Open source approaches to integration and analysis of geo-spatial data in water catchments

Example applications in hydrological and pollutant transport models

Peter Metcalfe[1(*)], Keith Beven[1] and Barry Hankin[2]

[1] Lancaster Environment Centre, Lancaster, UK [2] JBA Trust, Skipton, UK *Corresponding author: p.metcalfe@lancaster.ac.uk

The problem

Huge quantities of geo-spatial data now available for integrated catchment research :

- Topography; digital elevation models (ASTER, GDEM) ; channel networks;
- Land cover and use ;
- Soils classification (e.g. HOST);
- Hydrology and climate - discharge and rainfall records at point locations

Such data may have many projections, resolutions and extents.

How to combine these to produce geographically-meaningful input to integrated catchment models ***and retain their geo-spatial context*** *in analysis and outputs?*

Integration using standard GIS tools powerful *but* time-consuming and expensive (training & software)

Results in proprietorial data formats difficult to integrate with other tools and to distribute and validate with catchment stakeholders

Our approach

Use raster processing and geo-spatial capabilities from the `raster`, `sp` and other libraries within the free, open-source R language and environment.

Employ spatially-enabled databases (SQL Server, POSTGRES) for maintenance of large sets of environmental data

Leverage familiar tools such as Google Earth for visualisation of model results

Emphasise open data exchange formats: KML, XML and ASCII text for distribution of model results to other specialists and catchment stakeholders

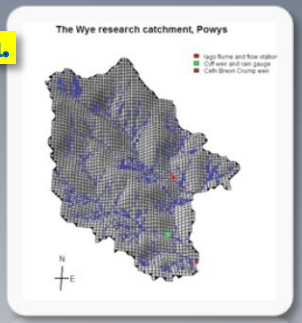
Investigate transforms of raster data into structured, hierarchical data structures that maintain geo-referencing information, for example `Igraph`

Case study: integrating geo-spatial data within a hydrological model

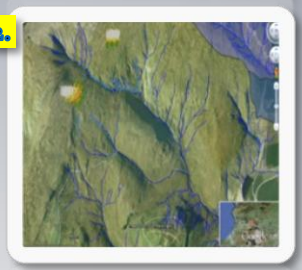
R implementation of semi-distributed Dynamic TOPMODEL (Beven and Freer 2001)

- Inputs: geo-referenced raster of catchment elevations and soil types plus hydrological times series (precipitation, evapotranspiration)
- Outputs : simulated flows across desired time range, HSU storages and fluxes ; rasters of storage deficits and saturated areas.
- Output presented using R; can be exported to Google Earth via KML

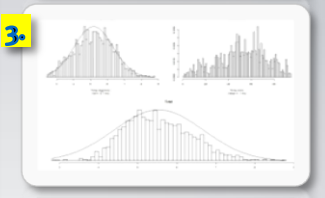
1.



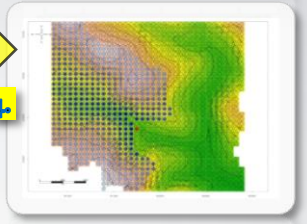
2.

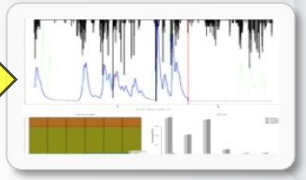


3.



4.





Case study: estimation of catchment travel times and pollutant fluxes using a particle tracking approach

1. Obtain catchment DEM (raster) and Digital River Network (ESRI shape) from CEH
2. "Seed" DEM with conceptual "polluted" water particles taking fast overland or slower subsurface pathways
3. Route particles semi-randomly overland with weightings derived from multi-directional downslope flow algorithm (Quinn *et al.* 1991). Particles retain geographical context, spatial query allows estimation of time to intersect DRN. Integrate particle loads to obtain pollutant fluxes into channel reaches
4. DRN converted to directed graph structure and used to route particles to outlet with fixed wave velocity (beven 1979)

Conclusions and further developments

- Open source approach supported for implementation of established hydrological model. Prepare for release as CRAN package after testing, and calibration and validation of results
- Add solute transport to the model - see Page *et al.* (2007) modelling Cl signal in upper Severn.
- Consider paths in ungauged subcatchments?

Acknowledgements

We gratefully acknowledge from the Centre for Global Eco-innovation and the JBA Trust the funding and support for the PhD studentship of which this research is a part, and for the assistance of Dr.Rob Lamb , Dr Duncan Whyatt and others at JBA and LEC.

References

Beven, K. J. and J. Freer (2001). "A Dynamic TOPMODEL." Hydrological Processes 15(10): 1993-2011.
Robert J. Hijmans & Jacob van Etten (2012). raster: raster: Geographic data analysis and modeling. R package version 2.0-41. http://CRAN.R-project.org/package=raster
Csardi G, Nepusz T: The igraph software package for complex network research, InterJournal, Complex Systems 1695. 2006.http://igraph.sf.net
Page, T., K. J. Beven, et al. (2007). "Modelling the chloride signal at Plynlimon, Wales, using a modified dynamic TOPMODEL incorporating conservative chemical mixing (with uncertainty)." Hydrological Processes 21(3): 292-307.
Pebesma, E.J., R.S. Bivand, 2005. Classes and methods for spatial data in R. R News 5 (2), http://cran.r-project.org/doc/Rnews/.
Quinn, P., K. Beven, et al. (1991). "The Prediction of Hillslope Flow Paths for Distributed Hydrological Modeling Using Digital Terrain Models." Hydrological Processes 5(1): 59-79.988-2003