GRASS GIS: a peer-reviewed scientific platform and future research repository

Yann Chemin[1], Václav Petráš[2], Anna Petrášová[2], Martin Landa[3], Sören Gebbert[4], Pietro Zambelli[5], Markus Neteler[6], Peter Löwe[7], and Margherita Di Leo[8]

[1]IWMI, Sri Lanka; [2]NCSU, USA; [3]FCE CTU in Prague, Czech Republic; [4]TICSA, Germany; [5]EURAC, Italy; [6]CRI, FEM, Italy; [7]TIB Hannover, Germany; [8]EC-JRC, Italy

Introduction

Over the last decades, GIS has become a key driver in geospatial science, research and application. GIS software which is licensed under a free and open source software (FOSS) licence is more than just a mere tool for spatial analysis.

GRASS GIS (Neteler et al., 2012 [17]), a free and open source GIS, is used by many scientists directly or as a backend in other projects such as R or QGIS to perform geoprocessing tasks.

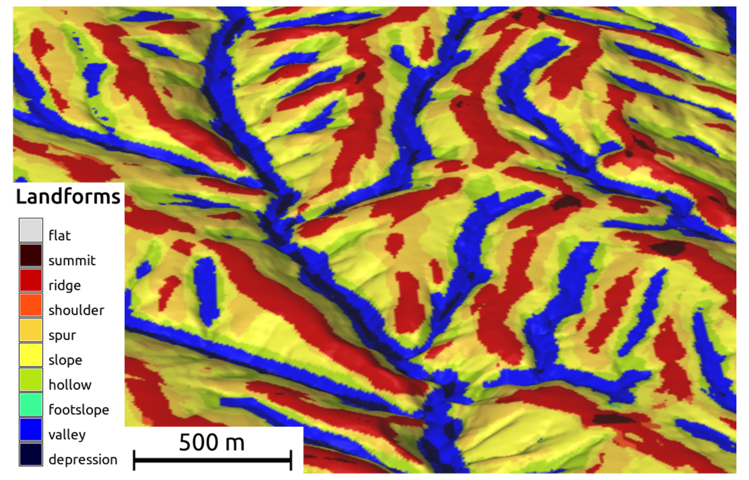
Thanks to the user and developer community, submitted code is evaluated in different fields of application beyond the field of expertise of the original authors, and different scales of magnitude of data. This exceeds the established review process for scientific writing in a given journal or a data publication in a defined field of science.

Immediate access to the software repository enables instant quality checking of the current software version both by continuous automated tests (Petras, 2014 [18]), and code review by human experts.

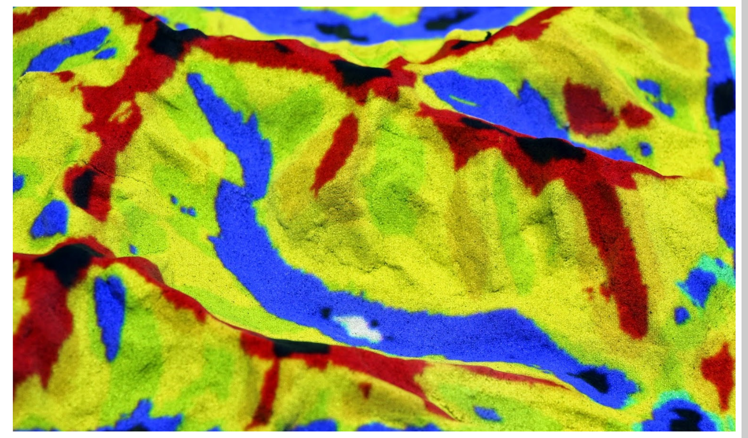
New scientific algorithms can be developed against the reviewed functionalities already provided by the GRASS GIS codebase. This avoids unnecessary overheads, by re-implementation, ensures quality by use of trusted components and allows reuse and long term preservation within the project software repository. Integrating scientific algorithms into GRASS GIS helps to preserve reproducibility of scientific results over time as the original author designed it (Rocchini & Neteler, 2012 [22]).

Landform detection: Geomorphons

Jasiewicz and Stepinski [13] developed a method and module *r.geomorphon* which provide orientation-invariant and relief-invariant method to classify landforms in a scale-independent way. Ashtekar et al. [1] used geomorphons to study soil properties in northwestern South America.



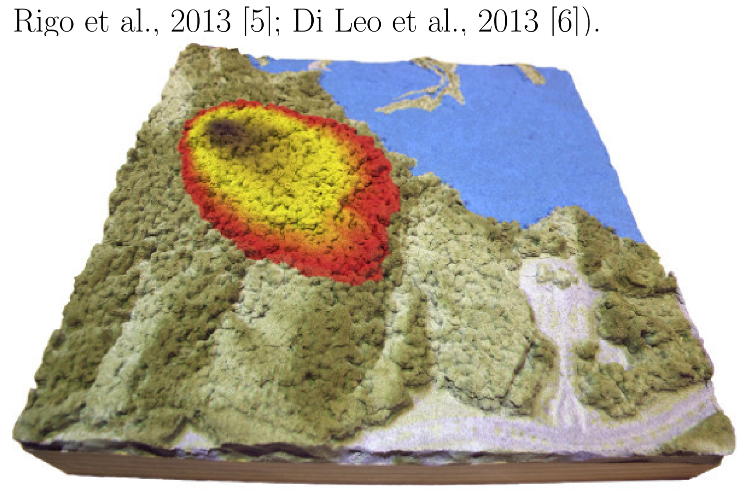
Geomorphons for part of Yakima Training Center (area 5x3 km, USA)



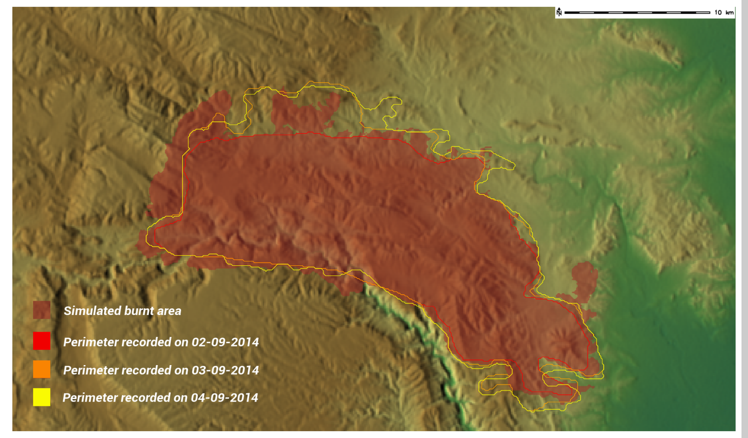
Ridges and valleys detected on the sand physical model of High Tatras (Slovakia) using *r.geomorphon* integrated in Tangible Landscape system

Natural Hazards: Wildfire Spread

The wildfire simulation toolset, originally developed by Xu (1994 [27]) implements Rothermel's model [24]. It is available through the GRASS GIS modules *r.ros* and *r.spread* and is object of active research. It has been extensively tested and recently adapted to European fuel types (Rodriguez-Aseretto et al., 2013 [23]; de Rigo et al., 2013 [5]; Di Leo et al., 2013 [6]).



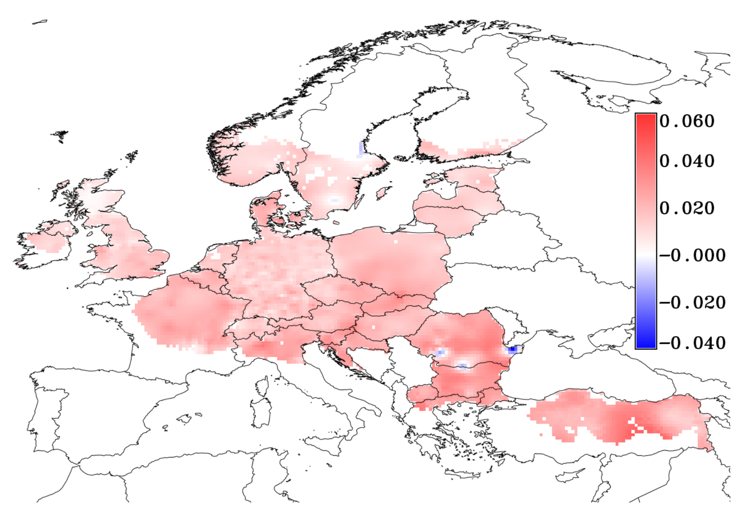
Wildfire simulation in Tangible Landscape environment



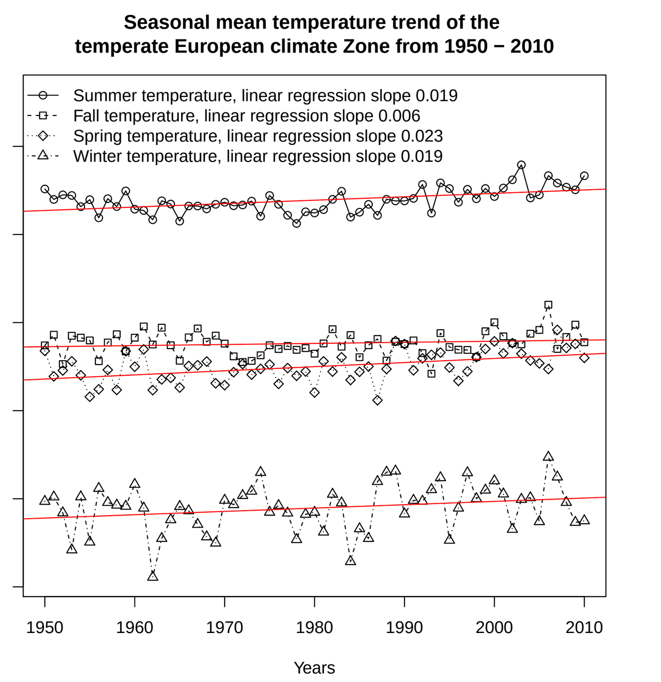
Major wildfire event near Valencia (Spain) between June 28th and July 4th 2012. The actual perimeters recorded by JRC-EFFIS are shown, in comparison with burnt area simulated in GRASS GIS.

GRASS GIS as Temporal GIS

The time dimension was introduced in GRASS GIS version 7 for raster, 3D raster and vector map layers, transforming it into a full featured temporal GIS (Gebbert and Pebesma, 2014 [9]). Time series of map layers are managed in space time datasets, a new data type in GRASS GIS. Based on the GRASS GIS Temporal Framework, more than 45 modules were implemented to manage, analyze, process and visualize space time datasets. The temporal enabled GRASS GIS is capable of efficiently handling more than 100,000 map layers; e.g., it was used to analyze the European Climate Assessment & Dataset ECA&D (Haylock et al. [12]) for climate change indicators.



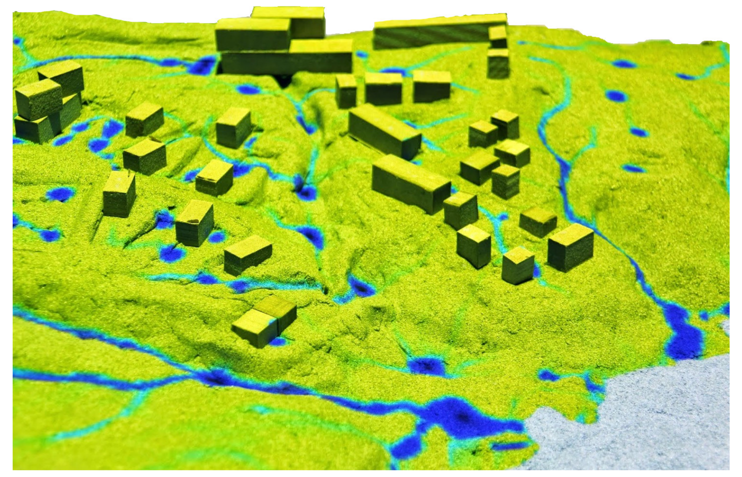
The linear regression slope computed with t.rast.series of the temperate climate zone in the European Union for summer season from 1950 - 2011 (Gebbert and Pebesma, 2014 [9]). Red color indicates rising temperature, blue indicates falling temperature. Units are degree Celsius per year.



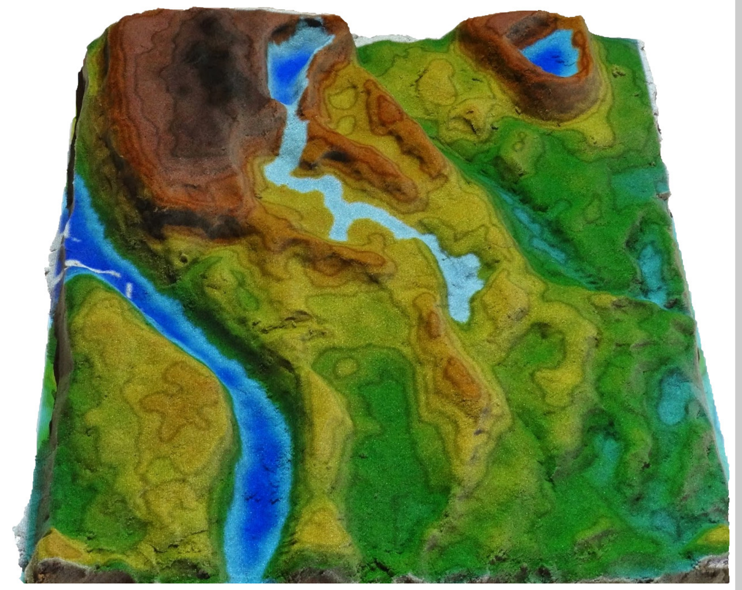
Seasonal mean temperature trend for the temperate climate zone of the European Union (Gebbert and Pebesma, 2014 [9]). The modules *t.rast.aggregate*, *t.rast.univar* and the open source statistical software system R were used to create this plot.

Natural Hazards: 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 [15]) 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 [20]). 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. The *r.sim.water* module was also utilized by Petrasova et al., 2014 [19] and is now part of *Tangible Landscape*, a tangible GIS system, which also incorporated *r.damflood*, a dam break inundation simulation [3].



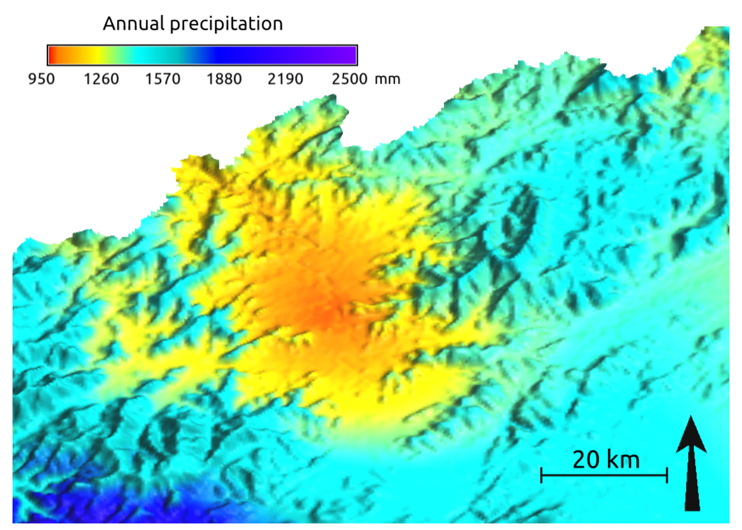
Overland flow used for landscape architecture design in Tangible Landscape environment



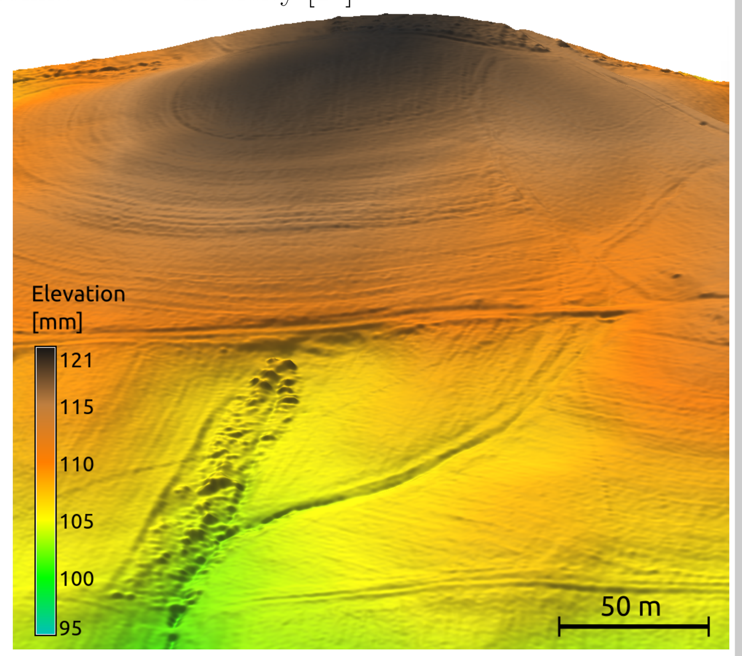
Coal ash pond breach in Tangible Landscape environment using *r.damflood* module

Spatial interpolation

The module *v.surf.rst* for spatial interpolation was developed approximately 20 years ago, since then it has been improved several times [11]. It is now an important part of GRASS GIS and is even taught at geospatial modeling courses, for example at North Carolina State University [16].



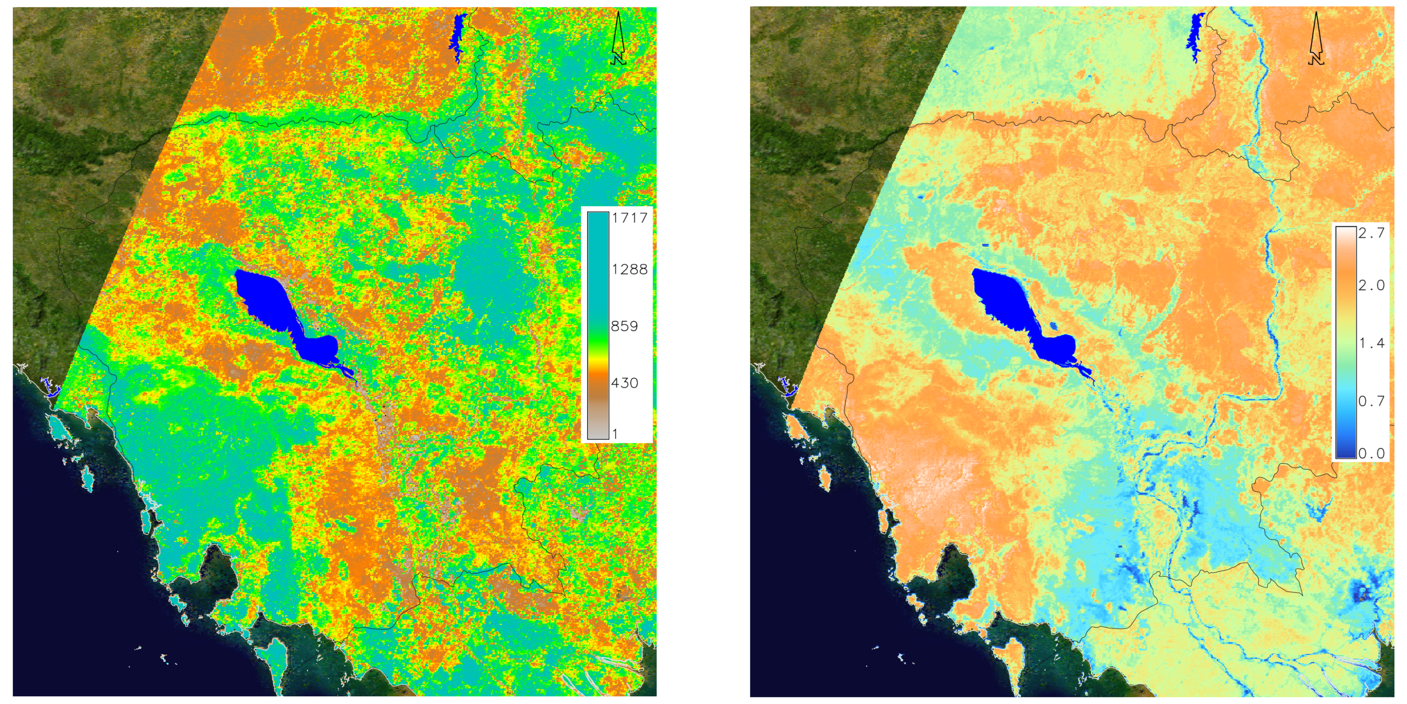
Precipitation interpolated from meteorological stations in 3D space using *v.vol.rst* in the area of North Carolina mountains (USA)



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)

Evapotranspiration (ET)

With the various types of actual ET models being developed in the last 20 years, it becomes necessary to inter-compare methods. Most already published ETa model comparisons address a low number of models, and small to medium areas (Chemin, 2014 [4]; Gao and Long, 2008 [7]; Garcia et al., 2007 [8]; Suleiman et al., 2008 [25]; Timmermans et al., 2007 [26]). With the large amount of remote sensing data covering the Earth, and the daily information available for more than twelve years (i.e. Aqua/Terra-MODIS) for each pixel location, it becomes paramount to have a more complete comparison, in space and time.

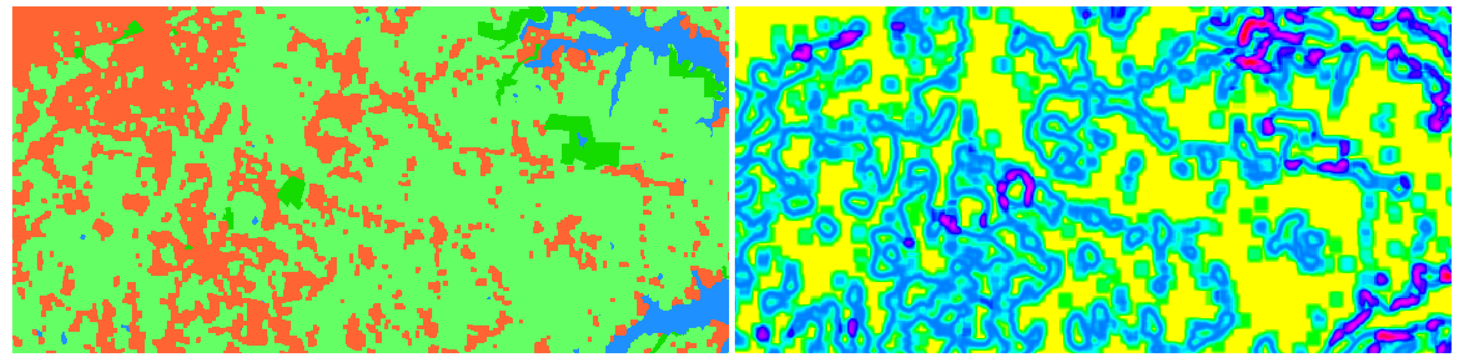


Biological fraction of 2014 ET (cumul. mm/y) & January average gap (mm/m), Tonle Sap, Cambodia.

To address this new experimental requirement, a distributed computing framework was designed and created (Chemin, 2012 [4]). The architecture design was built from original satellite datasets to various levels of processing until reaching the input dataset requirements of various ETa models. Each input product is computed once and reused in all ETa models requiring such input. This permits standardization of inputs as much as possible to reduce variations of models to their own internals/specificities. All of the ET models are available in the new GRASS GIS version 7 as imagery modules and replicability is complete for future research.

Landscape structure

A set of modules for multiscale analysis of landscape structure was added in 1992 by Baker et al. [2], who developed the *r.le* model similar to FRAGSTATS [14], see manual. The modules were gradually improved to become *r.li* in 2006. Further development continued, with a significant increase in speed [10] and a new interactive user interface. Rocchini et al. [21] used *r.li* modules to implement high level tool for calculating landscape diversity.



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

Conclusions

- Algorithms and models, included in GRASS GIS remain available long term (already for 30 years).
- The GRASS GIS development team takes care of API and operating system related changes.
- Scientists can use highly specialized tools implemented by others as a result of having both scientific publications and source code at hand.
- The long term preservation of knowledge allows scientists to build new research and tools upon existing know-how.

References and Acknowledgements

[1] J. M. Ashtekar, P. R. Owens, R. A. Brown, H. E. Winzeler, M. Dorantes, Z. Libohova, M. Dasilva, and A. Castro. Digital mapping of soil properties and associated uncertainties in the Llanos Orientales, South America. In *GlobalSoilMap: Basis of the global spatial soil information system*, page 367. CRC Press, 2014.

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

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

[4] Y. Chemin. *A Distributed Benchmarking Framework for Actual ET Models*. INTECH Open Access Publisher, 2012.

[5] D. de Rigo, D. Rodriguez-Aseretto, C. Bosco, M. Di Leo, and J. San-Miguel-Ayanz. An architecture for adaptive robust modelling of wildfire behaviour under deep uncertainty. In Jiří Hřebíček, Gerald Schimak, Miroslav Kubásek, and Andrea Rizzoli, editors, *Environmental Software Systems. Fostering Information Sharing*, volume 413 of *IFIP Advances in Information and Communication Technology*, pages 367–380. Springer Berlin Heidelberg, 2013.

[6] M. Di Leo, D. de Rigo, D. Rodriguez-Aseretto, C. Bosco, T. Petroliagkis, A. Camia, and J. San-Miguel-Ayanz. Dynamic data driven ensemble for wildfire behaviour assessment: A case study. *IFIP Advances in Information and Communication Technology ISSN:1868-4238. Special issue: "Environmental Software Systems. Fostering sharing information"*, 413:11–22, 2013.

[7] Y. Gao and D. Long. Intercomparison of remote sensing-based models for estimation of evapotranspiration and accuracy assessment based on SWAT. *Hydrological processes*, 22(25):4850–4869, 2008.

[8] M. García, L. Villagarcía, S. Contreras, F. Domingo, and J. Puigdefábregas. Comparison of three operative models for estimating the surface water deficit using ASTER reflective and thermal data. *Sensors*, 7(6):860–883, 2007.

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

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

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

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

[13] J. Jasiewicz and T. F. Stepinski. Geomorphons – a pattern recognition approach to classification and mapping of landforms. *Geomorphology*, 182:147–156, 2013.

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

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

[16] NCSU. GIS/MEA582: Geospatial modeling and analysis, 2015. http://courses.ncsu.edu/gis582/common/grass/interpolation_2.html.

[17] M. Neteler, M.H. Bowman, M. Landa, and M. Metz. GRASS GIS: a multi-purpose Open Source GIS. *Environmental Modelling & Software*, 31:124–130, 2012.

[18] V. Petras and S. Gebbert. Testing framework for GRASS GIS: ensuring reproducibility of scientific geospatial computing, 2014. Poster presented at 2014 Fall Meeting, AGU, San Francisco, Calif., 14-18 Dec.

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

[20] V. Raghavan, S. Choosumrong, D. Yoshida, and P. Vinayaraj. Deploying dynamic routing service for emergency scenarios using pgRouting, GRASS and ZOO. In *In Proceedings of FOSS4G Europe, Jacobs University, Bremen, Germany*, July 2014.

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

[22] D. Rocchini and M. Neteler. Let the four freedoms paradigm apply to ecology. *Trends in Ecology & Evolution*, 27(6):310–311, 2012.

[23] D. Rodriguez-Aseretto, D. de Rigo, M. Di Leo, A. Cortés, and J. San-Miguel-Ayanz. A data-driven model for large wildfire behaviour prediction in Europe. *Procedia Computer Science*, 18:1861–1870, January 2013.

[24] R. C. Rothermel. How to predict the spread and intensity of forest and range fires. Technical Report INT-143, US Forest Services, 1983.

[25] A. Suleiman, J. Al-Bakri, M. Duqqah, and R. Crago. Intercomparison of evapotranspiration estimates at the different ecological zones in Jordan. *Journal of Hydrometeorology*, 9(5):903–919, 2008.

[26] W. J. Timmermans, W. P. Kustas, M. C. Anderson, and A. N. French. An intercomparison of the surface energy balance algorithm for land (SEBAL) and the two-source energy balance (TSEB) modeling schemes. *Remote Sensing of Environment*, 108(4):369–384, 2007.

[27] J. Xu. *Simulating the spread of wildfires using a geographic information system and remote sensing*. PhD thesis, Rutgers University, New Brunswick, New Jersey, 1994.

Acknowledgements

We acknowledge Matthew Horvath for *Spill impacts from coal ash pond using GRASS GIS* and Eva Stopkova for physical model of High Tatras.

GRASS GIS project is a Open Source Geospatial Foundation (OSGeo) project and is using OSGeo infrastructure for Web site and code repository.

grass.osgeo.org   grasswiki.osgeo.org