**Global Ecosse overview**

**Introduction**

Global Ecosse consists of a set of programme built on Python modules some of which have been derived from Mark Richard’s Elum[[1]](#footnote-1) scripts which have been expanded, restructured and elaborated upon and others which have been written from scratch. Each programme uses a GUI (graphical user interface) to allow users to set values for each stage of the simulation runs.

In each programme folder there is a setup file, e.g. *global\_ecosse\_setup.txt*, which sets parameters such as pathnames and needs to be edited appropriately for each new environment.

**Data Sources**

* Soil data is from the Harmonized World Soil Database ([HWSD](http://webarchive.iiasa.ac.at/Research/LUC/External-World-soil-database/HTML/)) Granularity of 30 arc seconds.
* Land cover data is from the European Environment Agency ([EEA)](https://www.eea.europa.eu/) is an agency of the EU and the [Copernicus](http://land.copernicus.eu/pan-european/corine-land-cover/clc-2000) land monitoring service.
* Historic and future weather data is from the [UEA](https://www.uea.ac.uk/)’s Climatic Research Unit([CRU](http://www.cru.uea.ac.uk/data)) high-resolution gridded datasets which comprise NetCDF format files. Granularity of 0.5 degrees.
* GADM database of Global Administrative Areas ([GADM](http://www.gadm.org/)) maintained by California based institutes.
* It is not yet possible to specify a file comprising yield data per grid cell for the AOI (area of interest).

**The Preparation Phase**

This phase consists of stages 1 and 2 which are designed to limit the volume of downstream processing and to simplify the complexity of the simulations phase. The programmes enable a user to define an AOI (area of interest) e.g. a country such as Poland or a province such as Alberta in Canada, and to create a CSV of grid points with lookup indices for the HWSD for the AOI. “Cookie cutter” methodology is optionally employed to cut out adjacent land not belonging to the AOI. The user may also constrain the AOI to a bounding box defined by two latitude/longitude pairs.

**Stage 1: *ReformShapes***

Settings for this programme are defined in the setup file, ***reform\_setup.txt****,* in the programme folder.

Prepare national or province boundaries for use ***CookieCut*** programme in stage 2.

Take a GADM[[2]](#footnote-2) country or province level shapefile and make it more manageable by reducing the numbers of boundary points by “thinning” i.e. tidying the polygons.

The Python wrapper drives a C programme.

A new shape file is created in the *C:\temp* folder which must already exist.

**Stage 2a:** ***CookieCut***

Settings for this programme are defined in the setup file, ***cookie\_cut\_setup.txt****,* in the programme folder.

Create an AOI consisting of a CSV[[3]](#footnote-3) file. Each line consists of a coordinate set and a HWSD “**mu\_global**” mapping unit identifier extracted from the HWSD **.bil**[[4]](#footnote-4) file.

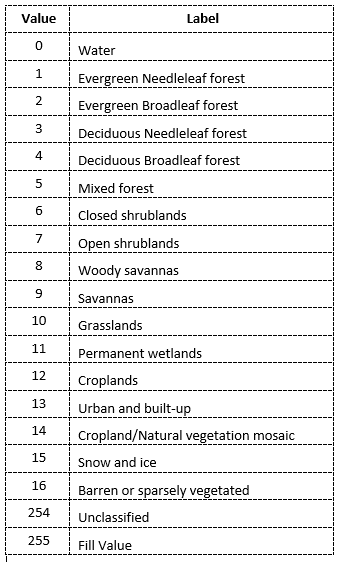
The file will have the name: **<province>****\_hwsd.csv** and will be created in the folder specified by the "images\_dir" entry in the setup file.

The user selects the shape file generated by ***ReformShapes*** and specifies a bounding box consisting of lower left longitude and latitude and upper right longitude and latitude.

Use the ***Countries***utility to determine bounding boxes for countries and provinces within countries.

The output file will consist of cells which lie within the (polygon) boundaries defined by the shapefile and will be clipped to the specified bounding box.

**Table 1 – MODIS code values**

**Stage 2b: *ModisHwsd* Land Use**

This programme uses the Global Land Cover Facility[[5]](#footnote-5) to enable the user to subsample HWSD CSV files created by the ***CookieCut*** programme with the name form of: **<province>\_hwsd.csv**

Settings for this programme are defined in the setup file, ***read\_lu\_setup.txt***, in the programme folder.

A configuration file is also used, ***hwsd\_modis\_config.txt***, which is recorded in a location specified in the setup file.

The user writes and selects a “JSON”[[6]](#footnote-6) file which specifies one or more provinces and one or more land cover types as defined in **table 1**.

Each line consists of a coordinate set, a HWSD “**mu\_global**” identifier as used by the HWSD and the land cover index.

The output file will consist of grid cells with the specified province and land cover types. For example, the Holy Cross province in Poland covers an area of 11,672 km2 which corresponds to 21,428 HWSD cells of which 12,052 cells for land cover types 10 and 12 (grasslands and croplands).

**Simulations Generation phase**

This phase generates the ECOSSE input file sets required for the ECOSSE programme.

The first step is to write ECOSSE formatted past and future meteorological files to a directory structure. Future climate is based on one of four future climate scenarios, namely **A1B\_MG1**, **A2\_\_MG1**, **B1\_\_MG1** and **B2\_\_**MG1 described in appendix A below

The second step reads the previously created HSWD CSV file and climate files to generate limited data site simulation files suitable for ECOSSE.

A ‘manifest’ file is written for each simulation set which reduces the amount of numbers of file sets required for a given region by identifying clusters of consecutive HWSD grid cells with the same latitude, soil and climate data.

**Stage 3: GlblEcosse**

Settings for this programme are defined in the setup file, ***global\_ecosse\_setup.txt***, in the programme folder.

Read the **\_hwsd.csv** AOI file from stage 2 and write ECOSSE input file sets for each entry but skipping adjacent entries on the same latitude with the same HWSD “**mu\_global**” identifier. Adjacent entries are recorded in a unique “manifest file”.

Historic and future climate files are generated for 200 years from 1901 until 2100 and are written to the corresponding location under the simulations folder, "**sims\_dir**" in global\_ecosse\_setup.txt, e.g.: "E:\\SpatialOutputs\\EcosseSims\\A1B\_MG1"

Generation is performed by dividing the AOI into 0.15 degrees Latitude bands corresponding to 36 rows of HWSD data.

**Spatial ECOSSE phase**

This phase uses the multi-site submission script (**spec.py**) which is tested, reliable and efficiently makes use of multi-core CPUs. The script runs the ECOSSE program which writes a **SUMMARY.OUT** file for each soil for each site. The HWSD permits a maximum of 9 soils for each grid cell.

**Stage 4:** **DevSpec**

The ECOSSE programme is executed for each of the input file sets generated in stage 3.

**Post-processing phase**

The user is given the option to aggregate the data from the **SUMMARY.OUT** files to CSV flat files for each desired metric.

Alternatively the user may request that outputs are written to geo-referenced  [NetCDF](https://en.wikipedia.org/wiki/NetCDF) format files. These files may be viewed using the Panoply[[7]](#footnote-7) data viewer.

**Stage 5:** **DevSpec** (post processing options)

Two options are available both of which comprise the serial reading of the SUMMARY.OUT files written in stage 4

Create a tab-delimited CSV file[[8]](#footnote-8) for each requested variable.

Create a geo-referenced NetCDF file which includes all requested variables.

**Resources consumed:**

Typically, for Ireland with a land area of 70,273km2 a total of 35,751 folders are generated with 5 files per folder plus 10,923 manifest files taking up 610Mb of disk space. The total time taken for all stages is approx. 3.5 hours.

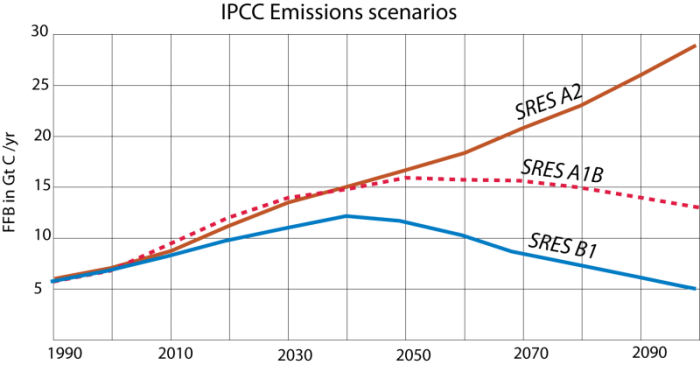
**Other utilities: GlblEcsseUtils**

DelSimsGUI.py – enables user to clean up sims directories. Faster than using Windows File Explorer as it uses the Python **shutil** module which offers a number of high-level operations on files and collections of files.

NB Deleted folders and files are not transferred to the Recycle Bin.

**Appendix A**

These scenarios are from SRES ( [Special Report on Emissions Scenarios](https://en.wikipedia.org/wiki/Special_Report_on_Emissions_Scenarios))



|  |  |
| --- | --- |
| A2 | A2 is commonly known as business-as-usual — in other words, it leads to a continuation of increased annual carbon emissions that follows the recent history. In effect, this scenario represents a somewhat divided world, one in which we just can't reach agreements on what to do about limiting emissions of CO2, so each country does what seems to be in its own best interests. This world is characterized by independently operating, self-reliant nations. This world also includes a continuously increasing population — it does not level off during this time period. Above all, in this world, decisions are based primarily on perceived economic interests, and the assumption is made that these interests do not include the development of alternative energy sources. |
| A1B | A1B is more optimistic. This scenario envisions an integrated world characterized by rapid economic growth, a population that reaches 9 billion by 2050 and then declines gradually, and the rapid development of alternative energy sources that facilitate increased economic growth while limiting and eventually reducing carbon emissions. This scenario also assumes that there will be rapid development and sharing of technologies that help us reduce our energy consumption. One of the keys to this scenario is that countries are integrated — they act together and find ways to improve the conditions for everyone on Earth. At this point in time, A1B is an optimistic but realistic scenario |
| B1 | B1 represents an even more integrated, more ecologically friendly world, but one in which there is still steady and strong economic growth. As in scenario SRES A1B, the population in this scenario peaks at 9 billion in 2050 and then declines. One way to think of this scenario is that it represents a rapid, strong, and global commitment to the reduction of carbon emissions — it represents the best we could possibly do, and yet it does not rely on miracle technologies. The only real miracle it requires is that we all quickly figure out how to think and act globally and not focus solely on our own national interests. |
| B2 | Most projections ignore scenario B2, as the combination of regional and environmental strategies is highly unlikely. |

Problems with DevSpec

<https://stackoverflow.com/questions/19981140/clean-python-multiprocess-termination-dependant-on-an-exit-flag>

maxwell2.abdn.ac.uk

1. The Elum project was a study into the sustainability impacts of land-use change to bioenergy crops in the UK. [↑](#footnote-ref-1)
2. Global Administrative Areas (GADM) goal is to map the administrative areas of all countries, at all levels. – see: <http://www.gadm.org/> [↑](#footnote-ref-2)
3. A comma-separated values (CSV) file stores tabular data (numbers and text) in plain text. Each line of the file is a data record. Each record consists of one or more fields, separated by commas – see: <https://en.wikipedia.org/wiki/Comma-separated_values> [↑](#footnote-ref-3)
4. A “**BIL**” image **file** (.**bil**), which means “band interleaved by line,” is an uncompressed **file** containing the actual pixel values of an image – see: <http://desktop.arcgis.com/en/arcmap/10.3/manage-data/raster-and-images/bil-format-example.htm> [↑](#footnote-ref-4)
5. See: <http://glcf.umd.edu/data/lc> [↑](#footnote-ref-5)
6. 'A person-readable collection of data comprising states (or provinces) and land use types. JSON is an acronym for JavaScript Object Notation, and is a way to store information in an organized easy-to-access manner. ' [↑](#footnote-ref-6)
7. Panoply plots geo-referenced and other arrays from [netCDF](http://www.unidata.ucar.edu/packages/netcdf/) and other datasets.

   See: <https://www.giss.nasa.gov/tools/panoply/> [↑](#footnote-ref-7)
8. In principle Excel format may be used, however, Excel cannot exceed the limit of 1,048,576 rows and 16,384 columns. By default, Excel places three worksheets in a workbook file. Each worksheet can contain 1,048,576 rows and 16,384 columns of data, and workbooks can contain more than three worksheets if your computer has enough memory to support the additional data. [↑](#footnote-ref-8)