QUICK GUIDE TO GENERATION OF SOIL TOPOGRAPHIC INDEX AND MONTHLY SATURATION PROBABILITY MAPS IN ARCGIS 9.1

A Product of: "Developing GIS Water Quality Tools for Tompkins County"
Prepared by: Shannon Seifert, M.P.S. Candidate, Biological and Environmental Engineering

REPORT CONTENTS

I. Introduction
   A. Definitions
   B. Data
   C. ArcGIS Raster Calculations

II. Generating STI in ArcGIS

III. Calculating monthly saturation probability (P_{sat}) in ArcGIS

IV: Querying P_{sat} Layers in ArcGIS

V. Notes on primary data sources
   A. DEM
   B. Soils
   C. Other
   D. Data provided on CD-ROM

VI. Contacts

VII. Appendices
   A. Scripts
   B. Metadata
I. INTRODUCTION

A. Definitions

Topographic indices are used to map hydrologically sensitive areas (HSA), and are particularly applicable to hydrologic systems driven by shallow subsurface flow, such as in the Tompkins County area. This report distinguishes between the topographic index (TI) which is determined purely by topography (a topographic layer), and the soil-topographic index (STI), which is a modified TI that includes a soils component (Soil depth and saturated permeability data). The STI corresponds directly to “λ” in Agnew et al. (2006), from which saturation probability regressions in this project for Tompkins County originate.

The equations governing TI and STI generation are listed below.

\[ TI = \ln \left( \frac{a}{\tan \beta} \right) \]  
\[ \text{Where} \]  
\[ a = \text{upslope contributing area per unit contour length (m)}; \]  
\[ \tan \beta = \text{surface slope (rise/run) as a fraction (-)} \]  

\[ STI = \ln \left( \frac{a}{\tan \beta \times Ks \times D} \right) \]  
\[ \text{Where:} \]  
\[ Ks = \text{mean saturated conductivity of the soil (m/d)} \]  
\[ D = \text{soil depth (m)} \]  

These equations must be adapted to GIS applications; as discussed in detail below.

B. Data

From our definitions it is clear that the primary maps needed are a Digital Elevation Model (DEM) layer and a soils data layer with accompanying soil properties (saturated conductivity and depth). Table 1 lists the raw data and sources used in this project.

<table>
<thead>
<tr>
<th>Data</th>
<th>Filename</th>
<th>Source</th>
<th>Extent</th>
</tr>
</thead>
<tbody>
<tr>
<td>DEM</td>
<td>tompkins_10m</td>
<td><a href="http://www.iagt.org">www.iagt.org</a></td>
<td>Cayuga Lake Watersheds</td>
</tr>
<tr>
<td>Soils</td>
<td>SoilsTC</td>
<td>CUGIR</td>
<td>Tompkins County</td>
</tr>
</tbody>
</table>
C. ArcGIS Raster Calculations

From Equations 1 and 2 (see Definitions):

1. The raster value ‘a’ is the same as the ‘flow accumulation’, multiplied by the grid cell area (100 m²) and divided by the grid cell size (10 m) of the DEM.
   
   \[ a = \frac{\text{flow accumulation} \times (100 \text{m}^2)}{10 \text{m}} = \frac{\text{flow accumulation}}{10} \times 100 \text{m} \]

2. The raster 'slope' is calculated as percent slope (an output option of the ArcGIS “Slope” tool).

3. Regarding the soils dataset, ‘Ks’ is the permeability of a soil type and ‘D’ is the total depth of soil. These values are provided in the Tompkins County ‘Soils_TC’ dataset; the original data consisted of a range of permeability values (in inches per hour) and total depth in inches for each soil type. To generate the STI, permeability was calculated using the geometric mean of values for ‘Ks’ after converting units to meters/day; and a simple unit conversion was applied to ‘D’ to produce total depth in meters. This allowed for consistent raster units (meters).

As a result, generating a TI and STI requires a series of steps that are outlined below. Equations 1 and 2 are reconstructed in the following way for GIS calculations, and are mathematically consistent. The steps for generating both the TI and STI using the ArcGIS 9+ Model Builder extension are outlined in the flowchart that follows.

\[
TI = \ln \left( \frac{a}{\tan \beta} \right)
\]

Where

\[ a = \text{flow accumulation} \times \text{cell size} (a = A \times 10) \]

\[ \tan \beta = \frac{\text{Slope}\%}{100} \]

\[
STI = TI - \ln(\text{KsD})
\]

Where

“ln(KsD)” is an existing raster dataset created by converting the ‘SoilsTC’ vector dataset into a raster dataset (filename: lnksd)

II. GENERATING STI USING ARCGIS 9+ (MODEL BUILDER)

The Model Builder feature in ArcGIS allows a user without programming experience to construct a model process through a GUI by dragging and dropping existing ArcToolbox tools into a schematic window. The STI generating model is included on the CD ROM with this report and you are invited to import and edit the model if interested. In addition to the ArcTool, exported JavaScript, Python, and VBasic scripts of the STI model are included on the CD ROM. These will work with little to no editing as long as they are stored and run from within the same directory as the input datasets (input dataset names may have to be changed in the script if you plan to use datasets not included in this report). The schematic view of the STI model process is included in this report, below.
III. Calculating Monthly Saturation Probability

The monthly saturation probability layers are directly based on the probability vs. STI regressions in Agnew et al. (2006). These regressions are reproduced below in Table 2. Note that ‘STI’ in Table 2 is the same as “λ” in Agnew et al. (2006).

Table 2: Monthly soil saturation probability regression values

<table>
<thead>
<tr>
<th>Month</th>
<th>STI&lt;sub&gt;0&lt;/sub&gt;</th>
<th>STI&lt;sub&gt;100&lt;/sub&gt;</th>
<th>Psat(STI) For STI&lt;sub&gt;0&lt;/sub&gt; &lt; STI &lt; STI&lt;sub&gt;100&lt;/sub&gt;</th>
<th>R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jan</td>
<td>8.0</td>
<td>15.2</td>
<td>Psat = 13.9(STI) - 111.8</td>
<td>0.87</td>
</tr>
<tr>
<td>Feb</td>
<td>7.9</td>
<td>15.5</td>
<td>Psat = 13.1(STI) - 103.3</td>
<td>0.88</td>
</tr>
<tr>
<td>Mar</td>
<td>8.0</td>
<td>15.3</td>
<td>Psat = 13.6(STI) - 108.5</td>
<td>0.86</td>
</tr>
<tr>
<td>Apr</td>
<td>7.6</td>
<td>15.5</td>
<td>Psat = 12.7(STI) - 96.3</td>
<td>0.91</td>
</tr>
<tr>
<td>May</td>
<td>8.5</td>
<td>16.2</td>
<td>Psat = 13.0(STI) - 110.5</td>
<td>0.95</td>
</tr>
<tr>
<td>June</td>
<td>9.7</td>
<td>17.0</td>
<td>Psat = 13.7(STI) - 133.3</td>
<td>0.94</td>
</tr>
<tr>
<td>July</td>
<td>10.3</td>
<td>17.8</td>
<td>Psat = 13.4(STI) - 137.9</td>
<td>0.91</td>
</tr>
<tr>
<td>Aug</td>
<td>10.3</td>
<td>17.7</td>
<td>Psat = 13.4(STI) - 137.5</td>
<td>0.93</td>
</tr>
<tr>
<td>Sept</td>
<td>9.9</td>
<td>17.2</td>
<td>Psat = 13.8(STI) - 137.2</td>
<td>0.94</td>
</tr>
<tr>
<td>Oct</td>
<td>8.4</td>
<td>16.7</td>
<td>Psat = 12.0(STI) - 100.2</td>
<td>0.95</td>
</tr>
<tr>
<td>Nov</td>
<td>7.9</td>
<td>15.4</td>
<td>Psat = 13.3(STI) - 105.0</td>
<td>0.92</td>
</tr>
<tr>
<td>Dec</td>
<td>7.7</td>
<td>15.0</td>
<td>Psat = 13.7(STI) - 105.6</td>
<td>0.89</td>
</tr>
</tbody>
</table>

<sup>a</sup> Psat = 0% for STI ≤ STI<sub>0</sub>
<sup>b</sup> Psat = 100% for STI ≥ STI<sub>100</sub>

**Example: Generating saturation probability map for April**

A number of conditional statements can be used to generate this layer. In this example, we use the ‘con’ statement in ArcGIS with the Map Algebra tool. After generating the STI raster (as described above), use the conditional ‘CON’ statement as follows, using the published the regression limits and equation in Table 2 (above) for the month of April:

CON (STI <= 7.6, 0, CON (STI >= 15.5, 100, (12.7 * STI - 96.3)))

The output raster should have values of 0 to 100, representing the percent chance of saturation in a given pixel.
IV. QUERYING P_{SAT} LAYERS IN ARCGIS

Raster layers can be queried using several different methods. For the purpose of this report, we will focus on a method that will select pixels from a raster within a range of values and then export those pixels to a polygon layer. The polygon layer can then be used to calculate total area or provide a map layer of specific $P_{sat}$ range for a given month.

To Query $Psat$ raster values:
1. Open all Psat tables in ArcMap
2. Follow ArcDocs steps for “Raster to Feature Conversion”
   a. Convert raster surface to categorical data: Use Spatial Analyst $\rightarrow$ Reclass $\rightarrow$ Reclassify and choose your input raster and reclass field (value) and classigy according to your desired percent ranges.
      i. You may want to change the raster data from floating-point to integer values using the “Int” Math tool before reclassifying. This step removes the decimal part of your raster cell value.
   b. Convert categories to polygons: Use “Conversion Tools” $\rightarrow$ “From Raster” $\rightarrow$ “Raster to Polygon” to generate a simple polygon shapefile. Check the simplify polygons box: the output shapefile will then represent each grid cell with a small square polygon. You now have a shapefile you can manipulate.
3. Use the Selection $\rightarrow$ Select by Attributes tool to select and export only the polygons you desire to work with.

V. NOTES ON PRIMARY DATASETS

A. Digital Elevation Model (DEM) – Raster Data filename: “fill_dem_10m”
1. Ideally, the DEM layer should cover all of the watersheds that intersect or fully encompass Tompkins County proper. This is so that flow accumulation along county boundaries is not underestimated. Such a DEM was available from the Institute for the Application of Geospatial Technology, or IAGT, at Cayuga community College (http://www.iagt.org). It is important that the DEM raster used for this analysis is in ‘floating-point’ and not ‘integer’ units.
2. Two intermediate steps were completed to correct DEM errors prior to generating the STI.
   a. A thin, linear artifact of “No Data” cells bisecting the DEM was removed using the ‘focalmean’ function in Map Algebra with parameters: “rectangle, 3, 3, data.” This removed the artifact and filled the No Data pixels with local mean values from adjacent cells. The artifact was probably a result of joining two DEMs with less than 3 pixel tolerance at edges.
   b. The Hydro tool “Fill” was used to correct small errors in the DEM by filling flow sinks. The output DEM used for generating STI is “fill_dem_10m.”

B. Soils
   Vector data filename: “SoilsTC”
   Raster data for ln($K_s D$) filename: “lnksd”
1. Missing permeabilities and depths: For some map units (i.e., bedrock), no permeabilities or depths are listed in the vector data layer. In such a case, the arbitrary value of 0.001 was assigned as $ln(K_s D)$ instead of zero before converting the soils vector layer to a raster layer. This is done to avoid mathematical errors in later raster calculations.
2. *Calculating geometric mean permeabilities:* In cases where either minimum or maximum permeability values were missing from the vector layer, the sole listed permeability was taken as the mean and used for the $ln(K_{sD})$ value.

C. Other

*Mathematic Errata:* The logarithm of zero, or 1/0, is not determinable. For this reason constants of 0.001 have been added individually to the flow accumulation and slope percent data, as shown in the flow chart in Figure 1.

A more detailed description of the process for generating $lnK_{sD}$ values from the Tompkins Co. Soil Survey data (SCS, 1965) is included in the MS Word file: “Treatment of Tompkins Co Soils Data for STI Generation” included on the DVD.

D. Data provided on DVD-ROM

<table>
<thead>
<tr>
<th>Data Files</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Original DEM</td>
<td>tompkins_10m</td>
</tr>
<tr>
<td>Corrected DEM</td>
<td>fill_dem_10m</td>
</tr>
<tr>
<td>Soils vector</td>
<td>SoilsTC</td>
</tr>
<tr>
<td>Soils raster</td>
<td>lnksd</td>
</tr>
<tr>
<td>Topographic Index raster</td>
<td>ti</td>
</tr>
<tr>
<td>Soil Topographic Index raster</td>
<td>sti</td>
</tr>
<tr>
<td>January Psat raster</td>
<td>psat_jan</td>
</tr>
<tr>
<td>February Psat raster</td>
<td>psat_feb</td>
</tr>
<tr>
<td>March Psat raster</td>
<td>psat_mar</td>
</tr>
<tr>
<td>April Psat raster</td>
<td>psat_apr</td>
</tr>
<tr>
<td>May Psat raster</td>
<td>psat_may</td>
</tr>
<tr>
<td>June Psat raster</td>
<td>psat_jun</td>
</tr>
<tr>
<td>July Psat raster</td>
<td>psat_jul</td>
</tr>
<tr>
<td>August Psat raster</td>
<td>psat_aug</td>
</tr>
<tr>
<td>September Psat raster</td>
<td>psat_sep</td>
</tr>
<tr>
<td>October Psat raster</td>
<td>psat_oct</td>
</tr>
<tr>
<td>November Psat raster</td>
<td>psat_nov</td>
</tr>
<tr>
<td>December Psat raster</td>
<td>psat_dec</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Model Files</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Schematic</td>
<td>sti.bmp</td>
</tr>
<tr>
<td>STI Toolbox</td>
<td>STI</td>
</tr>
<tr>
<td>VBasic script</td>
<td>sti.vbs</td>
</tr>
<tr>
<td>Java script</td>
<td>sti.js</td>
</tr>
<tr>
<td>Python Script</td>
<td>sti.py</td>
</tr>
</tbody>
</table>
VI. CONTACTS

Kate Hackett  
Senior Planner  
Tompkins County Planning Department  
121 East Court Street  
Ithaca, NY 14850  
Email: khackett@tompkins-co.org  
Telephone: (607) 274-5560  
Fax: (607) 274-5578

Crystal Buck  
Senior Planner  
Tompkins County Planning Department  
121 East Court Street  
Ithaca, NY 14850  
Email: cbuck@tompkins-co.org

David Carr  
Inst. for Applied Geospatial Technology  
Email: dcarr@iagt.org

Nicholas Hollingshead  
Cayuga Lake Watershed Network  
Email: nicholashollingshead@yahoo.com

Nathan W. Krause  
Geospatial Info. Technology Specialist  
Institute for the Application of Geospatial Technology  
199 Franklin St, Suite 300  
Auburn, NY 13021-3025  
(315)283-9454  
nkrause@iagt.org  
www.iagt.org

Arthur J. Lembo, Jr.; Ph.D.  
Department of Crop and Soil Sciences  
Cornell University  
1001 Bradfield Hall  
Ithaca, New York 14853  
Email: ajl53@cornell.edu  
(607) 255-6328

Vishal K. Mehta  
Ph.D. Candidate  
Dept. of Crop and Soil Science  
1007 Bradfield Hall, Cornell University  
Email: vkm2@cornell.edu  
ph: 607 255 1706 (campus)

Shannon Seifert  
M.P.S. Candidate  
Department of Biological & Environmental Engineering  
Cornell University  
Ithaca, NY 14853-5701  
Email: shannon.seifert@gmail.com  
(607) 533-4298

Dr. M.Todd Walter  
Department of Biological & Environmental Engineering  
Cornell University  
Ithaca, NY 14853-5701  
Email: mtw5@cornell.edu  
(607)255-2488 office  
(607)255-2463 lab

Reference

VII. Appendix

A. Python Script Example:

```python
# STI_final.py
# Created on: Thu Nov 30 2006 04:43:39 PM
# (generated by ArcGIS/ModelBuilder)
# Usage: STI_final <Soil_Topographic_Index> <Topographic_Index> <fill_dem_10m> <lnksd>
# ---------------------------------------------

# Import system modules
import sys, string, os, win32com.client

# Create the Geoprocessor object
gp = win32com.client.Dispatch("esriGeoprocessing.GpDispatch.1")

# Check out any necessary licenses
gp.CheckOutExtension("spatial")

# Load required toolboxes...
gp.AddToolbox("H:\Program Files/ArcGIS/ArcToolbox/Toolboxes/Spatial Analyst Tools.tbx")

# Script arguments...
Soil_Topographic_Index = sys.argv[1]
if Soil_Topographic_Index == '#':
    Soil_Topographic_Index = "H:\Documents and Settings\mtw5\Local Settings\Temp\STI" # provide a default value if unspecified

Topographic_Index = sys.argv[2]
if Topographic_Index == '#':
    Topographic_Index = "H:\Documents and Settings\mtw5\Local Settings\Temp\TI" # provide a default value if unspecified

fill_dem_10m = sys.argv[3]
if fill_dem_10m == '#':
    fill_dem_10m = "H:\GIS\Shannon\TompkinsCo\TompkinsCo_GIS\TompkinsCo\DEM\fill_dem_10m" # provide a default value if unspecified

lnksd = sys.argv[4]
if lnksd == '#':
    lnksd = "H:\GIS\Shannon\TompkinsCo\TompkinsCo_GIS\TompkinsCo\lnksd" # provide a default value if unspecified

# Local variables...
Watershed_Slope = "H:\Documents and Settings\mtw5\Local Settings\Temp\wsslope"
a = "H:\Documents and Settings\mtw5\Local Settings\Temp\Plus_flowacc1"
Slope__ = "H:\Documents and Settings\mtw5\Local Settings\Temp\Plus_2"
slope_per_unit_contour_width = "H:\Documents and Settings\mtw5\Local Settings\Temp\Divide_Plus_1"
v100 = "100"
a_10 = "H:\Documents and Settings\mtw5\Local Settings\Temp\Times_Plus_f1"
v10 = "10"
v_a_10___tan_B_ = "H:\Documents and Settings\mtw5\Local Settings\Temp\Divide_Times1"
constant_0_0001 = "0.0001"
constant_0_0001__2_ = "0.0001"
flowdir = "H:\Documents and Settings\mtw5\Local Settings\Temp\flowdir"
Output_drop_raster = ""
Flow_Accumulation_Layer = "H:\Documents and Settings\mtw5\Local Settings\Temp\flowacc"

# Process: Flow Direction...
gp.FlowDirection_sa(fill_dem_10m, flowdir, "FORCE", Output_drop_raster)

# Process: Flow Accumulation...
```
gp.FlowAccumulation_sa(flowdir, Flow_Accumulation_Layer, "")

# Process: Plus...
gp.Plus_sa(Flow_Accumulation_Layer, constant_0.0001, a)

# Process: Times...
gp.Times_sa(a, v10, a_10)

# Process: Slope...
gp.Slope_sa(fill_dem_10m, Watershed_Slope, "PERCENT_RISE", "1")

# Process: Plus (2)...
gp.Plus_sa(constant_0.0001_2, Watershed_Slope, Slope__)

# Process: Divide (2)...
gp.Divide_sa(Slope__, v100, slope_per_unit_contour_width)

# Process: Divide...
gp.Divide_sa(a_10, slope_per_unit_contour_width, v_a_10__tan_B__)

# Process: Ln...
gp.Ln_sa(v_a_10__tan_B__, Topographic_Index)

# Process: Minus...
tempEnvironment0 = gp.extent
gp.extent = "DEFAULT"
gp.Minus_sa(Topographic_Index, lnksd, Soil_Topographic_Index)
gp.extent = tempEnvironment0