Integrated platform source code on Github as main

Authors : Dr. Manfred HAFNER (HEAS), Giacomo Falchetta (IIASA), Muhammad Awais (IIASA), Edward Beyers (IIASA), Vittorio Giordano (PoliTo), Manfred Hafner (HEAS), Gregory Ireland (UCT), Alison Hughes (UCT), Simon Rukera-Tabaro (UR), Francesco Semeria (PoliTo), Marta Tuninetti (PoliTo), Adriano Vinca (IIASA), Ackim Zulu (UNZA)

LEAP-RE - Contract Number: 963530

Project officer: Bernardo Luis ABELLO GARCIA

Document title	Integrated platform source code on Github as main
Author(s)	Dr. Manfred HAFNER, Giacomo Falchetta (IIASA), Muhammad Awais (IIASA), Edward Beyers (IIASA), Vittorio Giordano (PoliTo), Manfred Hafner (HEAS), Gregory Ireland (UCT), Alison Hughes (UCT), Simon Rukera-Tabaro (UR), Francesco Semeria (PoliTo), Marta Tuninetti (PoliTo), Adriano Vinca (IIASA), Ackim Zulu (UNZA)
Number of pages	12
Document type	Deliverable
Work Package	WP12
Document number	D12.3
Issued by	HEAS
Date of completion	2023-03-02 18:03:20
Dissemination level	Error (13) !

Summary

This supporting document accompanies the GitHub code and Zenodo data repositories (which are the main objects of the deliverable) containing the source code of the RE4AFAGRI nexus modelling platform (WaterCROP, M-LED, OnSSET and MESSAGE-NEST). It shortly illustrates the structure and directory tree of the repositories, and it provides a concise description of how to download the code and the data to replicate the analysis on a local computer. The next deliverable (D12.4), forthcoming in early June 2023, will then provide a detailed user guide for the utilisation of the platform and its models.

Approval

Date	By
2023-03-02 18:03:40	Dr. Manfred HAFNER (HEAS)
2023-03-02 18:08:14	Mr. Léonard LéVêQUE (LGI)

Integrated platform source code on GitHub as main deliverable of Task 12.2

Accompanying document

Deliverable D12.3

WP12 of LEAP-RE (RE4AFAGRI)

www.leap-re.eu

https://www.leap-re.eu/re4afagri/

Authors: Giacomo Falchetta (IIASA), Muhammad Awais (IIASA), Edward Beyers (IIASA), Vittorio Giordano (PoliTo), Manfred Hafner (HEAS), Gregory Ireland (UCT), Alison Hughes (UCT), Simon Rukera-Tabaro (UR), Francesco Semeria (PoliTo), Marta Tuninetti (PoliTo), Adriano Vinca (IIASA), Ackim Zulu (UNZA)

Acronyms

IIASA: International Institute for Applied Systems Analysis

LEAP-RE: Long-Term Joint Research and Innovation Partnership on Renewable Energy between the European Union and the African Union

M-LED: Multisectoral Latent Electricity Demand assessment tool

MESSAGE: Model of Energy Supply Systems And their General Environmental Impact

NEST: NExus Solutions Tool

OnSSET: Open Source Spatial Electrification Tool

POLITO: Politecnico di Torino

RE4AFAGRI: Renewable Energy for African Agriculture – Modelling Excellence and Robust Business Models

SSA: Sub-Saharan Africa

UCT: University of Cape Town

WaterCROP: Water Crop assessment model

LEAP-RE Project –*Long-Term EU-AU Research and Innovation Partnership on Renewable Energy.*

1. Background

The RE4AFAGRI project seeks to support sub-Saharan African (SSA) smallholder farmers and communities to grant community-wide access to energy services and water for crop irrigation and human use with the ultimate goal of fostering rural development. It will provide African research institutions and public and private decision makers with the tools and expertise necessary to operate a multi-scale modelling platform that can support the design and implementation of integrated solutions for the energy and water nexus in rural areas. In parallel, RE4AFAGRI will set the ground for a multi-stakeholder discussion platform about the business models and enabling environment (policy and regulation) to promote the involvement of the private sector in water-energy-agriculture integrated solutions.

This report accompanies the fourth scientific deliverable of the RE4AFAGRI project (or WP12 of LEAP-RE). This deliverable is the first official deliverable of Task 12.2 "*Integration of modelling infrastructure*" (coordinated by IIASA), albeit also being tightly linked to Task 12.3 "*Field data collection, calibration and validation*" (coordinated by UCT), which are two of the core research tasks in the development of the nexus modelling platform for RE4AFAGRI.

The RE4AFAGRI platform is a multi-model framework to analyse deficits, requirements, and optimal solutions for integrated land-water-agriculture-energy-development nexus interlinkages in developing countries.

A more comprehensive background on the design and principles behind the RE4AFAGRI platform is found in Falchetta, G., Adeleke, A., Awais, M., Byers, E., Copinschi, P., Duby, S., ... & Hafner, M. (2022). *A renewable energy-centred research agenda for planning and financing Nexus development objectives in rural sub-Saharan Africa*. Energy Strategy Reviews, 43, 100922. https://doi.org/10.1016/j.esr.2022.100922

The deliverable consists of the GitHub code and Zenodo data repositories of the RE4AFAGRI platform, and it is accompanied by this short introductory note.

LEAP-RE Project –*Long-Term EU-AU Research and Innovation Partnership on Renewable Energy.*

2. The repositories

This supporting document accompanies the GitHub code and Zenodo data repositories (which are the main objects of the deliverable) containing the source code of the RE4AFAGRI nexus modelling platform (WaterCROP, M-LED, OnSSET and MESSAGE-NEST; see Figure 1). It shortly illustrates the structure and directory tree of the repositories, and it provides a concise description of how to download the code and the data to replicate the analysis on a local computer. The next deliverable (D12.4), forthcoming in early June 2023, will then provide a detailed user guide for the utilisation of the platform and its models.
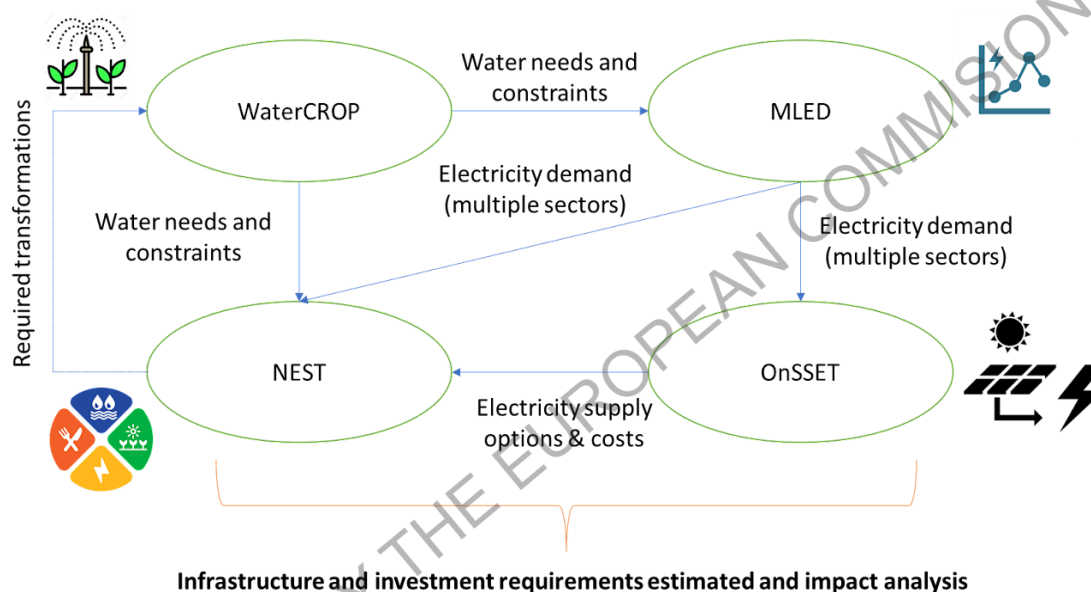

Figure 1: Schematic diagram of the RE4AFAGRI modelling platform

For a more detailed characterisation of the RE4AFAGRI models and platform development, the user is invited to refer to Deliverable 12.5 *"Joint EU-AU report on the platform testing and validation activity as a deliverable of Task 12.3 Field data collection, calibration, and validation"* and Deliverable 12.2 "*Joint EU-AU report on agreed project vision*", both retrievable on the official LEAP-RE website at https://www.leap-re.eu/.

The RE4AFAGRI platform is made publicly available on Github at https://github.com/iiasa/RE4AFAGRI_platform. As seen from Figure 2, the repository currently contains four folders, each containing the source code of:

- *watercrop*
- *m-led*
- *onnset*
- *nest*

The home page of the repository contains a general readme introductory file, which provides a general background on the platform and on the four modelling tools. It then provides different sections guiding the beginner user to (i) downloading the database to run the platform; (ii) setting up the local computer environment to run the models, in terms of required software dependencies; operating the platform, in terms of required manual steps (e.g. definition of working directory) necessary to successfully run the models; (iii) soft-linking the models, in terms of ensuring that the output of each model is successfully read and incorporated into the run of the downstream model; as well as (iv) details on how to customise the analysis with user-defined data and assumptions and visualising the results.
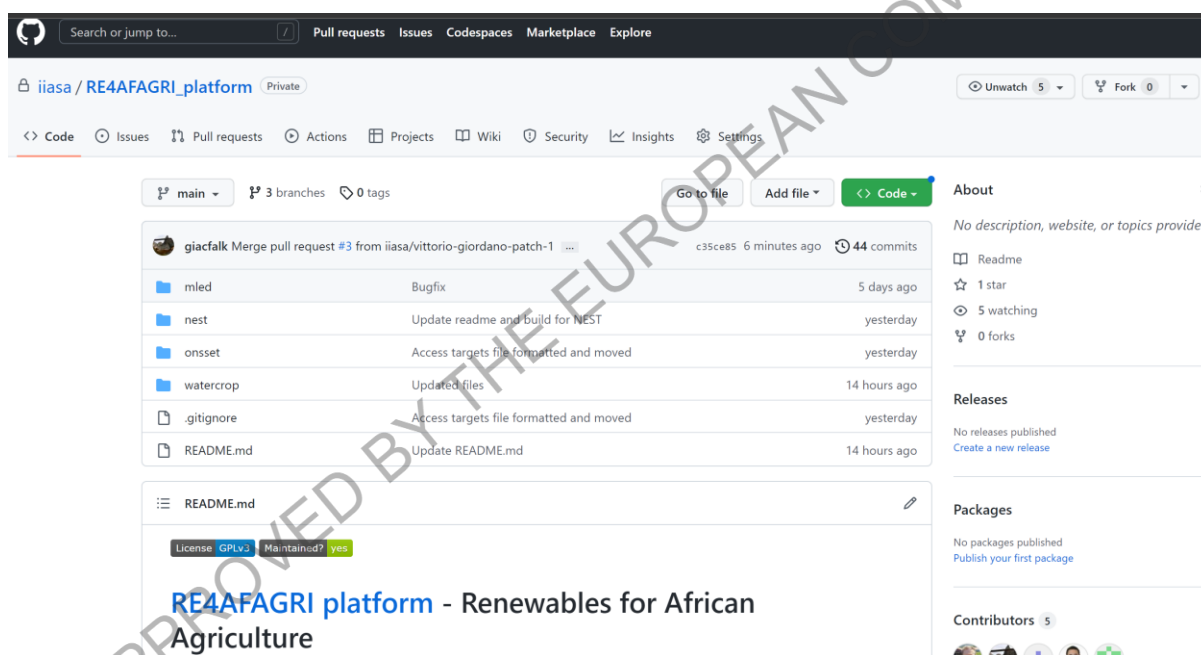

Figure 2: Screenshot of the development version of the RE4AFAGRI platform's official code repository

For general queries and issues on the models, the user is recommended to open an issue on the repository following this link: https://github.com/iiasa/RE4AFAGRI_platform/issues/new

The underlying databases to each model can be access through a unique dedicated access point, i.e. the official Zenodo repository of the RE4AFAGRI platform, accessible at: https://doi.org/10.5281/zenodo.7534846.

Figure 3: Screenshot of the development version of the Zenodo RE4AFAGRI platform's official data repository

The Zenodo repository (Figure 3) contains four zip files, each referring to one of the four RE4AFAGRI model. The repository contents can both downloaded in bulk or individual file(s) can be selected. Once downloaded, the compressed folder needs to be extracted and the corresponding local path must be replaced in each model source code as detailed in the readme file found at the home page of the RE4AFAGRI platform Github repository presented above.

The data allows replicating the Zambia country-study model (for M-LED; OnSSET; NEST) and the Africa-wide analysis (for WaterCROP), and it represents an important starting point for the scaling and replication of the analysis onto other countries of sub-Saharan Africa.

The next deliverable (D12.4), forthcoming in June 2023, will then provide a detailed user guide for the utilisation and tailoring of the platform and its models, which will be hosted on the Wiki page of the RE4AFAGRI Github repository, at https://github.com/iiasa/RE4AFAGRI_platform/wiki.

Altogether, the Github and Zenodo repository represent a crucial research output both in itself and for the forthcoming capacity building and stakeholder engagement and dissemination activities taking place in 2023-24.

The key responsible persons of each model development can be contacted under:

- For WaterCROP: contact marta.tuninetti@polito.it
- For M-LED: contact falchetta@iiasa.ac.at
- For OnSSET: contact gregoryireland@gmail.com
- For NEST: contact vinca@iiasa.ac.at

3. Instructions

3.1. Downloading the database

The database to run the platform for the pilot country of Zambia is avaiable at the official Zenodo repository of the RE4AFAGRI platform.

Once downloaded, the database(s) (a zipped folder for each of the four models) should be extracted. The exact full path to the database (e.g. *C:/Users/[yourusername]/Documents/RE4AFAGRI_database/...* should be parsed onto the different model at the following positions:

- For WaterCROP: at *lines 1, 4, 21, 55, 65, 68 - and subsequently for each crop - and 785* of the 'WaterCROP1_ETactual.mat' file at *lines 9, 14, 20, 94* of the 'WaterCROP2_Irrigation_requirements.mat' file
- For M-LED: at *line 10* of the MLED_hourly.R file, defining the db_folder parameter
- For OnSSET: include the OnSSET replication data folder unzipped in onsset\onsset_replication (more details below)
- For NEST: The database only raw data needed in the pre-processing phase. The data needed to run the model is already included in the Github repository

LEAP-RE Project –*Long-Term EU-AU Research and Innovation Partnership on Renewable Energy.*

3.2. Setting up the environment

The platform has been developed and tested in a Windows 10 environment connected to the Internet.

Each models is developed in a specific programming language and has thus specific software requirements, which are listed below.

- **First - Get the code and repository structure:**

  • Download or clone this entire repository either using, Github Desktop, git clone or downloading it as a .zip folder with all of the subfolders for each model included

- **For WaterCROP:**

  • Have 'Matlab' (version >= R2017a) installed on your local computer

- **For M-LED:**

  • Have R (version >=3.6) installed on your local computer: https://cran.r-project.org/bin/windows/base/
  • Have a recent version of RStudio installed on your local computer: https://posit.co/download/rstudio-desktop/
  • Open the MLED_hourly.r file in RStudio
  • Update your email adrress in line 12. Note that you need to enable the address to use Google Earth Engine, the procedure can be completed on https://signup.earthengine.google.com
  • Run lines 1-75. This will automatically run the backend.R file, which will take care of installing all the required package dependencies
  • During this procedure (to be carried out only the first time M-LED is run), please follow the instructions prompted in the first run to install and reply accordingly to the user prompt requests.

- **For OnSSET:**

  • Have Python (version 3+) and the conda package manager installed on your local computer:
  • If you do not have this, then download and install Anaconda for your operating system from here: https://www.anaconda.com/ (Many useful Jupyter Notebook tutorials are also available there if you are unfamiliar with Jupyter)
  • Then open "Anaconda Prompt" and navigate to this repository and into the onset folder, and run the following commands:
  • conda env create --name gep_onsset_env --file gep_onsset_env.yml (This might take a while and download 100+MB of Python packages)
  • Then run the following commands:
    – conda activate gep_onsset_env

LEAP-RE Project –*Long-Term EU-AU Research and Innovation Partnership on Renewable Energy.*

– jupyter notebook or jupyter lab (if you are familiar with JupyterLab and know how to ensure your correct environment is activated)

• This will open up the Jupyter Notebook in a browser window.

• Include the "onsset_replication" data from the [the official Zenodo repository of the RE4AFAGRI platform](#) Unzip the database and then take the onsset data into onsset\onsset_replication (the folder will exist in the code but will be empty when downloading the code from github). After completing this correctly the folder should have 3 sub-folders clusters, mled, and onsset_input_files and no longer be as .zip file. If done incorrectly the code in the next steps will likely fail.

• Go to the instructions below to run the different OnSSET parts of the model.

▪ **For NEST:**

• The requirements to run NEST are the same for running MESSAGEix, please check the [MESSAGEix documentation](#)

• Python 3.7 or newer

• GAMS with an active licence for the solver cplex

• R only needed to run pre-processing scripts

• once the setup is installed, also install the content of this folder as a `message-ix-models` package from source, following ths instructions in the `message-ix-models` [documentation](#)

3.3. Operating the platform

▪ **For WaterCROP:**

• Open 'WaterCROP1_ETactual.mat' file in MATLAB and select the scenario to run; then run the code.

• Open 'WaterCROP2_Irrigation_requirements.mat' and run the code. The code will produce the input files produces netcdf files of irrigation water requirements and yield growth potential for all African countries.

▪ **For M-LED:**

• Open the MLED_hourly.r file in RStudio

• After having run run lines 1-75 to configure the environment and the required dependencies (as discussed above), run line 80 to start running the scenarios specified in the MLED_hourly.r preamble in sequence (see below for more details on scenarios definition).

▪ **For OnSSET:**

• Make sure the replication data is downloaded from [the official Zenodo repository of the RE4AFAGRI platform](#) as described above.

LEAP-RE Project –*Long-Term EU-AU Research and Innovation Partnership on Renewable Energy.*

- In the root of the onsset folder first open MLED_extraction_to_OnSSET.ipynb and run all of the cells. This will extract the MLED demands and create OnSSET compatible input files for use in the next step. The process may take a few minutes. You should find them as .CSV files in the onsset\mled_processed_input_files folder with the names of the scenarios.
- Then, navigate into the onsset/onsset sub folder which also includes the .py Python files used by OnSSET.
- Open the OnSSET_Scenario_Running.ipynb Notebook and run all of the cells. This will take a few minutes and will run the different scenarios and calculate the least-cost electrification options for the entire country. It will output it's results into several folders as .CSV files both as full results files for every population cluster in the country as well as summary files (also used later by NEST).

▪ **For NEST:**

- After having installed the folder by source (see *Setting up the environment*) run the script `nest\message_ix_models\project\leap_re_nest\build.py`
- around line 126, at * 4) add water structure* follow the instruction and run the command in the command prompt

3.4. Soft-linking the models

RELEASE NOTE: the current release is still not including all the exchange of data for future scenarios, which will be included in the next release.

▪ **WaterCrop to M-LED:**

- WaterCrop produces netcdf files of irrigation water requirements and yield growth potential for all African countries. These files are contained (and can be updated) in the ./MLED_database/input_folder/watercrop folder and corresponding subfolders for each crop. These files are then read in the scenario_countryname.R file of M-LED.

▪ **WaterCROP to NEST:**

- NEST uses the same input as M-LED

▪ **M-LED to OnSSET:**

- The details of transferring the demand data from the MLED modelling into OnSSET compatible files are completed in the MLED_extraction_to_OnSSET.ipynb notebook in the onsset root folder

▪ **OnSSET to NEST:**

- Summary files of the OnSSET geospatial results are created when running the OnSSET scenarios in the OnSSET_Scenario_Running.ipynb inside the onsset/onsset

LEAP-RE Project –*Long-Term EU-AU Research and Innovation Partnership on Renewable Energy.*

Python code folder. These are used by NEST to split the demands appropriately for each region, rurality, year, technology type, and scenario

3.5. Customising the analysis

Consult the [Wiki documentation](#) (to be released in May 2023) for a detailed characterisation of the modules and key input datasets, parameters, and scenarios definition and updating.

3.6. Examining the results

Currently, each model has own reporting methods and formats, although a joint reporting module is currently under development.

In particular results can be examined by:

▪ **For WaterCROP:**

- WaterCROP outputs are found in 'results2' folder, and in each specific crop subfolder, which are automatically created after a model run.
- WaterCROP writes output data at 5 arcmin for the whole African continent
- The results files can be analysed using MATLAB and can be vsualized in GIS softwares

▪ **For M-LED:**

- M-LED outputs are found in the results folder which is automatically created inside the m-led home folder after a model run.
- For each scenario run, M-LED writes output data at four levels of aggregation, serving both model interlinkage purposes as part of the RE4AFAGRI modelling platform, interactive visualisation purposes in the RE4AFAGRI dashboards under development, and as static output summary files. Note that the output geopackages are reporting monthly demand for each timestep and each sector in kWh/month, while the summary csv files are reporting units in TWh/year:
  − OnSSET output geopackage, containing demand for all the original population cluster, the unit of analysis of M-LED
  − NEST output geopackage, aggregating the population clusters results at the NEST nodes level. In particular two files are written as NEST outputs, one total and one urban/rural stratified output for each NEST node.
  − GADM level 2 output geopackage, aggregating the population clusters results at the second level of administrative boundaries; useful for visualisation of aggregated results in the online dashboards and for informing policymakers
  − Summary CSVs and figures of results aggregated at the country level, disaggregated by sector, scenario, and year

LEAP-RE Project –*Long-Term EU-AU Research and Innovation Partnership on Renewable Energy.*

- **For OnSSET:**

  - The results files can be analysed using Python and Pandas in the "OnSSET_Scenario_Running.ipynb" notebook or with custom notebooks.
  - Otherwise they can be visusalised in GIS software such as QGIS: https://download.qgis.org/. They can be linked back to the cluster .gpkg shape files using a join on the "id" variable to visualise the shapes in addition to the electrification optimization information.
  - Go to the RE4AFAGRI visualisation platform (coming soon) to see the existing scenarios.

- **For NEST: work in progress**

LEAP-RE Project –*Long-Term EU-AU Research and Innovation Partnership on Renewable Energy.*