Spatial simulation infrastructure for Indonesia v.0.1

Data descriptor

RESTORE+

Technical Report

Supported by:

Federal Ministry for the Environment, Nature Conservation and Nuclear Safety

based on a decision of the German Bundestag

IIASA
World Agroforestry Centre
INPE
ipea Instituto de Pesquisa Econômica Aplicada
UN environment
WCMC
WWF
MCC
EDF
WRI INDONESIA
LSE
Grantham Research Institute on Climate Change and the Environment

Spatial simulation infrastructure for Indonesia v.0.1

Data descriptor

International Institute for Applied Systems Analysis (IIASA)
Laxenburg, August 2022

The RESTORE+ project is implemented by the International Institute for Applied Systems Analysis (IIASA), World Agroforestry Centre (ICRAF), Brazil National Space Research Agency (INPE), Brazil Institute for Applied Economic Research (IPEA), UN Environment-World Conservation Monitoring Centre (UNEP-WCMC), World Resources Institute (WRI) Indonesia, World Wildlife Fund (WWF) Indonesia, Mercator Research Institute on Global Commons and Climate Change (MCC), Environment Defense Fund (EDF) and London School of Economics (LSE) Grantham Research Institute on Climate Change and the Environment.

The project is part of the International Climate Initiative (IKI). The German Federal Ministry for the Environment, Nature Conservation and Nuclear Safety (BMU) supports this initiative on the basis of a decision adopted by the German Bundestag.

Authors:
Rastislav Skaský, International Institute for Applied Systems Analysis
Juraj Balkovič, International Institute for Applied Systems Analysis
Christian Folberth, International Institute for Applied Systems Analysis
Florian Kraxner, International Institute for Applied Systems Analysis
Ping Yowargana, International Institute for Applied Systems Analysis

Publication Editing and Layouting:
Charlotte Kottusch, Enrico Confienza

Required Citation:
Skaský, R., Balkovič, J., Folberth, C., Kraxner, F., Yowargana, P., (2022). *Spatial simulation infrastructure for Indonesia v.0.1*. RESTORE+ Technical Report, Vienna.

Executive Summary

- This document describes version 0.1 of the Spatial simulation infrastructure for the bio-physical modelling in Indonesia and the procedures of its creation.

- Primary purpose of the dataset was to provide a consistent interface for effective integrating of the data coming from variety of resources, as well as to provide the RESTORE+ project bio-physical and economic optimization modelling suite (EPIC-IIASA, G4M, WaNuLCAS, and GLOBIOM) with necessary inputs on terrain, soil, and sub-national administrative units.

Table of Contents

1. Introduction

An important part of the RESTORE+ project activities in Indonesia is process-based, bio-physical modelling of agricultural, forestry, and agro-forestry systems. Bio-physical models (EPIC-IIASA, G4M, WaNuLCAS), as well as some agricultural sector economic and land use optimization models (GLOBIOM) require spatially allocated data on all their mandatory inputs.

Usually, along with the daily weather data, the data on terrain (altitude, slope inclination, slope length) and soil (soil properties, such as profile distribution of the soil organic carbon, pH, or cation exchange capacity) provide direct mandatory inputs or boundary conditions. Other data elements, such as administrative unit's borders, provide a link to other more comprehensive, and multi-layered data linked to them, such as crop production or land use and land use change statistics. All these data are an important input for setting up the bio-physical simulation scenarios, performing regional calibrations, and/or validating the simulation outputs.

To support the communication among the models, either at the level of shared data inputs (for bio-physical models) or input/output exchange (e.g., with economic optimization models such as GLOBIOM), it is useful to setup a system of Spatial simulation units (SimU). The SimU can be setup as regularly (grid) or irregularly shaped spatial domains, the latter based on the similarities in the geographical conditions (e.g., topography and soil properties). SimU provide the RESTORE+ project with consistent interface for integrating the input data for models coming from variety of resources and thus, differing in their contents, spatial resolutions, and formats. The minimum data content necessary for running the bio-physical models, including the data on terrain, soil, and administrative unit's borders was attributed to the SimU for making the spatial data infrastructure to be ready for application/usage with the RESTORE+ bio-physical and economic optimisation modelling suite.

2. Input data and data processing

2.1 Spatial reference – Spatial simulation units (SimU)

Geographical grid of regular grid cells with spatial resolution of 5 x 5 arcmin (about 9.25 km at the equator) covering all administrative regions of Indonesia has been chosen to serve as basic spatial reference for all geographical data (inputs/outputs of the models, auxiliary and supporting data).

2.1.1 Geometry of the SimU

The rectangular grid of regular square elements was constructed in a GIS as a mesh (or fishnet) with the respective size of 5 arcmin. The geometry of the individual spatial elements was either polygon (square) or point (centroid of the square). Together they constitute regional coverage of polygon features (further referred to as grid) or point features (further referred to as lattice) fully covering whole national territory of Indonesia (including seas and ocean).

2.1.2 Spatial definition of the SimU

Each individual grid cell of the 5 x 5 arcmin spatial resolution grid represents a unique Spatial simulation unit (SimU) with a purpose to provide the models with spatially allocated and harmonized inputs, as well as to serve as a geographical reference for displaying outputs of the models.

A land surface mask was constructed based on the coastlines as available from the national administrative units dataset (BPS - Statistics Indonesia, https://www.bps.go.id/). All the grid cells intersecting the land surface mask were selected and taken as SimU for the final dataset (IND SimU v0.1) counting 26,883 spatial elements (grid cells, grid centroids). All other spatial elements, mostly representing sea and ocean surface or areas outside national territory of Indonesia, were set as not relevant and excluded from the SimU coverage (Figure 2.1).

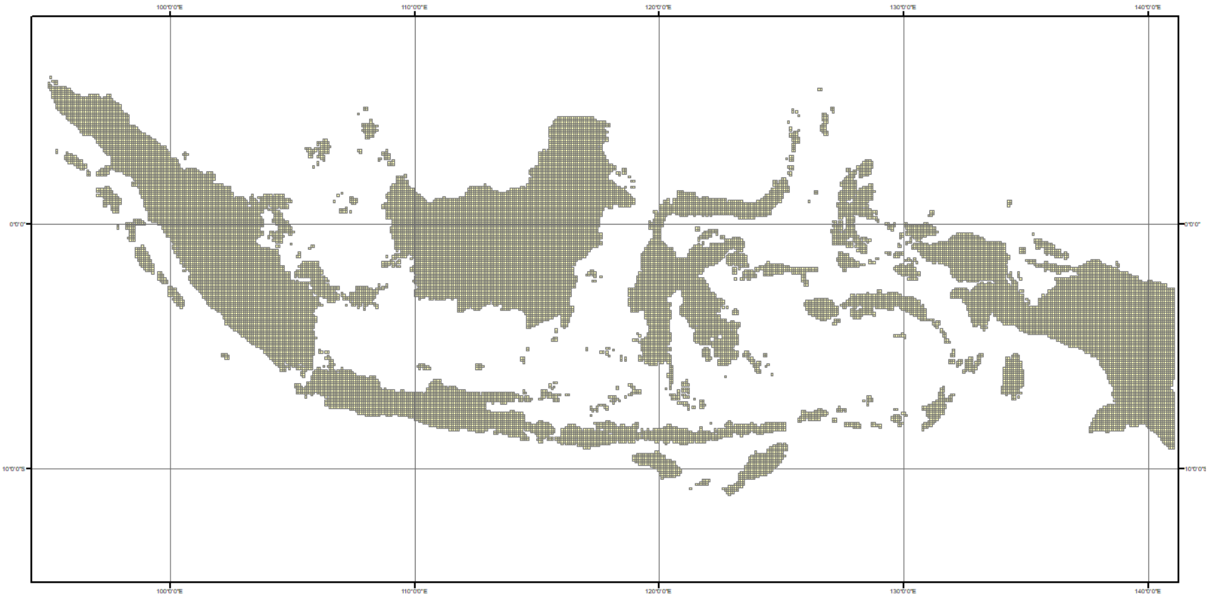

Figure 2.1 Spatial reference for input/output data, grid of 5x5 arcmin spatial resolution grid cells constituting spatial simulation units (SimU) for the models.

2.1.3 Indexing of the SimU

A unique integer number was assigned automatically by GIS software to each SimU and then sampled to the SimU centroid point locations, securing that spatially corresponding polygons and points (SimU and SimU centroids) had the same identification number. The column and row address, with columns and rows counted from upper left corner of the global coverage dataset, was also assigned to each SimU (grid, lattice). Colum and row identification of the SimU or SimU centroids keeps the direct link to the legacy data (the GEO-BENE project global dataset):

Min col/Min row = 1/1 (upper-left corner),

Max col/Max row = 4320/2160 (lower-right corner) of the global coverage.

2.1.4 Real area of the SimU

Geographical grid of regular grid cells with spatial resolution of 5 x 5 arcmin for Indonesia (the SimU coverage) was constructed using global geographic coordinate system WGS 1984 (further referred to as WGS).

Directly referring to a spherical shape of the Earth, the WGS coordinate system measures distances by angular units (decimal degrees, or dd). It does account for spherical shape of the Earth projected to the plane and keeps the angle-based distances between points also in the two-dimensional maps. As a result, all spatial elements (polygons) to the north or to the south of the equator always appear on the map larger than they are in reality (or two points appear to be located further away from each other).

Albers equal-area conic projected coordinate system for South Asia (https://spatialreference.org/ref/esri/asia-south-albers-equal-area-conic/) was used to get the SimU coverage projected into geographic coordinate system which enables the real area calculations, and real areas in hectares (ha) were calculated of all SimU (Figure 2.2).

Figure 2.2 Real areas (ha) calculated for each SimU after re-projecting the WGS geographical coordinate system into the equal-area geographic regional projection (Albers equal-area conic projected coordinate system for South Asia).

2.2 Input data for models

An important reason for the design and construction of the SimU spatial reference was to create a data harmonization frame for the inputs for all bio-physical models assumed for the RESTORE+ project (EPIC-IIASA, G4M, WaNuLCAS) which would provide consistent interface for integrating the data coming from variety of resources and thus, differing in their contents, spatial resolutions, and formats.

Several mandatory data elements are required by most of the bio-physical models, as well as by agricultural sector economic and land use optimization models (such as GLOBIOM). Usually the data on terrain (altitude, slope inclination, slope length) and soil (soil properties, such as profile distribution of the soil organic carbon, pH, or cation exchange capacity) provide direct basic inputs, boundary conditions, or drivers for the simulations. Other data elements, such as administrative unit's borders, provide a link to other more comprehensive, and multi-layered data linked to them, such as crop production or land use and land use change statistics being important inputs for setting up the simulation scenarios, doing regional calibrations, and/or validating the simulation outputs.

The minimum data content necessary for running the bio-physical models, including the data on terrain, soil, land cover, and administrative unit's borders was attributed to the SimU, after it was spatially harmonized from the original datasets as described below.

2.2.1 Terrain data

Elevation

The 3-arsec Shuttle Radar Topography Mission (SRTM, void filled by CGIAR .4.1, Jarvis et al. 2008; Reuter, Nelson, and Jarvis 2007) was used as a source of altitude data for Indonesia. Zonal means of underlying SRTM pixels were taken as representative altitude values for the SimU (Figure 2.3).

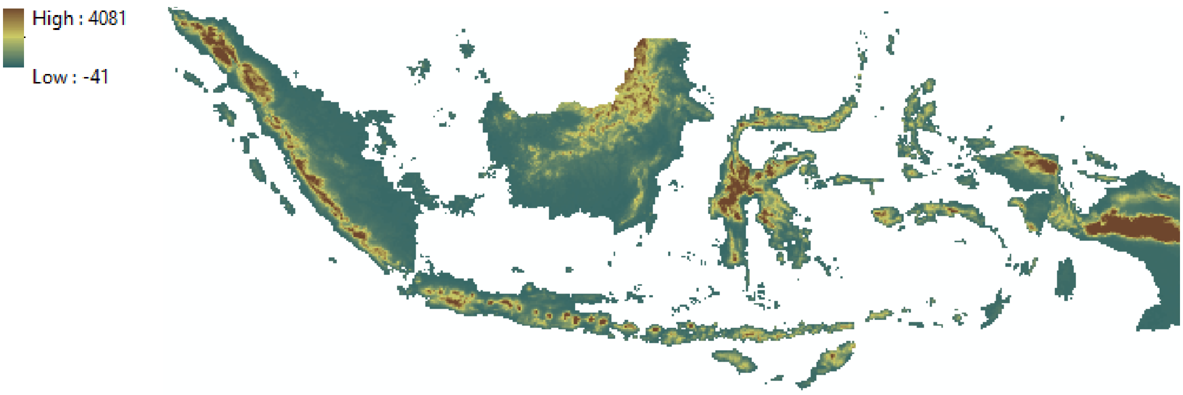

Figure 2.3 Map of Indonesia showing altitude (m) aggregated from finer-resolution elevation raster for each SimU.

<u>Slopes</u>

Slope inclinations (both degree and per cent) were calculated from the 3-arsec Shuttle Radar Topography Mission (SRTM, void filled by CGIAR .4.1, , Jarvis et al. 2008; Reuter, Nelson, and Jarvis 2007). Because of the angle distances used with input elevation model (decimal degree) a correction factor (Z-factor) was used to get correct vertical distances for slope calculations. The Z-factor value of 0.00000905 (-) was calculated as an arithmetic mean between the published Z-factors (https://www.esri.com/arcgis-blog/products/product/imagery/setting-the-z-factor-parameter-correctly/) relevant for latitudes at 0 deg and 10 deg, respectively. Calculated slopes (degree) were classified into 6 classes (Table 2.1, HRU, see chapter 2.2.3) and dominant class assigned to each SimU (Figure 2.4).

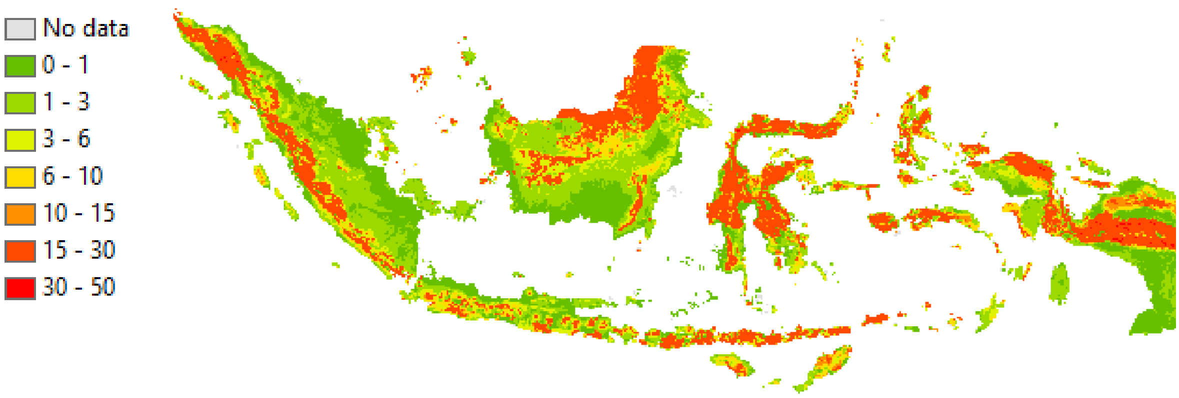

Figure 2.4 Map of Indonesia showing slope (deg) aggregated from finer-resolution slope raster for each SimU.

2.2.2 Soil data

Harmonized World Soil Dataset v1.1 (further referred to as HWSD, FAO/IIASA/ISRIC/ISS-CAS/JRC 2009) was taken as currently most up-to-date and official soil data product agreed and provided by the global soil community as a part of Global Soil Information System by FAO.

HWSD consist of polygons of soil mapping units (also represented as a raster), each described by one-to-many individual soil typological units and their relative area share and rank. Each soil typological unit is described via typical soil profile which provides further data on soil classification (soil name) and topsoil and subsoil analytical values of important soil properties, such as soil layer depth, clay, sand, stone or soil organic carbon content, soil pH, or cation exchange capacity.

Totally, 48,148 unique combinations of soil mapping units and associated typical soil profiles are recorded in HWSD and assigned with a unique ID. As a result, the same record of analytical values for subsoil and topsoil properties (typical soil profile) is in the HWSD data table repeated one-to-many times under different IDs.

An alternative list of typical soil profiles was created from the HWSD data table, listing each typical soil profile only once and assigning it a new unique ID (MySU_ID). A final list of N = 6,871 unique MySU_ID combinations was then used throughout the IND SimU v0.1 to link HWSD soil data to the SimU. A typical soil profile (MySU_ID) with highest area share and rank was selected from the HWSD soil mapping unit spatially intersected by the SimU centroid.

Soil hydrological properties values (water content at filed water capacity, water content at wilting point, and available water capacity) were additionally calculated for topsoil and subsoil layers of all typical soil profiles with pedo-transfer function (PTF) of Rawls, Brakensiek, and Saxtonn (1982).

A topsoil texture class was interpreted from HWSD data for each MySU_ID and assigned to each SimU (Figure 2.5) as an input for classification of the homogenous response units (see chapter 2.2.3).

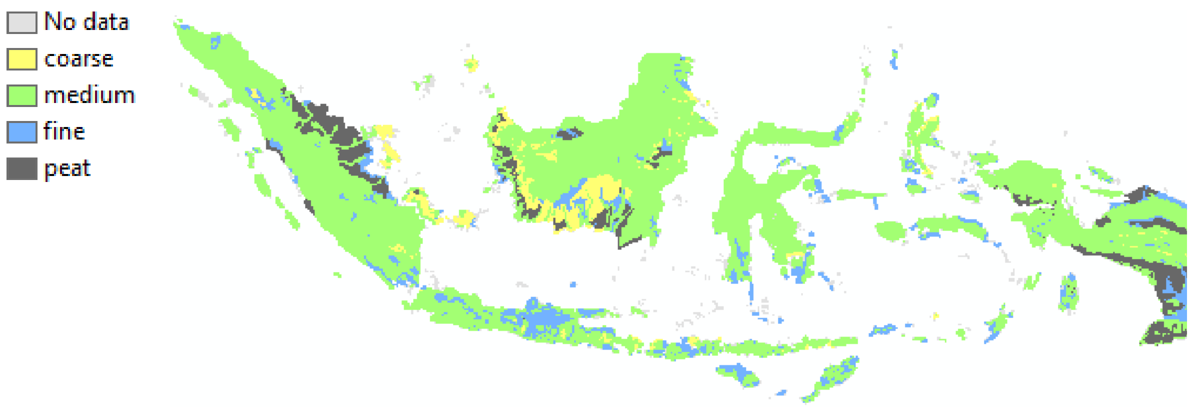

Figure 2.5 Map of Indonesia showing topsoil texture class for each SimU (coarse textured/light soils – sandy soils, medium textured soils – loamy or silty soils, fine textured/heavy soils – mostly clays, peat – organic soils).

2.2.3 Other data

Administrative units

The administrative unit layer for Indonesia (BPS - Statistics Indonesia, https://www.bps.go.id/) is a national spatial database of the administrative units of Indonesia. It was used to get sub-country border information (provinces, islands) for the SimU. The province borders polygons were intersected with SimU polygons and the information (the code and the name) about the area dominant province was assigned to each SimU (Figure 2.6).

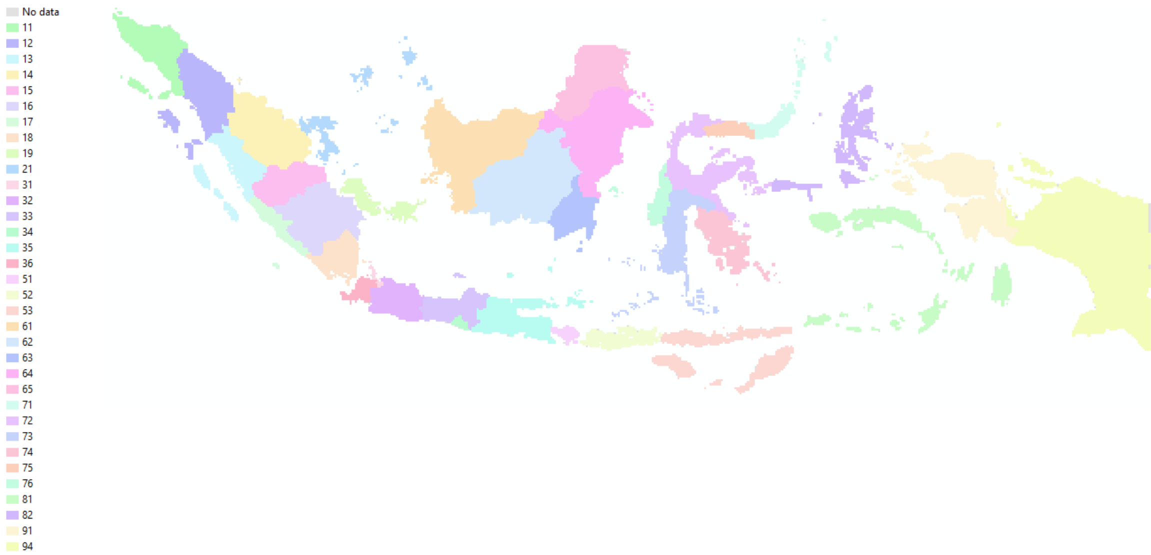

Figure 2.6 Area dominant province information (the BPS province numeric code) assigned to each SimU.

Homogenous response units (HRU)

The Homogenous Response Units (HRU) concept has been used with the GLOBIOM model legacy global dataset (Skalský et al. 2008) to provide land use optimization algorithms with landscape units defined by stable land characteristics which are only hard to be changed by climate or land management practices, and thus enabling for consistent analysing of the land use response to its assumed drivers.

Table 2.1 Altitude, slope and topsoil texture class criteria used for HRU classification.

Land characteristics	Unit	Class (class interval)
Altitude	meters	1 (0 – 300), 2 (300 – 600), 3 (600 – 1100), 4 (1100 – 2500), 5 (> 2500)
Slope	degree	1 (0-3), 2 (3-6), 3 (6-10), 4 (10-15), 5 (15-30), 6 (30-50), 7 (>50)
Soil	-	1 (coarse textured/light soils - sandy), 2 (medium textured soils - loamy, silty), 3 (fine textured/heavy soils - clays), 4 (organic soils/peat)

The concept of HRU was implemented with the GLOBIOM model legacy global dataset via the classification criteria based on the terrain and soil characteristics. A slightly modified classification (Table 2.1) was also used with each SimU using harmonised data on elevation, slope, and topsoil texture class, resulting classification of HRU for whole of Indonesia (Figure 2.7).

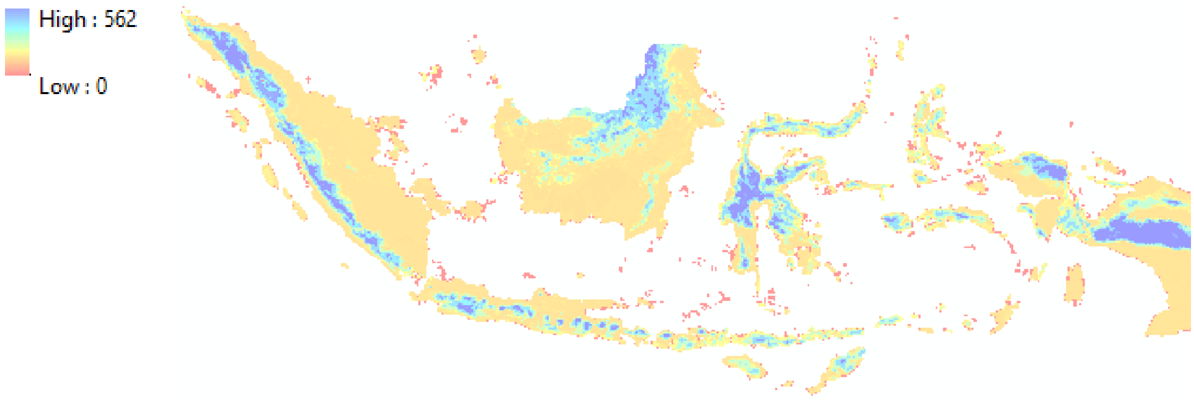

Figure 2.7 HRU map of Indonesia showing the SimU with the same or very similar topography (altitude, slope) and soil (topsoil texture) conditions which are only hard to be changed by climate or land use. Three-digit code with HRU = 0 for all SimU with missing topography or soil information.

<u>Legacy simulation units identification (SimU_GEOBENE)</u>

The spatial simulation units (SimU) were introduced with the GLOBIOM model legacy global dataset (Skalský et al. 2008) as a viable way of compromising the spatial resolution of the inputs and outputs of the bio-physical models (5 arcmin, wherever useful) and computational capabilities of the available infrastructure to handle the required number of simulations and storage of the simulation outputs. A set of about 200k SimU spatial units globally resulted from intersection between the 30 arcmin grid cells (identified by column and row address – Col_Row30), the 5 arcmin spatial resolution based HRU spatial units, and country borders, yielding spatial units ranging between 5 and 30 arcmin in their size.

The SimU polygons from the GLOBIOM legacy global dataset were intersected with each SimU set for Indonesia and identification number of the spatially corresponding SimU_GEOBENE read to the SimU centroid points. Due to slightly different assumptions and input data used for global dataset and regional spatial data infrastructure for Indonesia, some SimU did not have corresponding SimU_GEOBENE. The Col_Row30 identification was assigned to each SimU via SimU_GEOBENE related attributes from the GLOBIOM model legacy global dataset.

3. Data description

The Spatial simulation infrastructure for Indonesia v.0.1 (IND SimU v.0.1) data product consist of:

- the spatial reference data (grid and lattice) organized in two separate GIS layers (point and polygon features) covering Indonesia, and
- the table data organized in several separate (structured) text files corresponding to the spatial reference data which further describe the spatial elements (polygon/point feature attribute data).

3.1 Spatial reference data

The spatial reference data consist of regular spatial grids and geographically corresponding centroid point lattices at spatial resolutions of 5.0 (about 9.25 km on equator).

The geographical data of the IND SimU v.0.1 dataset covers the national territory of Indonesia. The spatial reference data coverage is defined by the geographical bounding box (decimal degrees) of S 11.0833, E 95.0000 (lower left), and N 6.0833, E 141.0833 (upper right) corners.

The list and description of the LMS RefGrids v.0.1 spatial reference data layers is in Table 3.1 with further details on attribute table fields in Table 3.2.

Table 3.1 List and general description of the IND SimU v.0.1 spatial reference data layers.

File name	Data format (feature type)	Description
IND_SimU_v01	ESRI Shape (polygon)	Borders of SimU organized in a grid of regular square spatial elements with 5 arcmin spatial resolution, each polygon described by SimUID, Col_Row, XLON, and YLAT.
IND_SimU_v01_cnt	ESRI Shape (point)	Centroid points of the SimU polygons organized in a lattice of points, each point described by SimUID, Col_Row, XLON, and YLAT.

Table 3.2 List and general description of the IND SimU v.0.1 spatial reference data attribute table fields.

Data field	Data type	Description
SimUID	Integer	Identification number of the spatial simulation unit for running process-based models (SimU), each SimU defined as one grid cell/centroid of the grid cell of the 5 arcmin spatial resolution grid.
Col_Row	Text	Alternative identification of the SimU. The Colum and row address of the grid cell counted from the upper left corner of the global coverage 5 arcmin grid, can be used as a link to legacy dataset used with the GLOBIOM model.
XLON	Double	Longitude (decimal degree), x-coordinate of the SimU centroid.
YLAT	Double	Latitude (decimal degree), y-coordinate of the SimU centroid.

3.2 Table data

Table data further describe spatial elements (polygon/point features) and provide users/models with a data on real areas, land surface real areas, administrative borders, and the terrain, soil, and land cover conditions within the grid cell borders (or at the centroid point).

The records in the data tables either directly correspond (one-to-one) to the individual grid cells or grid cell centroids (attribute tables) or provide an extended description of the classes used in the attribute tables (attribute domain tables).

The list and brief description of the LMS RefGrids v.0.1 attribute, attribute domain, and mapping tables is in Table 3.3. Detailed description of the individual data table attributes is in Table 3.4 - Table 3.7.

Table 3.3 List and general description of the IND SimU v.0.1 attribute and attribute domain tables.

File name	Data format	Description
IND_SimU_v01	Text file (semi-colon delimited)	Attribute table, description of the SimU, table content further described in Table 3.4.
D_ADM1	Text file (semi-colon delimited)	Attribute domain table, list of administrative units of Indonesia used for SimU description, table content further described in Table 3.5.
D_SLOPE	Text file (semi-colon delimited)	Attribute domain table, list of slope class definitions used for SimU description, table content further described in Table 3.6.
D_MySU_ID_HWSD11	Text file (semi-colon delimited)	Attribute domain table, list of soil units and soil profile data used for SimU descrition, table content further described in Table 3.7.

Table 3.4 List and general description of the IND SimU v.0.1 attribute table fields.

Data field	Data type	Description
SimUID	Integer	Identification number of the spatial simulation unit for running process-based models (SimU), each SimU defined as one grid cell/centroid of the grid cell of the 5 arcmin spatial resolution grid.
Col_Row	Text	Alternative identification of the SimU, the column and row address of the grid cell counted from the upper left corner of the global coverage 5 arcmin grid, can be used as a link to legacy dataset used with the GLOBIOM model.
XLON	Double	Longitude (decimal degree), x-coordinate of the SimU centroid.
YLAT	Double	Latitude (decimal degree), y-coordinate of the SimU centroid.
SimU_AreaHa	Double	Real area of the SimU (ha), calculated after reprojecting the geographical 5 arcmin spatial resolution grid into the equal-area geographic regional projection (Albers equal-area conic projected coordinate system for South Asia, https://spatialreference.org/ref/esri/asia-south-albers-equal-area-conic/).
Land_AreaPerc	Double	Relative area (%) of the SimU covered by land, estimated with the coastlines from the administrative unit layer for Indonesia (BPS - Statistics Indonesia, https://www.bps.go.id/).
Sea_AreaPerc	Double	Relative area (%) of the Simu covered by sea, estimated with the coastlines from the administrative unit layer for Indonesia (BPS - Statistics Indonesia, https://www.bps.go.id/).

Data field	Data type	Description
ADM1	Integer	Code of the 1st level sub-national administrative unit with the highest area share within the SimU (BPS - Statistics Indonesia, https://www.bps.go.id/), administrative units names in the separate table D_ADM1 (Table 3.5).
HRU	Integer	Three-digit numerical code of Homogenous Response Unit - HRU used with the GLOBIOM model legacy dataset to group grid cells based on the local conditions only hardly to be changed by climate or management, if HRU = -999 then altitude, slope, or soil information is missing for the SimU.
ALTITUDE	Integer	Mean altititude (m) of the SimU, if ALTITUDE = -999 then altitude information is missing for the SimU.
SLOPECLS	Integer	Area dominant slope inclination class within the SimU calculated from the underlying 3 arcsec SRTM v4.1 based slope raster, slope class definitions in the separate table D_SLOPE (Table 3.6), if SLOPECLS = -999 then slope inclination information is missing for the SimU.
SLOPEASP	Integer	Median of the slope aspect (deg) within the SimU calculated from underlying 3 arcsec SRTM v4.1 based slope aspect raster, if SLOPEASP = -999 then slope aspect information is missing for the SimU.
MySU_ID	Integer	Identification number of the area dominant unique combination of the HWSD v1.1 soil typological unit parameters (both soil classification and soil profile description) within the SimU, list of MySU_IDs and HWSD v1.1 data in the separate table D_MySU_ID_HWSD11 (Table 3.7), if MySU_ID = -999 then soil information is missing for the SimU.
MySU_ID_AreaPer	Double	Relative area (%) of the dominant MySU_ID from the total SimU area, if MySU_ID_AreaPer = -999 then soil information is missing for the SimU.
SimUID_GEOBENE	Integer	Link to legacy dataset used with GLOBIOM model, identification number of the global coverage spatial simulation unit, if SimUID_GEOBENE = 0 not any spatially corresponding legacy SimU_GEOBENE exists for the SimU.
ColRow30_GEOBENE	Text	Link to legacy dataset used with GLOBIOM model, the column and row address of the global coverage 30 arcmin spatial resolution grid cell, if ColRow30_GEOBENE = 0 not any spatially corresponding legacy 30 arcmin grid cell exists for the SimU.

Table 3.5 List and general description of the D_ADM1 attribute domain table fields.

Data field	Data type	Description
ADM1	Integer	Code of the 1st level sub-national administrative unit with the highest area share within the SimU (BPS - Statistics Indonesia, https://www.bps.go.id/).
ADM1_NAME	Text	Name of the 1st level sub-national administrative unit (BPS - Statistics Indonesia, https://www.bps.go.id/).
ISLAND	Text	Name of the main island on which the 1st level sub-national administrative unit is located (BPS - Statistics Indonesia, https://www.bps.go.id/).

Table 3.6 List and general description of the D_SLOPE attribute domain table fields.

Data field	Data type	Description
SLOPECLS	Integer	Area dominant slope inclination class within the SimU calculated from the underlying 3 arcsec SRTM v4.1 based slope raster.

Data field	Data type	Description
HRUCLS	Integer	Code for slope part of Homogenous Response Unit - HRU used with GLOBIOM model legacy dataset, based on slope inclination value (deg): 1 (0-3), 2 (3-6), 3 (6-10), 4 (10-15), 5 (15-30), 6 (30-50), 7 (>50).
DEGLOW	Double	Slope inclination (deg), slope class lower interval.
DEGUP	Double	Slope inclination (deg), slope class upper interval.
DEGMID	Double	Slope inclination (deg), slope class mid of interval.
PERMID	Double	Slope inclination (%), slope class mid of interval.
UPSL	Integer	Slope length (m), an arbitrary set value of the slope length associated with the slope class, an input for EPIC-IIASA model.
WSA	Integer	Field size (ha), an arbitrary set value of the typical field size associated with the slope class, an input for EPIC-IIASA model.

Table 3.7 List and general description of the D_MySU_ID_HWSD11 attribute domain table fields.

Data field	Data type	Description
MySU_ID	integer	Identification number of the unique combination of HWSD v1.1 soil typological unit parameters (both soil classification and soil profile description).
ISSOIL	integer	HWSD v1.1 attribute, identifies if the soil unit is the soil (1) or non-soil (0) body.
SU_SYM	short text	HWSD v1.1 attribute, Soil Mapping Unit Symbol.
SU_SYM_META	integer	HWSD v1.1 derived attribute, identification of soil classification used for SMU symbol: 74 (FAO 1974), 85 (FAO 1985), 90 (FAO 1990).
HG	integer	PTR derived attribute, USDA soil hydrological group, code (1 = A, 2 = B, 3 = C, 4 = D), derived from HWSD v1.1 SU_SYM and T_TEXTURE by pedo-transfer rules.
FC_TOP	double	PTF derived attribute, water content ($mm*mm^{-1}$) in topsoil at field water capacity (pF = 1.8), derived by pedotransfer function Rawls et al. (1982) Estimation of Soil Water Properties. Transactions of the ASAE. 25 (5): 1316-1320. (doi: 10.13031/2013.33720) applied to HWSD v1.1 data.
FC_SUB	double	PTF derived attribute, water content ($mm*mm^{-1}$) in subsoil at field water capacity (pF = 1.8), derived by pedotransfer function (Rawls et al. (1982) Estimation of Soil Water Properties. Transactions of the ASAE. 25 (5): 1316-1320. (doi: 10.13031/2013.33720) applied to HWSD v1.1 data.
WP_TOP	double	PTF derived attribute, water content ($mm*mm^{-1}$) in topsoil at wilting point (pF = 4.2), derived by pedotransfer function (Rawls et al. (1982) Estimation of Soil Water Properties. Transactions of the ASAE. 25 (5): 1316-1320. (doi: 10.13031/2013.33720) applied to HWSD v1.1 data.
WP_SUB	double	PTF derived attribute, water content ($mm*mm^{-1}$) in subsoil at wilting point (pF = 4.2), derived by pedotransfer function (Rawls et al. (1982) Estimation of Soil Water Properties. Transactions of the ASAE. 25 (5): 1316-1320. (doi: 10.13031/2013.33720) applied to HWSD v1.1 data.
AWC	double	PTF derived attribute, available water capacity (mm) in the 1m deep soil profile, calculated from FC_TOP/SUB and WP_TOP/SUB.
AWCred	double	PTF derived attribute, available water capacity (mm) in the 1m deep soil profile reduced by stone content, calculated from FC_TOP/SUB and WP_TOP/SUB and T_GRAVEL, S_GRAVEL.

Data field	Data type	Description
SOIL_HRU	integer	PTR derived code for soil part of Homogenous Response Unit (HRU) - a concept used with GLOBIOM model legacy dataset to group grid cells based on the local terrain and soil conditions only hard to be changed by land management, based on HWSD v1.1 T_TEXTURE, REF_DEPTH, T_GRAVEL and S_GRAVEL values, rules: IF REF_DEPTH < 100 OR ((T_GRAVEL+SGRAVEL)/2) > 35 THEN 4, IF T_TEXTURE = 0 THEN 5, ELSE T_TEXTURE) where T_TEXTURE 1 = coarse, 2 = medium, 3 = fine, 4 = organic.
TOP_DEPTH	integer	PTR derived topsoil depth (cm) as used for the EPIC-IIASA model, based on HWSD v1.1 REF_DEPTH value, rules: IF HWSD REF_DEPTH = 10 THEN 10, IF HWSD REF_DEPTH = 30 THEN 20, ELSE 30.
SUB_DEPTH	integer	PTR derived subsoil depth (cm) as used for the EPIC-IIASA model, based on HWSD v1.1 REF_DEPTH value, rules: IF HWSD REF_DEPTH = 10 THEN 30, IF HWSD REF_DEPTH = 30 THEN 30, ELSE 100.
T_TEXTURE	integer	HWSD v1.1 attribute, Topsoil Texture, code
DRAINAGE	integer	HWSD v1.1 attribute, Reference Soil Depth, code
REF_DEPTH	integer	HWSD v1.1 attribute, Drainage class (FAO), code
AWC_CLASS	integer	HWSD v1.1 attribute, Available Water Capacity Range, code
PHASE1	integer	HWSD v1.1 attribute, PHASE1, code
PHASE2	integer	HWSD v1.1 attribute, PHASE2, code
ROOTS	integer	HWSD v1.1 attribute, Obstacles to Roots (ESDB), code
IL	integer	HWSD v1.1 attribute, Impermeable Layer (ESDB), code
SWR	integer	HWSD v1.1 attribute, Soil Water Regime (ESDB), code
ADD_PROP	integer	HWSD v1.1 attribute, Other properties (gelic, vertic, petric), code
T_GRAVEL	integer	HWSD v1.1 attribute, Topsoil Gravel Content, %vol.
T_SAND	integer	HWSD v1.1 attribute, Topsoil Sand Fraction, % wt.
T_SILT	integer	HWSD v1.1 attribute, Topsoil Silt Fraction, % wt.
T_CLAY	integer	HWSD v1.1 attribute, Topsoil Clay Fraction, % wt.
T_USDA_TEX_CLASS	integer	HWSD v1.1 attribute, Topsoil USDA Texture Class, code
T_REF_BULK_DENSITY	double	HWSD v1.1 attribute, Topsoil Reference Bulk Density, kg/dm^3
T_BULK_DENSITY	double	HWSD v1.1 attribute, Topsoil Bulk Density, kg/dm^3
T_OC	double	HWSD v1.1 attribute, Topsoil Organic Carbon, % wt.
T_PH_H2O	double	HWSD v1.1 attribute, Topsoil pH (H2O), -log(H+)
T_CEC_CLAY	integer	HWSD v1.1 attribute, Topsoil Cation Exchange Capacity (clay), cmol/kg
T_CEC_SOIL	integer	HWSD v1.1 attribute, Topsoil Cation Exchange Capacity (soil), cmol/kg
T_BS	integer	HWSD v1.1 attribute, Topsoil Base Saturation, %
T_TEB	integer	HWSD v1.1 attribute, Topsoil Total Exchangeable Bases, cmol/kg
T_CACO3	double	HWSD v1.1 attribute, Topsoil Calcium Carbonate, % wt.
T_CASO4	double	HWSD v1.1 attribute, Topsoil Gypsum, % wt.
T_ESP	integer	HWSD v1.1 attribute, Topsoil Sodicity (Exchangeable Sodium Percentage), %
T_ECE	double	HWSD v1.1 attribute, Topsoil Salinity (Electrical Soil Conductivity), dS/m
S_GRAVEL	integer	HWSD v1.1 attribute, Subsoil Gravel Content, %vol.
S_SAND	integer	HWSD v1.1 attribute, Subsoil Sand Fraction, % wt.
S_SILT	integer	HWSD v1.1 attribute, Subsoil Silt Fraction, % wt.
S_CLAY	integer	HWSD v1.1 attribute, Subsoil Clay Fraction, % wt.
S_USDA_TEX_CLASS	integer	HWSD v1.1 attribute, Subsoil USDA Texture Class, code
S_REF_BULK_DENSITY	double	HWSD v1.1 attribute, Subsoil Reference Bulk Density, kg/dm^3
S_BULK_DENSITY	double	HWSD v1.1 attribute, Subsoil Bulk Density kg/dm^3
S_OC	double	HWSD v1.1 attribute, Subsoil Organic Carbon, % wt.
S_PH_H2O	double	HWSD v1.1 attribute, Subsoil pH (H2O), -log(H+)
S_CEC_CLAY	integer	HWSD v1.1 attribute, Subsoil Cation Exchange Capacity (clay), cmol/kg

Data field	Data type	Description
S_CEC_SOIL	integer	HWSD v1.1 attribute, Subsoil Cation Exchange Capacity (soil), cmol/kg
S_BS	integer	HWSD v1.1 attribute, Subsoil Base Saturation, %
S_TEB	integer	HWSD v1.1 attribute, Subsoil Total Exchangeable Bases, cmol/kg
S_CACO3	double	HWSD v1.1 attribute, Subsoil Calcium Carbonate, % wt.
S_CASO4	double	HWSD v1.1 attribute, Subsoil Gypsum, % wt.
S_ESP	integer	HWSD v1.1 attribute, Subsoil Sodicity (Exchangeable Sodium Percentage), %
S_ECE	double	HWSD v1.1 attribute, Subsoil Salinity (Electrical Soil Conductivity), dS/m

4. Limitations and potential impact

Primary purpose of the IND SimU v.0.1 dataset was to provide bio-physical and economic optimization models (EPIC-IIASA, WaNuLCAS, G4M, and GLOBIOM) with effective interface for harmonization of geographical and multi-layered data from different resources and modelling inputs/outputs exchange, and minimum set of mandatory terrain and slope data for running the models. In more details, the focus of data harmonization and interpretation was put on:

- providing consistent coverage of whole territory of Indonesia with geographical reference units for harmonizing and displaying all input and output (geographical) information for the models – a regular grid of 5 arcmin spatial resolution square elements (Spatial simulation units – SimU), and
- aggregating of underlying and finer-scale geographical information about terrain (altitude, slope inclination, and slope aspect) and soil (dominant soil typological unit) characteristics and providing a link to other multi-layered data (land use and agricultural statistics) via information about sub-national administrative unit attributed to each SimU.

Yet consistent and comprehensive in its spatial representation and attribute content, the IND SimU v.0.1 dataset use is strictly limited to its primary purpose and should not be used as a replacement of original fine-scale data for Indonesia. This mainly due to the assumptions and necessary spatial and attribute aggregations and simplifications done and resulting in lower geographical and attribute quality and accuracy compared to original global or national datasets.

To a limited extent, the spatial infrastructure, as well as attribute information provided with the spatial elements (SimU), is comparable and possible to be linked to the earlier large-scale products of IIASA used for bio-physical modelling and economic optimization of land use at global level (Skalský et al. 2008).

References

FAO/IIASA/ISRIC/ISS-CAS/JRC. 2009. "Harmonized World Soil Dataset - Major Soil Groups | Data Basin." 2009. https://databasin.org/datasets/92cd21cec56a48289ebaeccaa75d6ada/.

Jarvis, A., E. Guevara, H. I. Reuter, and A. D. Nelson. 2008. "Hole-Filled SRTM for the Globe : Version 4 : Data Grid." https://research.utwente.nl/en/publications/hole-filled-srtm-for-the-globe-version-4-data-grid.

Rawls, W. J., D. L. Brakensiek, and K. E. Saxtonn. 1982. "Estimation of Soil Water Properties. Transactions of the ASAE" 25 (5): 1316–20. https://doi.org/10.13031/2013.33720.

Reuter, H. I., A. Nelson, and A. Jarvis. 2007. "An Evaluation of Void-filling Interpolation Methods for SRTM Data." *International Journal of Geographical Information Science* 21 (9): 983–1008. https://doi.org/10.1080/13658810601169899.

Skalský, Rastislav, Zuzana Tarasovičová, Juraj Balkovič, Erwin Schmid, Michael Fuchs, Georg Kindermann, and Peter Scholtz. 2008. "GEO-BENE Global Database for Bio-Physical Modeling v. 1.0," 58.