How innovations thrive in GRASS GIS

Václav Petráš[1], Yann Chemin[2], Martin Landa[3], Thomas Leppelt[4], Pietro Zambelli[5], Luca Delucchi[6], Margherita Di Leo[7], Sören Gebbert, and Markus Neteler[8]

[1]North Carolina State University, USA (wenzeslaus@gmail.com, vpetras@ncsu.edu); [2]IWMI, Sri Lanka; [3]FCE CTU in Prague, Czech Republic; [4]DWD, Germany; [5]EURAC Research, Institute for Renewable Energy, Italy; [6]Fondazione Edmund Mach, Research and Innovation Centre, Italy; [7]European Commission, JRC, Italy; [8]mundialis GmbH & Co. KG, Germany;

Highlights

Poster topic highlights

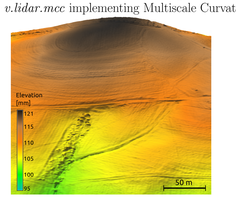
- Algorithms and models included in GRASS GIS remain available long term (Chemin et al., 2015).
- Analytical tools are not limited to one domain but spread across many fields.
- New tools can be built based on functionality or code of the existing ones regardless of the particular domain of problems they belong to.
- Both the functionality and the code are evaluated by the community of users and developers in different fields and scales.

General GRASS GIS highlights

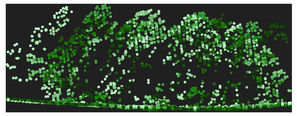
- The GRASS GIS development team takes care of interface and operating system related changes in the code provided by scientists and contributors.
- Cost estimated by Black Duck Open Hub is over $30,000,000.
- The free, libre and open source license used in GRASS GIS not only allows users to reduce GIS software costs to minimum, but also enables organizations to deploy cloud-based solutions without any license restrictions.
- GRASS GIS is used both directly and as a backend in other projects such as QGIS and R.

Lidar Data Processing

Filtering of ground and non-ground points was included into GRASS GIS early on in the lidar era in the *v.lidar.edgedetection* group of modules. The module *v.surf.rst* for spatial interpolation was developed approximately 20 years ago, but since then it has been improved several times (GRASS GIS Development Team, 2015b) including recent parallelization which will be included in GRASS GIS 7.4. Currently the module is used for both digital terrain model interpolation and interpolations in general.
The module *r.in.lidar* statistically analyzes massive point clouds. For advanced general point analysis GRASS GIS *v.outlier* implemented originally specifically for lidar data and recently also *v.cluster* implementing variety of clustering methods including DBSCAN (Density-Based Spatial Clustering of Applications with Noise) and OPTICS (Ordering Points to Identify the Clustering Structure). The *v.outlier* module serves as a base for a user contributed module *v.lidar.mcc* implementing Multiscale Curvature Classification (MCC).



Digital elevation model interpolated from lidar point clouds using *v.surf.rst*. Data are showing tillage in an agricultural field near Raleigh (North Carolina, USA)



Profile (vertical slice) of a small portion of a point cloud showing tree structure as captured by the returns (*v.profile.points*)

Acknowledgements

We acknowledge all members of the GRASS GIS community both users and developers. We would like to thank Markus Metz, Anna Petrasova, Stepan Turek, and Radim Blazek.

Open Source Geospatial Foundation (OSGeo) supports the collaborative development of open source geospatial software. GRASS GIS is a OSGeo project and OSGeo provides infrastructure for project websites, mailing lists and source code management.

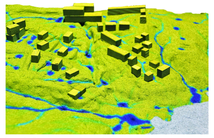
Initial development of *i.segment* module as well as several other developments were done as a part of the Google Summer of Code project. Google provides financial support to students and organizations participating in the Google Summer of Code project.

grass.osgeo.org

This poster is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License.

Water, Floods and Erosion

GRASS GIS entails several modules that constitute the result of active research on natural hazards. The *r.sim.water* simulation model (Mitas and Mitasova, 1998) for overland flow with spatially variable rainfall excess conditions was integrated into the Emergency Routing Decision Planning system as a WPS (Raghavan et al., 2014). The module *r.sim.water* together with the module *r.sim.sediment* for erosion-deposition modeling implements a path sampling algorithm which is robust and easy to parallelize. A unique least cost path algorithm which doesn't need any depression filling was implemented in *r.watershed* module in 1989 and it was updated for massive data sets in 2011. The *r.sim.water* module was also utilized by Petrasova et al. (2014) and is now part of *Tangible Landscape*, a tangible GIS system, which also incorporated *r.damflood*, a dam break inundation simulation by Cannata and Marzocchi (2012).

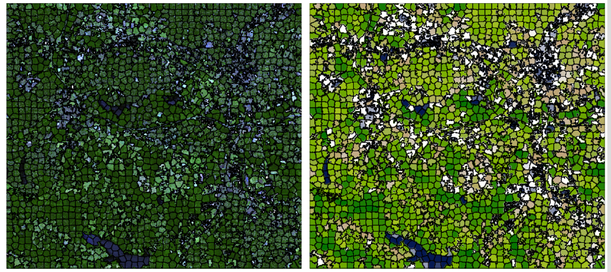


Overland flow simulated by *r.sim.water* used for landscape architecture design in Tangible Landscape environment (Historical master plan for Lake Raleigh, NC, USA)

Dam breach on Lake Raleigh (NC, USA) in Tangible Landscape environment simulated using *r.damflood* module

Image Segmentations

r.smooth.seg (formerly *r.seg*) was created by Alfonso Vitti (Vitti, 2008, 2012) and implemented a piece-wise smooth approximation of the original image according to Mumford and Shah (1989) and March and Dozio (1997) which can be used to reduce noise in the original image.
This supplemented *r.clump* available from 1980s which groups pixels with the same categories (or integer values). The latest version of *r.clump* coming in GRASS GIS 7.4 supports multiple image bands (or any rasters maps) as input. Clumping of cells based on threshold value and clumping of double precision floating point input is supported as well in the new version.
Eric Momsen implemented initial version of region-growing image segmentation in 2012 and Markus Metz then extended and optimized the code resulting in the inclusion of the *i.segment* module into GRASS GIS 7.0 and final replacement of old multi-resolution classification module of the same name which was available in 1992 (Zhuang, 1992). Another module called *i.segment.hierarchical* by Pietro Zambelli is based on *i.segment* and performs parallelized hierarchical segmentation.

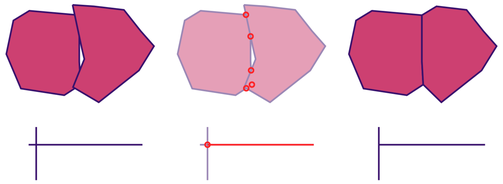


Superpixels (black outlines) on pseudo-color image of central Wake county, NC, USA (left) and the same superpixels colored according to the mean NDVI value per pixel (right).

In 2016 GRASS GIS implementation of SLIC Superpixels segmentation (Achanta, 2010; Achanta et al., 2012) was requested by the community and Rashad Kanavath and Markus Metz implemented *i.superpixels.slic* which provides users both with SLIC and SLIC0 methods.

Topology, Cleaning, Overlays, Attributes

Besides basic vector analysis tools such as *v.buffer* and *v.overlay*, suite of functionality for topology checking and cleaning is available through *v.build* and *v.clean* modules. The vector cleaning tools are particularly advantageous considering the use of non-topological exchange formats. The *v.clean* module was introduced in 2002 and several different contributors extended its functionality with last large set of improvements included in GRASS GIS 7.0. For that release all topological cleaning tools have been optimized with regard to processing speed, robustness, and system requirements. The processing speed of the *v.clean* module was substantially improved as well.

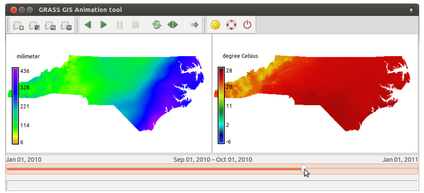


Original vector imported without cleaning (left), identified errors (middle), automatically topologically corrected vector (right)

An experimental module *v.feature.algebra* (formerly *v.mapcalc*) from 2002 was replaced by Python for GRASS GIS C libraries (PyGRASS) in GRASS GIS 7.0. However, new *v.mapcalc* module is available with richer syntax and functionality.
Attribute processing and queries in GRASS GIS can take an advantage of latest developments in database management systems (DBMSs) as several DBMSs are supported most notably SQLite and PostgreSQL. A PostGIS database can be connected including its geometry and topology. In parallel to that native support the OGC Simple Features is available.
The GRASS GIS 7 releases come with faster vector processing backend and more efficient internal vector format; both disk and memory requirements were reduced. The new spatial index performs queries faster (more than 10 times for large vectors compared to GRASS GIS 6).

Spatio-Temporal Data Analysis

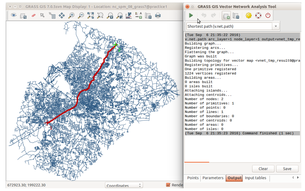
The time dimension was introduced in GRASS GIS 7.0 for vector maps, rasters, and 3D rasters which transformed GRASS GIS into a fully-featured temporal GIS (Gebbert and Pebesma, 2014; Gebbert and Leppelt, 2015). Time series of map layers are managed in space time datasets, a new data type in GRASS GIS, and are still accessible also as individual map layers. Based on the GRASS GIS Temporal Framework Python application programming interface (API), more than 50 modules were implemented to manage, analyze, process and visualize space time datasets. More than 100,000 map layers can be now handled efficiently in GRASS GIS. Example usages of this functionality include analysis of the European Climate Assessment & Dataset ECA&D (Haylock et al., 2008) and temperate climate zone in the European Union identification (Gebbert and Pebesma, 2014).
The new temporal modules (staring with *t.*) work beside well established *r.series* module and specialized modules such as *r.hants* implemented according to Roerink et al. (2000) or *r.seasons*. Latest addition includes raster and vector temporal algebra which can be used for tasks such as computing hydrothermal coefficient for a time series of climate data using the actual mathematical formula (Leppelt and Gebbert, 2015).



Creating a synchronized animation of monthly total precipitation and mean temperature for NC, USA

Vector Network Analysis

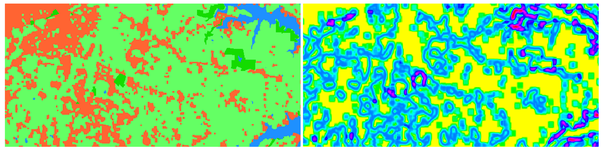
In 2003 Radim Blazek introduced shortest path analysis (*v.net.path*), traveling salesman (*v.net.salesman*), and several other modules for vector network analysis (staring with *v.net*). Over the years, the number of available algorithms increased to over 15 modules including, for example, centrality measures and connected components added in 2009 by Daniel Bundala. In 2014 Stepan Turek implemented turns support into all relevant vector network modules. Also since GRASS GIS 7.0 all vector network analysis tools provide fine control over node costs.



Shortest path between two points in the Wake County road network (NC, USA)

Landscape Structure

A set of modules for multiscale analysis of landscape structure was added in 1992 by Baker and Cai (1992), who developed the *r.le* model similar to FRAGSTATS (McGarigal and Marks, 1995). The modules were gradually improved to become *r.li* in 2006. Further development continued, with a significant increase in speed (GRASS GIS Development Team, 2015a) and a new interactive user interface. Rocchini et al. (2013) used *r.li* modules to implement high level tool for calculating landscape diversity.



Landuse classes and derived landscape diversity according to Shannon index near Charlotte, NC, USA

References

Achanta, R. (2010). SLIC Superpixels, EPFL Technical Report No. 149300.

Achanta, R., Shaji, A., Smith, K., Lucchi, A., Fua, P., and Süsstrunk, S. (2012). SLIC superpixels compared to state-of-the-art superpixel methods. *IEEE transactions on pattern analysis and machine intelligence*, 34(11):2274–2282.

Baker, W. L. and Cai, Y. (1992). The r. le programs for multiscale analysis of landscape structure using the GRASS geographical information system. *Landscape Ecology*, 7(4):291–302.

Cannata, M. and Marzocchi, R. (2012). Two-dimensional dam break flooding simulation: a GIS-embedded approach. *Natural hazards*, 61(3):1143–1159.

Chemin, Y., Petras, V., Petrasova, A., Landa, M., Gebbert, S., Zambelli, P., Neteler, M., Löwe, P., and Di Leo, M. (2015). GRASS GIS: a peer-reviewed scientific platform and future research repository. In *Geophysical Research Abstracts*, volume 17, page 8314.

Gebbert, S. and Leppelt, T. (2015). GRASS GIS: The first Open Source Temporal GIS. In *EGU General Assembly Conference Abstracts*, volume 17, page 7296.

Gebbert, S. and Pebesma, E. (2014). A temporal GIS for field based environmental modeling. *Environmental Modelling & Software*, 53:1–12.

GRASS GIS Development Team (2015a). GRASS GIS Trac: Revision log for r.li. http://trac.osgeo.org/grass/log/grass/trunk/raster/r.li?rev=63988&limit=999.

GRASS GIS Development Team (2015b). GRASS GIS Trac: Revision log for v.surf.rst. http://trac.osgeo.org/grass/log/grass/trunk/vector/v.surf.rst?rev=63988&limit=999.

Haylock, M. R., Hofstra, N., Tank, A. M. G. K., Klok, E. J., Jones, P. D., and New, M. (2008). A European daily high-resolution gridded data set of surface temperature and precipitation for 1950 - 2006. *Journal of Geophysical Research*, 113(D20).

Leppelt, T. and Gebbert, S. (2015). A GRASS GIS based Spatio-Temporal Algebra for Raster-, 3D Raster-and Vector Time Series Data. In *EGU General Assembly Conference Abstracts*, volume 17, page 11672.

March, R. and Dozio, M. (1997). A variational method for the recovery of smooth boundaries. *Image and Vision Computing*, 15(9):705–712.

McGarigal, K. and Marks, B. J. (1995). FRAGSTATS: spatial pattern analysis program for quantifying landscape structure. Technical Report PNW-351, USDA.

Mitas, L. and Mitasova, H. (1998). Distributed soil erosion simulation for effective erosion prevention. *Water Resources Research*, 34(3):505–516.

Mumford, D. and Shah, J. (1989). Optimal approximations by piecewise smooth functions and associated variational problems. *Communications on pure and applied mathematics*, 42(5):577–685.

Petrasova, A., Harmon, B., Petras, V., and Mitasova, H. (2014). GIS-based environmental modeling with tangible interaction and dynamic visualization. In Ames, D.P., Quinn, N.W.T., Rizzoli, A. E., editor, *7th International Congress on Environmental Modelling and Software*, San Diego, California, USA.

Raghavan, V., Choosumrong, S., Yoshida, D., and Vinayaraj, P. (2014). Deploying Dynamic Routing Service for Emergency Scenarios using pgRouting, GRASS and ZOO. In *In Proceedings of FOSS4G Europe, Jacobs University, Bremen, Germany*.

Rocchini, D., Delucchi, L., Bacaro, G., Cavalliini, P., Feilhauer, H., Foody, G. M., He, K. S., Nagendra, H., Porta, C., Ricotta, C., et al. (2013). Calculating landscape diversity with information-theory based indices: A GRASS GIS solution. *Ecological informatics*, 17:82–93.

Roerink, G., Menenti, M., and Verhoef, W. (2000). Reconstructing cloudfree NDVI composites using Fourier analysis of time series. *International Journal of Remote Sensing*, 21(9):1911–1917.

Vitti, A. (2008). *Free discontinuity problems in image and signal segmentation*. PhD thesis, PhD thesis in Environmental Engineering, Università di Trento, Italy.

Vitti, A. (2012). The Mumford-Shah variational model for image segmentation: An overview of the theory, implementation and use. *ISPRS journal of photogrammetry and remote sensing*, 69:50–64.

Zhuang, H. C. (1992). Image processing in GRASS.

LaTeX TikZposter