. From this plot, it is possible to elect any of these points as suitable solutions, always considering that the best compromise may consist in the lowest WEI possible at a limited reduction effort (the 'elbow' zone of the plot). Anyway, each point in the Pareto is mathematically an optimized point but corresponding to different implementation strategies for water management and allocation between the sectors and the regions. After confirming with ***OK***, E-Water will create for every solution *n* selected by user:

- A new set of single report plots into a dedicated folder (*ParetoSelected*). All image file names are followed by solution number in order to distinguish them each other (*Fig1_fresW_st<n>.jpeg*, etc.)
- A *.csv* table containing demand reduction, expressed in percentage, for irrigation, livestock and population (*DemandReductionStr<n>.csv*)

Conclusions

Although E-Water module can be considered a completed software, there are still improvements that can be made in order to enlarge its range of capabilities. At the current moment, it relies on a limited portion of data, mostly coming from modelled datasets, so the incoming updates may include the integration processes involving the analysis of satellite data. In addition, it takes to integrate a proper input storage system, in order to easily allow users to add any datasets whenever it is needed.

It has to be reminded that the final goal of E-Water is to serve as a global process platform, so his use is not restricted to the Mékrou Basin analysis but can be considered for any future research project, regardless of the geographic domain they are located in, possibly integrating it with bigger systems with the chance to deal with a more exhaustive variety of data.

Bibliography

1. **Neitsch, S.L., Arnold, J.G., Kiniry, J.R., Williams, J.R.,** (2011) Soil & Water Assessment Tool Theoretical Documentation Version 2009. *Texas Water Resour. Inst. doi:10.1016/j.scitotenv.2015.11.063*

2. **Wang, X., J. R. Williams, P. W. Gassman, C. Baffaut, R. C. Izaurralde, J. Jeong, J. R. Kiniry,** (2012). EPIC and APEX: Model Use, Calibration, and Validation. *Trans. ASABE. doi:10.13031/2013.42253*

3. **Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., Husak, G., Rowland, J., Harrison, L., Hoell, A. and Michaelsen, J.** (2015) The climate hazards infrared precipitation with stations — a new environmental record for monitoring extremes *Scientific Data volume 2, Article number: 150066 doi:10.1038/sdata.2015.66*

4. **Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N., Vitart, F.** (2011). The ERA-Interim reanalysis: configuration and performance of the data assimilation system*, Q. J. Roy. Meteorol. Soc., 137, 553– 597, doi:10.1002/qj.828.*

5. **Russo, S., Marchese, A. F., Sillmann, J., Immé, G.** (2016). When will unusual heat waves become normal in a warming Africa? *Environ. Res. Lett., 11, 054016, doi:10.1088/1748- 9326/11/5/054016, 2016.*

6. **Russo, S., Dosio, A., Graversen, R. G., Sillmann, J., Carrao, H., Dunbar, M. B., Singleton, A., Montagna, P., Barbola, P., Vogt, J. V.** (2014). Magnitude of extreme heat waves in present climate and their projection in a warming world*, J. Geophys. Res.- Atmos., 119, 12500–12512, doi:10.1002/2014JD022098.*

7. **McKee, T. B., N. J. Doesken, and J. Kleist** (1993). The relationship of drought frequency and duration of time scales. *Eighth Conference on Applied Climatology, American Meteorological Society, Jan17-23, 1993, Anaheim CA, pp.179-186.*

8. **C. T. Agnew** (2000) Using the SPI to Identify Drought. *Drought Network News Vol. 12, No. 1, Winter 1999–Spring 2000. Published by the International Drought Information Center and the National Drought Mitigation Center, School of Natural Resources, University of Nebraska – Lincoln.*

9. **Cordano, E.**(2017). RasterList: A Raster Where Cells are Generic Objects. *R Package version 0.5.8, Joint Research Centre (JRC) of the European Commission* https://CRAN.R-project.org/package=rasterList