ORNL DAAC
Distributed Active Archive Center for Biogeochemical Dynamics

NASA

Data

NACP MsTMIP: Unified North American Soil Map

Get Data

Revision Date: August 21, 2014

Summary:

This data set provides soil maps for the United States (US) (including Alaska), Canada, Mexico, and a part of Guatemala. The map information content includes maximum soil depth and eight soil attributes including sand, silt, and clay content, gravel content, organic carbon content, pH, cation exchange capacity, and bulk density for the topsoil layer (0-30 cm) and the subsoil layer (30-100 cm). The spatial resolution is 0.25 degree.

The Unified North American Soil Map (UNASM) combined information from the state-of-the-art US General Soil Map (STATSGO2) and Soil Landscape of Canada (SLCs) databases. The area not covered by these data sets was filled by using the Harmonized World Soil Database version 1.21 (HWSD1.21). The Northern Circumpolar Soil Carbon (NCSCD) database was used to provide more accurate and up-to-date soil organic carbon information for the high-latitude permafrost region and was combined with soil organic carbon content derived from the UNASM (Liu et al., 2013).

The UNASM data were utilized in the North American Carbon Program (NACP) Multi-Scale Synthesis and Terrestrial Model Intercomparison Project (MsTMIP) as model input driver data (Huntzinger et al., 2013, Wei et al., 2013). The driver data (Wei et al., 2013, 2014) were used by 22 terrestrial biosphere models to run baseline and sensitivity simulations. The compilation of these data was facilitated by the NACP Modeling and Synthesis Thematic Data Center (MAST-DC). MAST-DC was a component of the NACP (www.nacarbon.org) designed to support NACP by providing data products and data management services needed for modeling and synthesis activities.

There are 18 files provided in GeoTIFF (.tif) format with this data set: 16 maps of the eight soil attributes (one map per each attribute per soil layer), one map of maximum soil depth from the soil map sources, and one map of the percentage of the area of the first dominant soil component to the area of the cell.

Data and Documentation Access:

Get Data: http://daac.ornl.gov/cgi-bin/dsviewer.pl?ds_id=1242

Supplemental Information: There are two companion files with this data set:

- Liu_et_ al_2013.pdf: Provides additional attribute descriptions in Appendix A of the publication.
- NACP_MsTMIP_Unified_NA_Soil_Map.pdf: A .pdf of this document.

Related Data Products: Wei, Y., et al. 2014. NACP MsTMIP: Global and North American Driver Data for Multi-Model Intercomparison. http://dx.doi.org/10.3334/ORNLDAAC/1220

Data Citation:

Cite this data set as follows:

Table of Contents:

1. Data Set Overview:

Project: North American Carbon Program

This data set was collected as part of the NACP MsTMIP campaign. NACP (Denning et al., 2005; Wofsy and Harris, 2002) is a multidisciplinary research program to obtain scientific understanding of North America's carbon sources and sinks and of changes in carbon stocks needed to meet societal concerns and to provide tools for decision makers. Successful execution of the NACP has required an unprecedented level of coordination among observational, experimental, and modeling efforts regarding terrestrial, oceanic, atmospheric, and human components. The project has relied upon a rich and diverse array of existing observational networks, monitoring sites, and experimental field studies in North America and its adjacent oceans. It is supported by a number of different federal agencies through a variety of intramural and extramural funding mechanisms and award instruments.

This data set provides topsoil layer (0-30 cm) and subsoil layer (30-100 cm) maps of the US (including Alaska), Canada, Mexico, and a part of Guatemala. The maps include eight soil attributes: sand, silt, and clay content, gravel content, organic carbon content, pH, cation exchange capacity, and bulk density for each soil depth.

The Unified North American Soil Map (UNASM) combined information from the state-of-the-art US General Soil Map (STATSGO2) and Soil Landscape of Canada (SLCs) databases. The area not covered by these data sets was filled by using the Harmonized World Soil Database version 1.21 (HWSD1.21). The Northern Circumpolar Soil Carbon (NCSCD) database was used to provide more accurate soil organic carbon information for the high-latitude permafrost region and was combined with soil organic carbon content derived from the UNASM (Liu et al., 2013).

The UNASM data were utilized in the North American Carbon Program (NACP) Multi-Scale Synthesis and Terrestrial Model Intercomparison Project (MsTMIP) as model input driver data (Huntzinger et al., 2013, Wei et al., 2013). The driver data (Wei et al., 2013, 2014) were used by 22 terrestrial biosphere models to run baseline and sensitivity simulations. The compilation of these data was facilitated by the NACP Modeling and Synthesis Thematic Data Center (MAST-DC). MAST-DC was a component of the NACP (www.nacarbon.org) designed to support NACP by providing data products and data management services needed for modeling and synthesis activities.

2. Data Characteristics:

Temporal Coverage

Compilation from published soil databases between 1993/01/01 and 2010/12/31.

Temporal Resolution

None

Spatial Coverage

This data set provides soil maps for the US (including Alaska), Canada, Mexico, and a part of Guatemala.

Spatial Resolution

0.25-degree grid with a total of 142,080 North American grid cells.

Site boundaries: (All latitude and longitude given in decimal degrees)

Site (Region)	Westernmost Longitude	Easternmost Longitude	Northernmost Latitude	Southernmost Latitude
North America	-170	-50	84	10

Data File Information

There are 18 files with soil attribute information in this data set in GeoTIFF (.tif) format. There is one file for each attribute for the topsoil and subsoil layers, one file depicting maximum soil depth from the soil map sources, and one map of the percentage of the area of the first dominant soil component to the area

of the cell.

Table 1. Soil Attributes of topsoil layer (0– 30 cm) and subsoil layer (30–100 cm). For complete descriptions of the attributes and their synthesis from various sources, please see the companion file, Liu et al., 2013, Appendix A.

Soil Attribute	File_Name	Abbreviation	Units	General Description
First Dominant Soil Component	Unified_NA_Soil_Map_First_Dominant_Soil_Component_Area_Percentage.tif	sc	%	The percentage of the area of the first dominant soil component to the area of the cell
Maximum Soil Depth	Unified_NA_Soil_Map_Maximum_Soil_Depth.tif	Soil Depth	cm	Soil depth is the maximum soil depth of each cell before harmonization to standard layers. Maximum soil depth can be used as the approximate measure of the depth-to-bedrock.
Topsoil Sand Fraction	Unified_NA_Soil_Map_Topsoil_Sand_Fraction.tif	sand	% weight	Sand content is defined as the percentage of this particle size class based on weight. Sand comprises particles or granules ranging in diameter from 0.05mm to 2.0 mm.
Topsoil Silt Fraction	Unified_NA_Soil_Map_Topsoil_Silt_Fraction.tif	silt	% weight	Silt content is defined as the percentage of this particle size class based on weight. Silt size is between 0.002 and 0.0625 (In STATSGO2, the silt size is between 0.002 and 0.05 mm).
Topsoil Clay Fraction	Unified_NA_Soil_Map_Topsoil_Clay_Fraction.tif	clay	% weight	Clay content is defined as the percentage of this particle size class based on weight. Clay is composed primarily of fine-grained with the diameter less than 0.002 mm.
Topsoil Gravel Fraction	Unified_NA_Soil_Map_Topsoil_Gravel_Content.tif	gravel	% volume	Gravel content is the volume percentage gravel (diameter> 2 mm) in the soil. The depth-weighted average method was used to interpolate gravel content of different layers into two standard layers.
Topsoil Organic Carbon	Unified_NA_Soil_Map_Topsoil_Organic_Carbon.tif	oc	% weight	Soil organic carbon concentration is defined as the percentage of soil weight. The mass-weighted average of soil organic carbon concentration was calculated when the unified soil dataset was harmonized into two standard layers.
				pH is a measure for the acidity and alkalinity of the soil. pH was

Topsoil pH (H$_2$O)	Unified_NA_Soil_Map_Topsoil_pH.tif	ph		measured in a soil-water solution. The depth-weighted average method was used to interpolate pH values of different layers into two standard layers.
Topsoil Bulk Density	Unified_NA_Soil_Map_Topsoil_Reference_Bulk_Density.tif	bd	g cm^{-3}	Bulk density is the ratio of the mass of soil material to the total volume of solids plus pores. Since bulk density is measured and estimated using different methods, potential bias among the bulk density measurements from different sources might exist.
Topsoil Cation Exchange Capacity	Unified_NA_Soil_Map_Topsoil_Cation_Exchange_Capacity.tif	cec	meq/100g	Cation exchange capacity represents the total nutrient fixing capacity of a soil. Soils with low CEC have little resilience and cannot build up stores of nutrients. CEC values are represented as milliequivalents of hydrogen per 100 g of dry soil.
Subsoil Sand Fraction	Unified_NA_Soil_Map_Subsoil_Sand_Fraction.tif	sand	% weight	All descriptions for the subsoil attributes are the same as for the topsoil attributes
Subsoil Silt Fraction	Unified_NA_Soil_Map_Subsoil_Silt_Fraction.tif	silt	% weight	
Subsoil Clay Fraction	Unified_NA_Soil_Map_Subsoil_Clay_Fraction.tif	clay	% weight	
Subsoil Gravel Fraction	Unified_NA_Soil_Map_Subsoil_Gravel_Content.tif	gravel	% volume	
SubsoilOrganic Carbon	Unified_NA_Soil_Map_Subsoil_Organic_Carbon.tif	oc	% weight	
Subsoil pH (H$_2$O)	Unified_NA_Soil_Map_Subsoil_pH.tif	ph		
Subsoil Bulk Density	Unified_NA_Soil_Map_Subsoil_Reference_Bulk_Density.tif	bd	g cm^{-3}	
Subsoil Cation Exchange Capacity	Unified_NA_Soil_Map_Subsoil_Cation_Exchange_Capacity.tif	cec	meq/100g	

3. Data Application and Derivation:

The North American soil maps are useful to the modeling community for a data set of soil physical properties specifically created for North America by combining state-of-the-art soil information from STATSGO2, two versions of SLCs, and HWSD1.21. The soil data has been used in the data set: NACP MsTMIP: Global and North American Driver Data for Multi-Model Intercomparison project, archived at the ORNL DAAC.

4. Quality Assessment:

The maps were developed at 0.25 degree in latitude and longitude, which limits the flexibility for users to downscale to any spatial resolution (Liu et al., 2013).

Uniform soil taxonomy was not provided because the data were derived from different sources that used different soil taxonomies. Quantifying the uncertainty of the UNASM was complicated, as it was developed from four data sources without uncertainty for any of them. This data set uses a simple index as the approximation of the certainty of the UNASM. It provides the percentage of the area of the dominant component selected for each cell to the total land area of the cell. This percentage can be served to quantify the extent to which the selected dominant soil component represents the cell. The higher percentage implies that the selected soil component can be more representative of the soil properties within the cell and the soil information assigned for the cell is of more confidence. 51.3% of all land cells have the 1st dominant soil component that occupies over 50% of the land area of the cell. The 1st dominant soil

component selected for most of land cells can represent 40–70% of the cell area (Fig. 1).

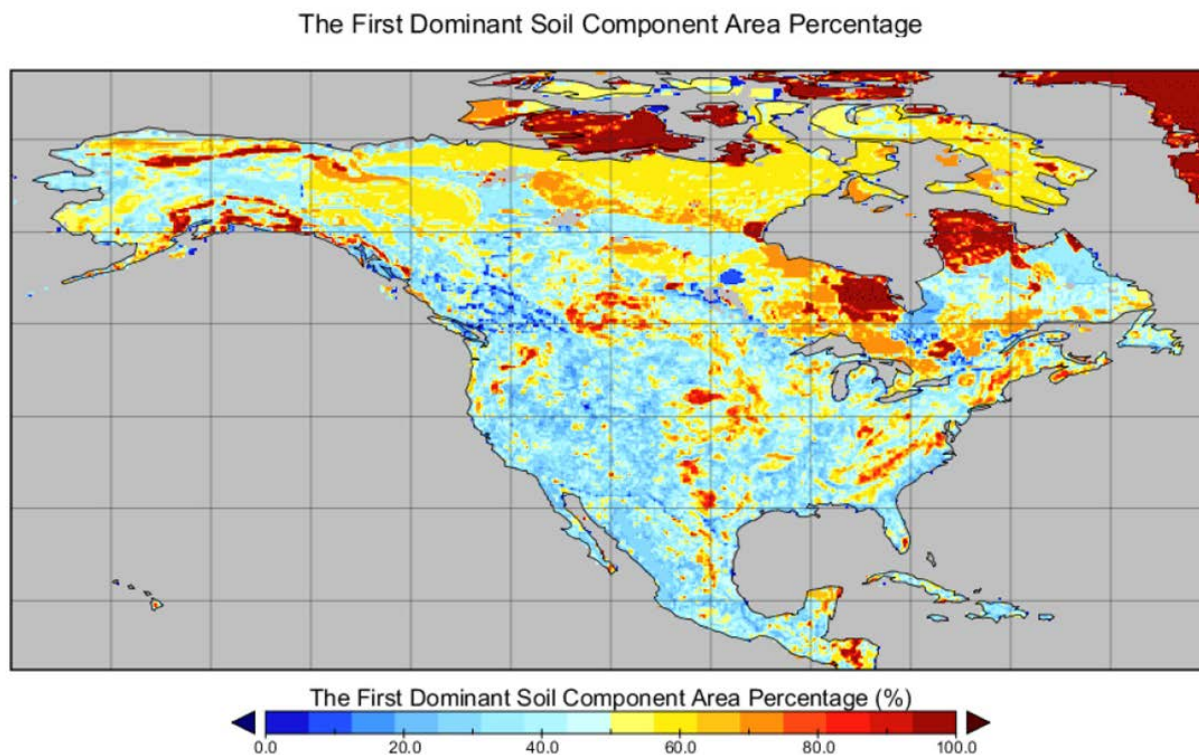## The First Dominant Soil Component Area Percentage

Figure 1. The percentage of the area of the first dominant soil component to the area of the cell (Lui et al. 2013). The source GeoTIFF file is provided as a data file Unified_NA_Soil_Map_First_Dominant_Soil_Component_Area_Percentage.tif

In terms of the source data quality, there is higher confidence in the area based on STATSGO2 and SLC3.2, and lower confidence in the area based on SLC2.2 and the area filled with HWSD1.21. Especially the high-latitudinal area filled with HWSD 1.21 was of less confidence due to the lack of observations when developing HWSD.

5. Data Acquisition Materials and Methods:

The soil maps were developed from a combination of information from the state-of-the-art US General Soil Map (STATSGO2), and the Soil Landscape of Canada (SLCs) databases. The area not covered by these data sets was filled by using the Harmonized World Soil Database version 1.21 (HWSD1.21). The Northern Circumpolar Soil Carbon (NCSCD) database was used to provide more accurate soil organic carbon information for the high-latitude permafrost region (Liu et al., 2013).

The maps were developed in a four step process: (1) STATSGO2, SLCs, and HWSD 1.21 were converted into 0.25- degree gridded format by selecting the dominant soil type in each cell; (2) STATSGO2, SLCs and HWSD1.21 were merged into a seamless map that can best represent soil in North America; (3) the soil map developed in stage 2 was harmonized into two standard layers, 0–30 cm and 30– 100 cm; and (4) Quality control. The steps are described below:

1. Conversion of STATSGO2, SLCs, and HWSD 1.21 into 0.25-degree gridded format:

- The required soil properties for STATSGO2 and SLCs were linked to the soil maps, which contained polygon features of soil and nonsoil map units on the landscape.
- The unique soil types and their combined area fractions were evaluated in a 0.25-degree cell. The soil type with the highest area fraction was selected as the dominant soil type in the cell, and the soil component was selected that had the largest area of the dominant soil type.
- The selected soil component's vertical soil layers information along with their detailed soil properties were assigned to the target 0.25-degree grid cell. The data structure of HWSD 1.21 is similar to those of STATSGO2 and SLCs. The HWSD 1.21 is represented in a 30 arc-second resolution raster map instead of vector polygons. A similar approach was applied to that used for the two polygonal data sets to select the dominant soil type and soil component for each target 0.25-degree cell and then linked to the detailed soil attributes.

2. Integration of the gridded STATSGO2, SLCs and HWSD 1.21 into a 0.25-degree North American grid with a total of 142,080 grid cells:

- The HWSD for North America has not been updated since the FAO Soil Map of the World was developed in 1970s. Thus, HWSD was given the lowest priority.
- STATSGO was first published in 1994, and in 2006 the STATSGO spatial and tabular data (which contains soil attribute values) were revised and updated, and was renamed to STATSGO2.
- SLC3.2, published in 2011, provides soil data for the significant agricultural regions of Canada. SLC3.2 is based on SLC3.0 (released in 2005), but SLC3.2 added some additional agricultural areas with new soil attribute data. SLC2.0, released in 1996, not only covers the agricultural region but also some other areas in Canada. However, SLC2.0 is less reliable than SLC3.2 based on the personal contact with the officers in Agriculture and Agri- Food

Canada.

- Based on whether the data was reliable and up-to-date, different soil data sources were merged with the following priority: STATSGO2>SLC 3.2>SLC 2.2>HWSD 1.21.
- For the cells that had both STATSGO2 and SLC3.2, the fraction of the cell within Canada and U.S. was calculated. If most of the area of a cell is within the US, the STATSGO2 data was assigned to the cell. Otherwise, SLC3.2 was assigned to the cell. The goal of this selection was to take advantage of more precise soil information from STATSGO2 and SLCs, and fill the rest of the area with the HWSD 1.21

3. In order to harmonize the vertical structure within the three data sets, the profile data were converted into two standard layers that were consistent with the HWSD. After integrating the different soil data sets, the seamless data set from Step 2 still had inconsistent vertical soil layers and the thickness of soil layers also varied across cells. To harmonize the vertical structure within these three data sets, the profile data were converted into two standard layers consistent with the HWSD 1.21:

- The top layer is from 0 to 30 cm and the sub layer is from 30 to 100 cm. The depth-weighted averaging method was used to interpolate the layers from the Step 2 data set to the two standard layers for the volume-related properties gravel fraction and pH, and the mass-weighted averaging method was used for the other soil properties that are related to soil weight.
- The standard map layers were compared to each layer in the Step 2 soil data set. If the map layer was entirely contained within a single unified soil data layer from Step 2, we used the Step 2 layer value for the standard layer. Otherwise, all layers from Step 2 that were fully or partially included within the standard map layers were identified, and a portion to each of the standard layers was used as the weighting to determine the standard layer properties.
- If the soil thickness was less than 30 cm or 100 cm, the maximum soil depth would be used for harmonization rather than extending the attributes to 100 cm. Therefore, if the soil thickness was less than 30 cm, the weighted average of soil properties in different layers would be assigned to the topsoil layer, but the subsoil layer (30–100 cm) would be filled with missing value, which is −999.0 in this study. The maximum soil depth derived from the data sources was kept as a separate variable.

4. Quality control:

- All fields were checked for the minimum and the maximum, which were then compared to the value range for each soil property in the source data sets. The cell values falling out of soil property value ranges were treated as errors.
- Missing values resulting from empty entries in the data source were filled with −999.0, except for soil texture fields in the surface organic layer.
- Zeros in sand, silt, clay content in the surface organic layer from STATSGO2 and SLCs are valid values for soil since soil texture is not applicable in this case. The sum of sand, silt and clay fractions in top and subsoil layers was corrected to 100% in the cases where necessary due to rounding errors. Similar to what was done in the development of the HWSD 1.21, when the sum was less than 100 %, the largest percentage was increased to obtain 100 %. When the sum exceeded 100 %, the highest value was reduced to obtain a sum of 100 %.

Improvement of soil organic carbon in high-latitude areas:

The data sets above do not provide sufficient and accurate soil organic carbon information for the high-latitude permafrost region. The Northern Circumpolar Soil Carbon (NCSCD) database links organic carbon measurements from 1,647 pedons in the northern permafrost regions to several digitized regional/national soil maps to produce a combined circumpolar coverage (Hugelius et al., 2013). To provide a more accurate estimate of the soil organic carbon storage in North America, the NCSCD was combined with soil organic carbon content derived from the UNASM (Liu et al., 2013).

6. Data Access:

This data is available through the Oak Ridge National Laboratory (ORNL) Distributed Active Archive Center (DAAC).

Data Archive Center:

Contact for Data Center Access Information:

E-mail: uso@daac.ornl.gov
Telephone: +1 (865) 241-3952

7. References:

Denning, A.S., et al. 2005. Science implementation strategy for the North American Carbon Program: A Report of the NACP Implementation Strategy Group of the U.S. Carbon Cycle Interagency Working Group. U.S. Carbon Cycle Science Program, Washington, DC. 68 pp.

FAO/IIASA/ISRIC/ISS-CAS/JRC: Harmonized World Soil Database (version 1.21), FAO, Rome, Italy and IIASA, Laxenburg, Austria, 2013.

Huntzinger, D.N., C. Schwalm, A.M. Michalak, A.W. King, K. Schaefer, A. R. Jacobson, Y. Wei, S. Liu, R.B. Cook, and W.M. Post. 2013. The North American Carbon Program Multi-scale synthesis and Terrestrial Model Intercomparison Project -- Part I: Overview and experimental design. Geosci. Model Dev., 6, 3977-4008. DOI:10.5194/gmdd-6-3977-2013.

Liu, S., Y. Wei, W. M. Post, R. B. Cook, K. Schaefer, and M. M. Thornton. 2013. The Unified North American Soil Map and its implication on the soil organic carbon stock in North America. Biogeosciences, 10, 2915–2930. DOI:10.5194/bg-10-2915-2013.

Soil Landscapes of Canada Working Group: Soil Landscapes of Canada version 3.2. Agriculture and Agri-Food Canada, (digital map and database at 1:1 million scale), 2010.

Soil Survey Staff, Natural Resources Conservation Service, United States Department of Agriculture, US General Soil Map (STATSGO2), available at: http://www.nrcs.usda.gov/wps/portal/nrcs/detail/soils/survey/?cid=nrcs142p2_053629, last access: August 2011.

Wei, Y., S. Liu, D. Huntzinger, A.M. Michalak, N. Viovy, W.M. Post, C. Schwalm, K. Schaefer, A.R. Jacobson, C. Lu, H. Tian, D.M. Ricciuto, R.B. Cook, J. Mao, and X. Shi. 2013. The North American Carbon Program Multi-scale Synthesis and Terrestrial Model Intercomparison Project – Part 2: Environmental driver data Geosci. Model Dev. Discuss., 6, 5375-5422. DOI:10.5194/gmdd-6-5375-2013

Wei, Y., S. Liu, D. Huntzinger, A.M. Michalak, N. Viovy, W.M. Post, C. Schwalm, K. Schaefer, A.R. Jacobson, C. Lu, H. Tian, D.M. Ricciuto, R.B. Cook, J. Mao, and X. Shi. 2014. NACP MsTMIP: Global and North American Driver Data for Multi-Model Intercomparison. Data set. Available on-line [http://daac.ornl.gov] from Oak Ridge National Laboratory Distributed Active Archive Center, Oak Ridge, Tennessee, USA. http://dx.doi.org/10.3334/ORNLDAAC/1220

OAK RIDGE
National Laboratory

Privacy Policy | Feedback | FAQs | Site Map

Follow @ORNLDAAC

☐ **Home**

About Us
Who We Are
User Working Group
Biogeochemical Dynamics
Data Citation Policy
News
Newsletters
Workshops

Products
Product Overview
Field Campaigns
Validation
Regional/Global
Model Archive

Data
Complete Data Set List
Search for Data
Field Campaigns
Validation
Regional/Global
Model Archive

Tools
Data Search
Site Search
Search by DOI
WebGIS
SDAT
MODIS Land Subsets
THREDDS

Help
FAQs
Tutorials
Data Management

☐ **Contact Us**