Spatial Data for Soil Mapping

By Tom D’Avello, GIS specialist/soil scientist, NRCS, National Soil Survey Center, Geospatial Research Unit, Morgantown, WV; Katey Yoast, Data Analyst, USDA Food and Nutrition Service, Robbinsville, NJ; Dylan Beaudette, Soil Scientist, NRCS, Davis, CA; and Suzann Kienast-Brown, GIS specialist/soil scientist, NRCS, Bozeman, MT

Soil scientists have been using Geographic Information Systems and Image Processing software as tools to assist with soil mapping since the late 1970s. However, widespread use within NRCS did not occur until the early 2000s with the adoption of ArcGIS as the GIS platform for USDA. Since that time, all soil survey offices have become equipped with GIS software and most soil scientists are comfortable using ArcGIS as needed for their job duties. Fortunately, many soil scientists are only aware of pre-GIS methods through the stories of seasoned soil scientists.

The best available tools for the given circumstances have been brought to the task of each generation’s soil mapping challenges. The software and data currently available to soil scientists offer many opportunities to aid the soil mapping process. Much has been written, discussed, researched, and taught regarding digital soil mapping. The best way to think of this without getting bogged down in terminology and methods is “soil mapping using geospatial software and data.” The end product is a soil map produced by soil scientists.

The challenge for the new generation of soil scientists is making use of the large volume of available data. The following table is meant to provide a brief description of data layers (covariates) commonly used for soil mapping activities grouped by soil-forming factor (Jenny, 1941; McBratney et al., 2003). In addition, a listing of software unique to the creation of certain covariates is presented along with links to specific datasets.

Type of Covariate	Covariate	Description	Resolution	Preferred Software	Application
Relief (R)	Elevation	Derived from a digital elevation model (DEM); elevation is represented as a continuous terrain relief surface	Varies depending on source; for most Order 2-3 soil survey applications, 5-10 m resolution is recommended		All areas where elevation is a factor in soil class or properties being modeled
R	Slope Gradient	The maximum rate of change in elevation calculated from a DEM to represent the hydraulic gradient acting upon overland and subsurface water flow through the influence of gravity. Neighborhood size and shape play a large role in resulting slope gradient output, as described in specific Job Aid.	Same as input digital elevation model	ArcSIE SAGA GRASS	All areas where slope gradient is a factor in soil class or properties being modeled
R	Slope Aspect	Direction of slope gradient depicting flow direction. Neighborhood size and shape play a large role in resulting slope gradient output, as described in specific Job Aid. The xTerrain Toolbox contains scripts for calculating northness, eastness, northwestness, and northeastness.	Same as input digital elevation model	ArcMap ArcSIE SAGA xTerrain Tools GRASS	Potentially all areas, except where slope gradient is < 3% and relief is < 150 m
R	Solar Radiation	Estimated incoming solar radiation for a user-specified time period, typically 1 year. May also be considered a proxy for a climatic covariate.	Same as input digital elevation model	ArcMap SAGA GRASS	Potentially all areas, especially where slope gradient is > 3% and relief is > 150 m
R	Heat Load Index	Index to correctly represent southwest aspects being warmer than southeast aspects (McCune and Keon, 2002)	Same as input digital elevation model	Geomor-phometry and Gradient Metrics Toolbox	Potentially all areas, especially where slope gradient is > 3% and relief is > 150 m

R	Profile Curvature	Curvature of slope gradient (direction of the steepest slope) depicting flow acceleration. Neighborhood size and shape play a large role in resulting profile curvature output, as described in specific Job Aid.	Same as input digital elevation model	ArcSIE SAGA GRASS	Potentially all areas
R	Tangential Curvature	Curvature perpendicular to slope gradient depicting flow convergence. Neighborhood size and shape play a large role in resulting tangential curvature output, as described in specific Job Aid.	Same as input digital elevation model	ArcSIE SAGA GRASS	Potentially all areas
R	Convergence Index	Convergence index is used to determine whether water flow from neighboring cells diverges or converges. Convergence is calculated using flow direction between adjacent cells based on the aspects of neighboring cells.	Same as input digital elevation model	SAGA	Potentially all areas
R	Downslope Distance Gradient	Quantifies downslope controls on local drainage, assuming flow accumulation in flat areas downslope topography	Same as input digital elevation model	SAGA	Potentially all areas
R	Morphometric Protection Index	A measure of the openness/protection calculated by analyzing the degree to which the surrounding relief protects the given cell	Same as input digital elevation model	SAGA	Potentially all areas, except where slope gradient is dominantly 0-3%
R	Multiresolution Index of Valley Bottom Flatness	Measure of flatness and lowness depicting depositional areas (Gallant and Dowling, 2003)	Same as input digital elevation model	SAGA	Potentially all areas
R	Multiresolution Index of Ridge Top Flatness	Measure of flatness and upness depicting stable upland areas (Gallant and Dowling, 2003)	Same as input digital elevation model	SAGA	Potentially all areas

R	Topographic Position Index	The deviation of a point elevation from the specified local mean, calculated by dividing the elevation difference by its standard deviation	Same as input digital elevation model	SAGA	Potentially all areas; good for delineating valley bottoms from hilltops and ridges
R	Terrain Surface Texture	A measure of fine (many) versus coarse (few) topographic spacing, calculated as the number of pits and peaks within a specified neighborhood (Iwahashi and Pike, 2007)	Same as input digital elevation model	SAGA	Potentially all areas; good for delineating crests and troughs regardless of amplitude and lithology breaks
R	Valley Depth	The vertical height below summit accumulation	Same as input digital elevation model	SAGA	Potentially all areas, except where slope gradient is dominantly 0-3%
R	Slope Height	The relative height above the closest modeled drainage accumulation	Same as input digital elevation model	SAGA	Potentially all areas
R	Normalized Height	The normalized difference between slope height and valley depth. Also referred to as relative position. Neighborhood size and shape play a large role in this output, as described in specific Job Aid.	Same as input digital elevation model	SAGA xTerrain Tools	Potentially all areas, especially those with a toposequence pattern
R	Vertical Distance to Channel	The vertical height above the channel network. Also known as Altitude Above Channel Network.	Same as input digital elevation model	SAGA	Potentially all areas
R	Overland Flow Distance to Network Channel	A measure of overland flow distances to a channel network based on a DEM and channel network information best depicting potential energy of flow	Same as input digital elevation model	SAGA	Potentially all areas

R	Topographic Wetness Index aka Wetness Index aka Compound Topographic Index	A measure of water accumulation or soil saturation calculated as: *ln(sca/sg)*, where *sca* is the upslope contributing area per unit contour length (or specific catchment area, SCA) and *sg* is the local slope gradient (Moore et.al., 1988)	Same as input digital elevation model	ArcSIE SAGA GRASS	Potentially all areas
R	SAGA Wetness Index	This is similar to the TWI but is calculated as *ln(fa/sg)*, where *fa* is flow accumulation and *sg* is the local slope gradient. In this way, the SAGA TWI does not consider flow as very thin film and therefore predicts cells situated in valley floors as having a small vertical distance to a channel and a more realistic, higher potential soil moisture compared to the standard TWI calculation.	Same as input digital elevation model	SAGA	Potentially all areas
R	Slope Length	A measurement of the distance from the origin of overland flow along its flow path to the location of either concentrated flow or deposition	Same as input digital elevation model	SAGA	Potentially all areas, except where slope gradient is < 3%
R	Stream Power Index	A measure of soil erosion calculated as *ln(sca*sg)*, where *sca* is specific catchment area and *sg* is local slope gradient (Moore et.al., 1988)	Same as input digital elevation model	SAGA	Potentially all areas, except where slope gradient is < 3%
R	Fuzzy Landform Element Classification	Landform classification technique that characterizes terrain patterns according to slope, maximum curvature, minimum curvature, profile curvature, and tangential curvature based on a linear semantic import model for slope and curvature and a fuzzy classification (Schmidt and Hewitt, 2004)	Same as input digital elevation model	SAGA	Potentially all areas

R	Terrain Surface Classification	Landform classification technique that characterizes terrain patterns into 8, 12, or 16 landforms using slope gradient, local convexity, and surface texture (Iwahashi and Pike, 2007)	Same as input digital elevation model	SAGA	Potentially all areas
R	Geomorphons	Scale-invariant landform classification technique that characterizes terrain patterns using a line-of-sight algorithm (Jasiewicz and Stepinski, 2013). There are 10 output landforms: flat, peak, ridge, shoulder, spur, slope, pit, valley, footslope, and hollow.	Same as input digital elevation model	GRASS	Potentially all areas
R	Depression Cost Surface	An index that is calculated as the distance from depression location using slope gradient as a cost surface	Same as input digital elevation model	xTerrain Tools	Terrain with depressional positions
R	Potential Drainage Density	An index that is calculated as the number of flowpaths per unit area as defined by a user-specified neighborhood (Dobos et. al, 2005)	Same as input digital elevation model	xTerrain Tools	Areas with broad valleys and braided or undefined drainage channels
R	Relief	The absolute value of elevation difference for a user-specified neighborhood (Riley et al., 1999)	Same as input digital elevation model	xTerrain Tools	Potentially where relief is a factor; relief and slope gradient usually have > 90% correlation; pick one or the other
R	Relief Ratio	The value of a cell when compared to the average relief for a user-specified neighborhood (Hammond, 1954)	Same as input digital elevation model	xTerrain Tools	Potentially areas where relief is a factor
R	Roughness by Relief and Aspect	The product of the standard deviation of elevation and aspect variety for a user-specified neighborhood	Same as input digital elevation model	xTerrain Tools	Areas where relief is a factor and degree of dissection or surface smoothness is variable

R	Roughness by Standard Deviation of Relief	Calculates the standard deviation of elevation for a user-specified neighborhood. High standard deviation corresponds to rougher terrain.	Same as input digital elevation model	xTerrain Tools	Potentially areas where relief is a factor
R	Slope Heterogeneity	Calculates the standard deviation of slope for a user-specified neighborhood (Evans, 1998)	Same as input digital elevation model	xTerrain Tools	Potentially areas where slope gradient is > 8%
Climate (C)	Precipitation	Typically mean annual, but seasonal time slices are also used	Varies, but 4km and 800m are common	PRISM	Potentially all areas if project domain is of a large spatial extent and if orographic effects warrant
C	Air Temperature	Typically mean annual, but seasonal time slices are also used	Varies, but 4km and 800m are common	PRISM	Potentially all areas if project domain is of a large spatial extent and if orographic effects warrant
C	Soil Temperature	Typically mean annual	Varies	Newhall model or modeled from sensor data	Potentially all areas if project domain is of a large spatial extent and if orographic effects warrant
C	Evapo-transpiration	Typically potential evapotranspiration	Varies, but 800m is common	Often derived from PRISM data or Newhall model	Potentially all areas if project domain is of a large spatial extent and if orographic effects warrant
C	Water Balance aka Effective Precipitation	The difference between precipitation and evapotranspiration for a given time period	Varies, but 800m is common	Often derived from PRISM data or Newhall model	Potentially all areas if project domain is of a large spatial extent and if orographic effects warrant
C	Solar Radiation	Estimates potential or actual incoming solar radiation for specified time interval (Hofierka & Suri, 2002)	Same as input digital elevation model	SAGA ArcGIS GRASS	Areas where climate is aspect driven

Organisms (O)	Landsat GeoCover 2000	Orthorectified Landsat band combinations of 7-4-2 and 4-3-2	15m		Potentially all areas where land cover data is needed
O	Blue spectral band	Landsat 5, 7—band 1; Landsat 8—band 2	30m		Bathymetric mapping, distinguishing soil from vegetation, and detecting cultural features
O	Green spectral band	Landsat 5, 7—band 2; Landsat 8—band 3	30m		Peak vegetation (plant vigor), separating vegetation from soil, and separating urban from non-urban areas
O	Red spectral band	Landsat 5, 7—band 3; Landsat 8—band 3	30m		Vegetation slopes
O	Near infrared spectral band	Landsat 5, 7—band 4; Landsat 8—band 4	30m		Biomass content, shorelines, distinguishing dry and moist soils
O	Short wave infrared spectral band (a) (SWIR)	Landsat 5, 7—band 5; Landsat 8—band 6	30m		Soil and vegetation moisture; land use
O	Thermal infrared spectral band	Landsat 5, 7—band 6; Landsat 8—band 10, 11	60m		Soil moisture
O	Normalized Difference Vegetation Index (NDVI)	$\frac{NIR - Red}{NIR + Red}$ Red and Near Infrared bands; Landsat 5, 7—bands 3, 4; Landsat 8—bands 4, 5	Resolution of imagery	ERDAS ArcGIS SAGA	Vegetative cover
O	Landfire (multiple datasets)	Comprehensive geospatial data and databases put together by multiple Federal agencies that describe vegetation, wildland fuel, and fire regimes across the U.S. and insular areas.	Varies		ESD and soil survey
O	Land Cover (NLCD)	National Land Cover Database	30m		

O	Canopy Density	Density of vegetative canopy	Varies	Derived from LiDAR point cloud; dependent on sufficient point density	Areas with relatively undisturbed vegetation where canopy density differences infer soil differences
O	Canopy Height	Height of vegetative canopy	Varies	Derived from LiDAR point cloud; dependent on sufficient point density	Areas with relatively undisturbed vegetation where canopy height differences infer soil differences
Parent Material (P)	Short wave infrared band (a) (SWIR)	Landsat 5, 7—band 5; Landsat 8—band 6	30m		Soil and vegetation moisture; land use
P	Short wave infrared band (b) (SWIR)	Landsat 5, 7—band 7; Landsat 8—band 7	30m		Hydrothermally altered rocks associated with mineral deposits; separating land from water
P	Carbonate Difference Ratio	Red - Green Red + Green Red and Green bands; Landsat 5, 7—bands 3, 2; Landsat 8—bands 4, 3	Resolution of imagery	ERDAS Model ArcGIS SAGA	Differentiate carbonate-rich areas
P	Clay Difference Ratio	SWIR(a) - SWIR(b) SWIR(a) + SWIR(b) SWIR bands; Landsat 5, 7—bands 5, 7; Landsat 8—bands 6, 7	Resolution of imagery	ERDAS Model ArcGIS SAGA	Differentiate areas of high clay hydroxyl influence
P	Ferrous Minerals Difference Ratio	SWIR(a) - NIR SWIR(a) + NIR SWIR bands; Landsat 5, 7—bands 5, 4; Landsat 8—bands 6, 5	Resolution of imagery	ERDAS Model ArcGIS SAGA	Differentiate areas of higher ferrous mineral influence
P	Iron Difference Ratio	Red - SWIR(a) Red + SWIR(b) Red and SWIR bands; Landsat 5, 7—bands 3, 7; Landsat 8—bands 4, 7	Resolution of imagery	ERDAS Model ArcGIS SAGA	Differentiate areas of higher iron mineral influence

P	Rock Outcrop Difference Ratio	SWIR(a) - Green SWIR(b) + Green SWIR and Green; Landsat 5, 7—bands 5, 2; Landsat 8—bands 6, 3	Resolution of imagery	ERDAS Model ArcGIS SAGA	Differentiate sedimentary rock (lime/dolostone) from igneous rock
P	Geology		1:250,000		Potentially all areas where parent material is pertinent
P	Aeroradiometric Data	Data generated by aerial sensing of radiation emanating from the earth's surface provides general estimates of the geographic distribution of uranium, thorium, and potassium in surficial and bedrock units.			Radiometric data may be helpful in differentiating various parent materials (Cook et al., 1996)
Soils (S)	Soil Map	Soil polygons or regions from SSURGO or gSSURGO	1:12,000-1:24,000	SSURGO, gSSURGO	Potentially all areas where data are available
S	Soil Pedon	Soil description with field observations and/or physical, chemical, and mineralogical characterization		KSSL and NASIS	Potentially all areas where data are available
S	Proximal Sensing	Data in x,y,z format, such as EM, GPR, Gamma, or Bathymetric surveys	Various spacing	User developed	Potentially all areas where data are available

Literature

Cook, S.E., R.J. Corner, P.R. Groves, and G.J. Grealish. 1996. Use of airborne gamma radiometric data for soil mapping. Soil Research 34(1):183-194.

Dobos, E., J. Daroussin, and L. Montanarella. 2005. An SRTM-based procedure to delineate SOTER Terrain Units on 1:1 and 1:5 million scales. EUR21571 EN, Office for Official Publications of the European Communities, Luxembourg.

Evans, I.S. 1998. What do terrain statistics really mean? *In* S.K. Land et al. (eds.) Landform Monitoring, Modeling and Analysis, John Wiley and Sons, Ltd., Chichester, UK, pp. 119-138.

Gallant, J.C., and T.I. Dowling. 2003. A multiresolution index of valley bottom flatness for mapping depositional areas. Water Resources Research 39(12).

Hammond, E.H. 1954. Small scale continental landform maps. Annals of the Association of American Geographers 44:32-42.

Hofierka, J., & Suri, M. (2002, September). The solar radiation model for Open source GIS: implementation and applications. In Proceedings of the Open source GIS-GRASS user’s conference. 1-19.

Iwahashi, J., and R.J. Pike. 2007. Automated classifications of topography from DEMs by an unsupervised nested-means algorithm and a three-part geometric signature. Geomorphology 86:409–440.

Jasiewicz, J., and T.F. Stepinski. 2013. Geomorphons—A pattern recognition approach to classification and mapping of landforms. Geomorphology 182:147-156.

Jenny, H. 1941. Factors of soil formation: A system of quantitative pedology.

McBratney, A.B., M.M. Santos, and B. Minasny. 2003. On digital soil mapping. Geoderma 117(1):3-52.

McCune, B. and Keon, D. 2002. Equations for potential annual direct incident radiation and heat load. Journal of vegetation science, 13(4):603-606.

Moore, I.D., G.J. Burch, and D.H. Mackenzie. 1988. Topographic effects on the distribution of surface soil water and the location of ephemeral gullies. Transactions of the ASAE 31(4):1098-1107.

Riley, S.J., S.D. DeGloria, and R. Elliot. 1999. A terrain ruggedness index that quantifies topographic heterogeneity. Intermountain Journal of Sciences 5:23-27.

Schmidt, J., and A. Hewitt. 2004. Fuzzy land element classification from DTMs based on geometry and terrain position. Geoderma 121(3):243-256.

Soil Science Division Staff. 2017. Soil survey manual. C. Ditzler, K. Scheffe, and H.C. Monger (eds.). USDA Handbook 18. Government Printing Office, Washington, D.C.