**Software paper for submission to the Journal of Open Research Software**

To complete this template, please replace the blue text with your own. The paper has three main sections: (1) Overview; (2) Availability; (3) Reuse potential.

Please submit the completed paper to: editor.jors@ubiquitypress.com

**(1) Overview**

Title

Moirai version 3: a data processing system to generate recent historical land inputs for global modelling applications at various scales

Paper Authors

1. Di Vittorio, Alan V.

2. Vernon, Chris

3. Shu, Shijie

Paper Author Roles and Affiliations

1. Lawrence Berkeley National Laboratory, Berkeley, CA, US

2. Joint Global Change Research Institute, Pacific Northwest National Laboratory, College Park, MD, US*.*

3. University of Illinois at Urbana-Champaign, Urbana-Champaign, IL, US

Abstract

The Moirai land data system is designed to produce recent historical land inputs for an integrated assessment model. The primary function of Moirai is to combine spatially explicit input data (e.g., raster images) with tabular input data (e.g., crop price table) to generate spatially-referenced tabular data of crop production, crop harvested area, land value, irrigated and rainfed crop area, water footprint, soil and vegetation carbon density of unmanaged land, and historical land use/cover. These data are aggregated to user-defined geographic boundaries within 231 countries, and the default boundaries are defined globally by 235 watersheds. The production, harvested area, and land value outputs reconstruct those available from the Global Trade Analysis Project, while the other outputs provide additional information for various applications. Furthermore, Moirai is a modular system that can be updated and customized through replacement and addition of source data.

Keywords

area, crop, data, GCAM, GTAP, HYDE, IAM, land, land cover, land use, moirai, model, pasture, production, system, urban, water

Table 1. Raster inputs (these have been accessed in June 2018 unless noted otherwise)

|  |  |  |
| --- | --- | --- |
| Data | Details | Source |
| Crop yield and harvested area | 5 arcmin,  175 crops-same as GTAP,  circa 2000,  area provided as fraction of land area in grid cell | Monfreda et al., 2008;  http://www.earthstat.org/data-download/ |
| Cropland physical extent | 5 arcmin,  circa 2000,  provided as fraction of land area in grid cell | Ramankutty et al., 2008;  http://www.earthstat.org/data-download/ |
| Irrigated and rainfed crop harvested area | 5 arcmin, hectares,  26 crop classes,  circa 2000 | Portmann et al., 2010;  https://www.uni-frankfurt.de/45218031/data\_download/ |
| Crop water footprint data | 5 arcmin, mm/yr,  18 crop types,  3 water types,  circa 2000 | Mekonnen and Hoekstra, 2011;  http://waterfootprint.org/en/resources/ water-footprint-statistics/ |
| Fraction of land area in grid cell for crop and water footprint data above | 5 arcmin,  spherical earth with  WGS84 mean radius | D. Plouff and N. Ramankutty provided these data corresponding to the above cropland data (in late 2013). Note that these are the same data used to provide the area values in the current crop yield and harvested area data above |
| Potential vegetation | 5 arcmin,  thematic,  15 vegetation types, circa 2000 if no historical land use had occurred | Ramankutty and Foley, 1999;  http://www.earthstat.org/data-download/ |
| Land use area | 5 arcmin, km2,  1700-2016 (decadal up to 2000),  HYDE 3.2.000 baseline,  12 land use types | Klein Goldewijk et al., 2017;  ftp://ftp.pbl.nl/hyde/hyde3.2/2017\_beta\_release/ |
| Land area in grid cell | 5 arcmin, km2,  circa 2000,  spherical earth with WGS84 mean radius,  with Greenland and several arctic islands added based on fraction of land area in grid cell for crop area and potential vegetation;  this is the working grid | Klein Goldewijk et al., 2017;  ftp://ftp.pbl.nl/hyde/hyde3.2/2017\_beta\_release/ |
| Total grid cell area | 5 arcmin, km2,  spherical earth with WGS84 mean radius,  with Greenland and several arctic islands added based on fraction of land area in grid cell for crop area and potential vegetation;  this is the working grid | Klein Goldewijk et al., 2017;  ftp://ftp.pbl.nl/hyde/hyde3.2/2017\_beta\_release/ |
| 234 Country boundaries | 5 arcmin,  from VMAP0 vector data (the source of FAO country boundaries),  added East Timor based on a map,  and merged some countries to reflect FAO data | VMAP0: http://gis.ess.washington.edu/data/raster/GlobalData/ (last accessed in 2013, now restricted to UW, but these data are currently available in four parts at http://gis-lab.info/qa/vmap0-eng.html); |
| Original AEZ boundaries | 5 arcmin,  1961-1990 data,  160 country boundaries,  GTAP Land Use Database, Release 2.1 | Lee et al., 2005; Lee et al., 2009; Monfreda et al., 2009; https://www.gtap.agecon.purdue.edu/resources/res\_display.asp?RecordID=1900 |
| Output Geographic Land Unit (GLU) boundaries | 5 arcmin,  thematic,  235 water basins | Developed for the water module of the Global Change Assessment Model, aggregated from a 1/8-degree global watershed data set |
| Land cover area data | half-degree,  1800-2016 (decadal up to 2000),  23 land cover types,  fraction of grid cell and grid cell area | Meiyappan and Jain, 2012;  Produced specifically for Moirai using HYDE 3.2.000 data;  alternate public version available here: https://www.atmos.illinois.edu/~meiyapp2/datasets.htm |
| Protected land area | 5 arcmin,  thematic,  protected or not-protected | Derived from a previous version of the World Database on Protected Areas;  (current version available at: https://www.iucn.org/theme/protected-areas/our-work/world-database-protected-areas) |

Table 2. Text inputs as comma separated value files (these have been accessed in June 2018 unless noted otherwise)

|  |  |  |
| --- | --- | --- |
| Data | Details | Sourcea |
| 87 economic regions | Tabular,  GTAP Land Use Database, Release 2.1 | Lee et al., 2005; Lee et al., 2009; Monfreda et al., 2009; https://www.gtap.agecon.purdue.edu/resources/res\_display.asp?RecordID=1900 |
| Land rent for 13 use sectors | Tabular,  87 regions by 18 AEZs,  GTAP Land Use Database, Release 2.1 | Lee et al., 2005; Lee et al., 2009; Monfreda et al., 2009; https://www.gtap.agecon.purdue.edu/resources/res\_display.asp?RecordID=1900 |
| FAO 235 countries | Tabular,  with some added countries to match the VMAP0 data,  there are two input files containing these data: one maps the countries to the economic regions and the other maps the countries to the raster country data | http://faostat.fao.org/; accessed Aug 2013 |
| Geographic Land Unit (GLU) list | Thematic codes and names for the GLU raster data | Developed for the water module of the Global Change Assessment Model, aggregated from a 1/8-degree global watershed data set |
| GCAM region list | Names and integer codes for GCAM regions,  used in some diagnostics | GCAM-source? |
| Country to GCAM region mapping | Cross-reference table mapping FAO countries to GCAM regions | FAO country and GCAM region data |
| HYDE3.2.000 list | Thematic codes and names for HYDE3.2.000 raster data | Klein Goldewijk et al., 2017;  ftp://ftp.pbl.nl/hyde/hyde3.2/2017\_beta\_release/ |
| Land cover to land use mapping | Cross-reference table mapping land cover to potential vegetation and land use | Land cover, land use, and potential vegetation data |
| GTAP product use list |  | Lee et al., 2005; Lee et al., 2009; Monfreda et al., 2009; https://www.gtap.agecon.purdue.edu/resources/res\_display.asp?RecordID=1900 |
| 175 crop to FAO crop and GTAP use mapping | Cross-reference table mapping 175 crops to FAO crops to GTAP use | 175 Crop, FAO crop, and GTAP use data |
| FAO production data | Tabular,  up to 169 crops,  1997-2007 | http://faostat.fao.org/; accessed Aug 2013 |
| FAO yield data | Tabular,  up to 160 crops,  1997-2007,  for diagnostics only | http://faostat.fao.org/; accessed Aug 2013 |
| FAO harvested area data | Tabular,  up to 161 crops,  1997-2007,  for year recalibration only | http://faostat.fao.org/; accessed Aug 2013 |
| FAO Crop producer prices data | Tabular,  up to 205 crops,  1997-2007 | http://faostat.fao.org/; accessed Aug 2013 |
| USD-year conversion list | Factors to convert input FAO 2005 USD to output 2001 USD | Derived from consumer price index centered at 1982-1984 |
| Potential vegetation list | Thematic codes and names for potential vegetation raster data | Ramankutty and Foley, 1999;  http://www.earthstat.org/data-download/ |
| Vegetation carbon density for potential vegetation | Average vegetation carbon densities for the potential vegetation types | Literature review |
| Soil carbon density for potential vegetation | Average soil carbon densities for the potential vegetation types | Literature review |