Hydrological functions and applications of GIS

Techniques for spatial data analysis
Last week

- We looked at slope, and its components \textit{gradient} and \textit{aspect}.
- \textbf{Slope} was the tangent plane to a surface – we assume continuity when calculating it...
- We briefly defined \textit{profile} and \textit{plan curvature}.
- Slope is an \textbf{important factor} in many processes, from \textit{house prices} to \textit{food sources} for red deer.
- We looked at a range of algorithms for calculating \textit{gradient} and \textit{aspect} on \textit{rasters} and \textit{tins}.
- Finally, we compared some applications of these parameters...
Outline

• Briefly look at why **hydrology** is so **important**
• Examine algorithms for calculating hydrological derivatives from terrain (so called **primary indices**)
• Look at **compound indices** derived from these primary indices and consider their physical meaning
• Examine a paper investigating **resolution effects** on **channel initiation**
Learning objectives

You will:

- **know** what is meant by the **D8 algorithm** and other **flow direction algorithms** (Rho8, FD8), and can **discuss** their **strengths** and **weaknesses**;

- **be able to describe** some standard GIS operators used in **hydrology** and **relate** them to **physical features** found in the **landscape**; and

- **know** what is meant by a **compound topographic index** and be able to **list examples** of these and **comment** on their use.
Why is hydrology so important?

- Topography shapes and is shaped by hydrological processes.
- Catchments are a key unit in understanding natural processes and anthropogenic influences on the landscape.
- We will look at a few applications of hydrological properties in the 2nd part of the lecture...
- Topography is fundamental to the flow characteristics within a catchment: e.g.: 
Influence of basin characteristics on hydrograph (from http://www.geog.ox.ac.uk/students/teaching/pg1_ith/)

Distributed modelling
Some examples of predictions based on hydrological properties...

- **Sediment** and **nutrient transport** in landscape
- Pollution and **transport of pollutants**
- **Variability of biological productivity** as a function of catchment
- **Flood forecasting** e.g. through snow melt, thunder storms etc...
- Prediction of **stream formation**
- Prediction of **erosion** (e.g. USLE)
Overland flow

• When no more water can be absorbed by soils then **overland flow** occurs.

• At the tops of hillslopes overland flow may be ‘**sheet flow**’, water flows as a thin sheet over the land surface.

• Sheet flow does not occur over very long distances – flow is concentrated into rills – the initiators of **channelled flow**...

• Channelled flow is **convergent**, sheet flow may be **divergent**...

What curvature property do you think would describe divergent flow?
Example topography – Glen Feshie

From: http://www.gla.ac.uk/medicalgenetics/feshie.htm#Glen_Feshie
Deriving flow direction

- **Flow direction** defines the route that water will take from one cell to others.
- **Deriving flow directions** from a DEM is a fundamental step in hydrological modelling.
- Many algorithms have been described in the literature, but one of the most basic remains the most commonly used (D8).
- What **types of relief** do you think would be challenging to derive flow direction in?
D8

- Assume direction of flow is direction of steepest drop to one of 8 neighbours (all flow goes to one cell)
- The flow direction here is the same as the aspect calculated with the steepest drop last week (but represented differently)

\[
\text{FlowDirection} = 2^{j-1} \quad \text{where} \quad j = i \quad \text{for} \quad \max_{i=1,8} \left\{ \phi(i) \frac{z_9 - z_i}{\lambda} \right\}
\]

<table>
<thead>
<tr>
<th>Cell addressing</th>
<th>Flow directions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Z₁ Z₂ Z₃ Z₄</td>
<td>32 64 128</td>
</tr>
<tr>
<td>Z₅ Z₆ Z₇ Z₈</td>
<td>16 1</td>
</tr>
<tr>
<td></td>
<td>8   4</td>
</tr>
</tbody>
</table>

\(\phi(i) = 1\) : for NSEW neighbours
\(= 1/\sqrt{2}\) : for diagonals

\(\lambda\) is cell spacing

Remember, \(2^n\), where \(0 \leq n \leq 7\), will give values of between 1 and 128 \(\Rightarrow\) Why might these values be better than 1,2,3,4,...?
More D8

- Flow directions can only have intervals of 45°
- Where steepest drop occurs in multiple directions, different strategies can be used:
  - Assign (all) these directions to cell (by summing directions) (ArcGIS – treated later as a sink – you saw this in the lab)
  - Assign 1st direction to cell
  - Flag direction as undefined
- This method is identical to that for calculating aspect with steepest drop
### D8 example...

#### Heights

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>4</th>
<th>6</th>
<th>12</th>
<th>20</th>
<th>44</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>4</td>
<td>8</td>
<td>11</td>
<td>13</td>
<td>24</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>8</td>
<td>15</td>
<td>20</td>
<td>25</td>
<td>39</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>14</td>
<td>22</td>
<td>32</td>
<td>37</td>
<td>47</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>18</td>
<td>31</td>
<td>39</td>
<td>44</td>
<td>52</td>
<td></td>
</tr>
</tbody>
</table>

#### D8 results

<table>
<thead>
<tr>
<th>xx</th>
<th>16</th>
<th>16</th>
<th>16</th>
<th>16</th>
<th>16</th>
</tr>
</thead>
<tbody>
<tr>
<td>64</td>
<td>32</td>
<td>16</td>
<td>32</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>64</td>
<td>64</td>
<td>32</td>
<td>64</td>
<td>64</td>
<td></td>
</tr>
<tr>
<td>64</td>
<td>32</td>
<td>32</td>
<td>32</td>
<td>32</td>
<td></td>
</tr>
<tr>
<td>64</td>
<td>32</td>
<td>16</td>
<td>32</td>
<td>32</td>
<td></td>
</tr>
</tbody>
</table>

#### Direction key

<table>
<thead>
<tr>
<th>32</th>
<th>64</th>
<th>128</th>
</tr>
</thead>
<tbody>
<tr>
<td>16</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>4</td>
<td>2</td>
</tr>
</tbody>
</table>

Note solution on boundaries (where 9 cells are not available)
D8 results on big grid...

59 unique values?!
Issues with D8

• Flow can only be modelled in 45° intervals – therefore rarely agrees with aspect (e.g. where aspects lie between 45° and 60° all flow lines will be modelled as 45°)
• Cannot model flow dispersion
• Flow lines tend to be parallel in areas of similar aspect (that is on planar slopes)
Parallel streams on planar slopes (at 45° increments)

See how streams are calculated later...
Rho8

- Rho8 algorithm attempts to deal with problems of parallel flow by changing the constant in the equation

\[
\text{FlowDirection} = 2^{j-1} \quad \text{where } j = i \text{ for } \max_{i=1,8} \left\{ \phi(i) \frac{z_i - z_j}{\lambda} \right\}
\]

\[
\phi(i) = 1 \quad \text{: for NSEW neighbours}
\]
\[
= \frac{1}{2-r} \quad \text{: for diagonals where}
\]

- Has the effect of ‘nudging’ flow lines out of parallel flow
- Still cannot model flow dispersion, but networks are more realistic
- Answers are potentially different every time
Modelling dispersion – FD8

- In upland areas where channels have not initiated, flow may be better modelled as **divergent** (areas of **sheet flow**)

- Moore proposed an **algorithm** which represents the fraction of the flow in a cell passed to a neighbour $i$ as

$$F_i = \frac{\max(0, \text{slope}_i^y)}{\sum_{j=1}^{8} \left[ \max(0, \text{slope}_j^y) \right]}$$

- $\nu$ is a constant (set to 1.1)

- Note we say slope here, not gradient – we give a **sign** to the gradient (up or down) and only downhill slope (+ve) is passed
FD8 example

<table>
<thead>
<tr>
<th>Heights</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
</tr>
<tr>
<td>3</td>
</tr>
<tr>
<td>5</td>
</tr>
<tr>
<td>9</td>
</tr>
<tr>
<td>12</td>
</tr>
<tr>
<td>14</td>
</tr>
</tbody>
</table>

- For the given DEM
  - Work out **which cells** will **receive flow** from their neighbours
  - Estimate, using **weighted arrows**, **how much** flow will be received
Modelling channel networks

• Having derived flow directions, we can derive channel networks

• Many different approaches exist to modelling channel networks
  – Those based around focal operators e.g. local maxima/ minima, convexities and concavities
  – “Hydrologically-based” algorithms using flow paths (calculated from flow directions)
  – Other approaches are derived from areas such as image processing (see Remote Sensing)

• We will look mainly at the second approach (briefly at the first)
Focal operators

• Early approach from Peucker and Douglas (1975)
• To find potential points on channels
  – **Flag highest point** in 2x2 window for every cell
  – Unflagged cells are potential channel points
• **Simple** and **fast** to implement
• Does not give a **complete network** (only discontinuous points)
• Very susceptible to **noise** in the DEM
<table>
<thead>
<tr>
<th>1</th>
<th>4</th>
<th>6</th>
<th>12</th>
<th>20</th>
<th>44</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>3</td>
<td>8</td>
<td>11</td>
<td>13</td>
<td>24</td>
</tr>
<tr>
<td>5</td>
<td>8</td>
<td>15</td>
<td>20</td>
<td>25</td>
<td>39</td>
</tr>
<tr>
<td>9</td>
<td>14</td>
<td>22</td>
<td>32</td>
<td>37</td>
<td>47</td>
</tr>
<tr>
<td>12</td>
<td>18</td>
<td>31</td>
<td>39</td>
<td>44</td>
<td>52</td>
</tr>
<tr>
<td>14</td>
<td>26</td>
<td>36</td>
<td>43</td>
<td>48</td>
<td>58</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>C</th>
<th>F</th>
<th>C</th>
<th>F</th>
<th>F</th>
<th>F</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>C</td>
<td>F</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
</tr>
<tr>
<td>C</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
</tr>
<tr>
<td>C</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
</tr>
<tr>
<td>C</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
</tr>
<tr>
<td>C</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
</tr>
</tbody>
</table>

Heights

Results (C are potential channels)
• Method works because **concavities** are effectively treated as stream beds – flow converges in concave areas...
• Networks are **disconnected** – this an early method which is no longer used...
• Dual can be used to find **ridges**...
Calulating flow accumulation

- Most approaches to modelling channels use **flow direction** algorithms as a starting point.
- Flow accumulation is then calculated as the **sum of upstream elements** draining to a cell.
  - For every cell count how many neighbours drain to it.
    - For each of these cells do the same.
    - Repeat until all cells are either on the boundary of the grid or do not have upstream cells.
- When we calculate the flow accumulation for a cell, we can also (if we **flag** the contributing cells) calculate the **catchment boundary** for that cell.
- If we choose some **pour point** in the channel network, we can calculate the **catchment upstream** of that point.
Flow accumulation example

Flow direction key

<table>
<thead>
<tr>
<th>32</th>
<th>64</th>
<th>128</th>
</tr>
</thead>
<tbody>
<tr>
<td>16</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>4</td>
<td>2</td>
</tr>
</tbody>
</table>

FlowAcc = 2+1+1+2+1

Associated watershed
Potential problems with flow accumulation

- If flow directions converge on a single cell then flow accumulation network is broken.
- Such locations are local minima known as sinks and are generally considered to be undesirable artefacts of the terrain model.
- Two basic strategies to remove sinks:
  - 1) Find the nearest adjacent cell with elevation the same as or lower than the sink – hard code a flow direction to it (radial search).
  - 2) Increase the elevation of the sink to that of a neighbouring cell – check neighbouring cell then drains (sink filling) (ArcGIS approach).

NB: In ArcGIS, undefined flow directions are also treated as sinks.
Flow accumulation network broken by sinks

Note also problems in lake
Identifying channels

• Several approaches to identifying channels
  – Use of a threshold to identify channels (e.g. cells with a flow accumulation greater than some specific value are flagged as streams)
  – This threshold may be purely empirical (does the stream network look realistic) or based on some knowledge of hydrologic properties (see later)...
  – A second grid can be overlaid on the flow accumulation grid to simulate input precipitation
  – If we assume all this precipitation is overland flow then channel initiation can be calculated as some threshold of accumulated water...
Channels identified in a filled DEM with a simple threshold accumulation area of 250 cells.
Aside: TIN-based flow directions

• Calculating flow directions on a TIN generally uses a similar ‘steepest-path approach’ as D8
• Flow can be across or between triangles
• Steepest-path is not limited to 45° increments
• Flow between triangles (where two triangles slope down towards a shared edge) is treated as channelled flow
• Similar potential problems (flat triangles and pits) require handling as in rasters
Illustration of TIN-based flow

- Image from Martinoni, D. 1997: Extraction of hydrologic structures from triangulated DTM
- Note flow **across** and **between** triangles..
Summary

• Briefly considered why hydrology is so important – you should read/ know about this from your other courses!
• Examined different algorithms for flow direction calculation – important notions of convergence or divergence...
• Looked at how we can estimate stream position – e.g. using flow accumulation or focal operators
• Noted problem of sinks in deriving flow accumulation
• Finally, saw example of calculation of flow direction from TINs
Spatial analysis in hydrology

• We will look at 2 examples:
  – A set of so-called ‘compound’ topographic indices describing physically meaningful catchment properties – we will relate these to the USLE and particularly the LS parameter
  – A study looking at the impact of DEM resolution on derived stream network
Compound topographic indices

- **Compound indices** use elements derived directly from the DEM to generate some physically meaningful property which varies in space.
- Examples include potential incident radiation which we saw last week:

\[
I = [\cos \theta_0 \cos \beta + \sin \theta_0 \sin \beta \cos(\phi_0 - A)]S_0 \times \exp(-T_0 / \cos \theta_0)
\]

- Today we will look at two indices:
  - The **wetness index**; and
  - the **stream power index**.
Wetness index

\[ w_r = \ln \left( \frac{A_s}{T \tan \beta} \right) \]

or

\[ w = \ln \left( \frac{A_s}{\tan \beta} \right) \]

Here:
- \( A_s \) is the **specific catchment area** – defined as the upslope area per unit length of contour;
- \( \beta \) is the **gradient**; and
- \( T \) is the **transmissivity** of a saturated soil profile.

If we assume the soil to be **homogenous** over the catchment then the second form is used (\( T = 1 \))
Wetness index...
What does wetness index mean?

- **Correlation** found between $w$ (wetness index) and **distribution of surface soil water content** (in a small fallow catchment)

- Indices using the product of **plan curvature** (rate of change of aspect) and **aspect** also found to give good correlations with surface soil water

- **What physical reasons** might there be for these correlations?
Erosion processes

- **Stream power** can be defined as
  \[ \Omega = \rho g q \tan \beta \]
  where \( \rho \) is the density of water;
  
  \( g \) is the acceleration due to gravity;
  
  \( q \) the discharge per unit width.

- \( A_s \) (specific catchment area) is often considered to be **proportional to discharge**, so a **compound index** for stream power is (\( \rho \) and \( g \) are constants)
  \[ A_s \tan \beta \]
Stream power index

Where is stream power high?
An application of these indices

- Moore found that these two indices in combination were good predictors of location of **ephemeral**\(^1\) gullies
- For a semi-arid catchment in Australia they found gullies formed when \(w > 6.8\) and \(A_s \tan \beta > 18\); and \(w > 8.3\) for a catchment in Antigua
- The value of \(w\) varies because of different soil properties in the different catchments

---

\(^1\)That is in existence, power, favour, popularity, etc. for a short time only; short-lived; transitory.
Using indices in USLE

- Moore shows how the **physical potential** for **erosion** can be evaluated using a term equivalent to LS, where

\[
LS = (n + 1) \left( \frac{A_s}{22.13} \right)^n \left( \frac{\sin \beta}{0.0896} \right)^m
\]

with \( n = 0.4 \) and \( m = 1.3 \).

- This equation is suggested to be more amenable in complex topographies than the empirical equation.

- You have seen this form of LS in the lab.

Recall LS is the slope-length factor in the USLE...
Note how LS is high, **both in channels** and on **open slopes** – accounting for erosion by **sheet** and **channelled flow**...
Back to defining channels

• McMaster investigates how important **DEM resolution** is in defining the accuracy of predicted stream locations

• **Accurate stream location** (and definition) is important because:
  – it determines **travel distances** on hillslopes and **network link length**
  – this in turn influences **modelled flood lag times** and magnitudes

• Remember:
Threshold resolutions for stream networks

• McMaster refers to a result suggesting:

  “ratio of **average elevation change** per pixel to **average elevation error** greater than one”
  for

  “reasonably accurate hydrological network”

• Thus in low elevation areas absolute error must be less (obvious!?)

• This result implies we should **divide** DEMs into areas where this criteria is met...
Critical support area

- Earlier, I guessed a threshold accumulation area to define a channel network.
- In practice, the critical support area required to support a permanent stream is a function of slope characteristics, soil properties, groundwater interaction, surface cover and climatic conditions.
- Thus the critical support area is spatially variable.
- McMaster derived it for his area from topographic maps by averaging over 111 headwaters.
- The value he derived is site-specific.
Example critical support area...

Critical support area is the area of the watershed upstream of the stream entity on the map.

Data © SwissTopo
DEM resolution effects

- He then went on to look at the dependence of some stream parameters on resolution.
- Firstly, he explored how accuracy of streams generated from the DEM compared to mapped positions at different resolutions for D8 and $D_{\infty}$.

Accuracy of Stream network declines at grid size $\sim$200m)
Hillslope length

- Secondly he explored how **slope** varied with accumulation area
- Inflection point suggests **break in slope** - change from sheet to channelled flow

![Graph showing average slope length](image)
What does this mean?

- The **average hillslope length** appears to be a critical variable in defining a **meaningful resolution** for accurate channel definition.
- So appropriate resolutions will **vary** for different landscapes...
- If we look at some pictures this is easy to see...
Pictures from a presentation by Tarboton at http://www.crwr.utexas.edu/terrainAnalysis/presentations.cfm
Summary

• Looked at compound indices for wetness and stream power
• Saw how these could be used in calculating the LS factor of the USLE in a different way
• You should consider what would happen if you did this in your Anjeni work...
• We then looked at a paper which considers the importance of terrain resolution on deriving stream network position...
Next week

- Next week I will talk about **multi-criteria analysis** – using what we have seen up to now to make decisions (one of Openshaw’s 3 examples of spatial analysis)
References

• Moore, I.D. 1996. Hydrologic modelling and GIS. In Goodchild et. al. (eds.) *GIS and Environmental Modeling: Progress and Research Issues*.