Open-source Versus Proprietary GIS on Landscape Metrics Calculation: A Case Study

Wen-Chieh (Jeffrey) Wang [1], Jia-Liang Yang [2], Ya-Yun Lin [3]

[1] Department of Urban Planning and Landscape Architecture, Chaoyang University of Technology, Taichung, Taiwan, jeffwang@cyut.edu.tw
[2] Graduate School of Architecture and Urban Design, Chaoyang University of Technology
[3] Graduate School of Architecture and Urban Design, Chaoyang University of Technology

Abstract

Landscape ecology is a field that focuses on the ecology of landscape. Ecologists see a landscape as a land mosaic composed of landscape elements, particularly patches. A categorical map is the most common method to represent the patches in a landscape. Landscape metrics are algorithms that quantify certain characteristics of patches, classes of patches, or entire landscape.

Landscape metrics calculation is usually the key part of a landscape pattern analysis in various landscape ecology studies. Because of the necessity of handling spatial data, landscape metrics calculation often requires the aid of a GIS. Nowadays, since both open-source and proprietary GIS solutions are available for landscape metrics calculation, a comparison between these two solutions helps people understand the state of the art in this field.

This paper describes a case study that compares two particular landscape metrics calculation solutions. One is through the GRASS open-source GIS. The other is the combination of the ArcView proprietary GIS plus an external spatial pattern analysis program called FRAGSTATS. The case study is conducted in three steps, examining standard operation procedures, comparing metrics capable of being calculated, and testing with landscape pattern analysis samples.

The study shows that the proprietary but free FRAGSTATS itself turns out to be a more capable, mature, and updated program. With the help of a modern virtualization program, such as the open-source VirtualBox, FRAGSTATS can work with virtually all GIS. This makes devoting more efforts to update the veteran r.le or replace it with a newer one seemingly not so necessary.

1. Introduction

1. 1 Motivation

Landscape metrics are a set of measurements that is used to quantify the spatial pattern of a "landscape." Among various solutions, FRAGSTATS is probably the most popular computer programs that professionals use to calculate landscape metrics in recent years. Although FRAGSTATS is in the public domain and can be freely downloaded from the web, its newest version is a Windows-only program and not yet open-sourced (McGarigal, 2006). Besides working as a standalone program without any GIS (Geographic Information Systems) functionality, FRAGSTATS is also designed to work conveniently with a few GIS through the loose-coupling mechanism, *i.e.* supporting direct input and output of file formats of those GIS. However, all three GIS that FRAGSTATS directly supports are proprietary GIS.

For those who do not work on Windows platform or cannot afford the FRAGSTATS plus a proprietary GIS solution, a viable alternative is using the popular free and open-source GIS, GRASS. Before the original version of FRAGSTATS was published in 1995 (McGarigal & Marks, 1995), GRASS already had a suite of raster modules called r.le (Baker & Cai, 1992), which stands for landscape ecology, that can calculate some landscape metrics.

In addition to focus on Free and Open Source GIS (FOSG) as usual, the 2008 Free and Open Source Software for Geospatial (FOSS4G) conference also wants to encourage exposure, debate, and understanding among FOSS and proprietary communities (FOSS4G 2008, 2008). Therefore, it seems of interest to attendees of this conference to discuss an "applied" topic regarding landscape metrics calculation that used both open-source and proprietary GIS.

1. 2 Objectives

The objectives of this study is as follows.

1. Outline the concept of landscape metrics and explain the relationship between its calculation and GIS.
2. Introduce two particular landscape metrics calculation solutions: one is FOSG-based and the other is conveniently coupled with a proprietary GIS.
3. Evaluate these two solutions through a three-parts case study to uncover their pros and cons.
4. Propose the definite recommendation on selecting one solution over the other, and some suggestions regarding the future development of the FOSG-based solution.

2. Related Theories

2. 1 Land mosaics and patches

Established in late 19 century as a scientific discipline, ecology refers to the scientific study of the distribution and multitude of life and how the distribution and multitude are affected by interactions between the organisms and their environment (Wikipedia contributors, 2008a). Landscape ecology is a sub-discipline of ecology, as well as geography, that focuses on how spatial variation in the landscape influences ecological processes (Wikipedia contributors, 2008b). While traditional ecological studies frequently assume that systems are spatially homogenous, landscape ecology in contrast assumes spatial heterogeneity and thus must take spatial patterns into consideration (Turner & Gardner, 1991). Note that in the terminology of landscape ecology, "landscape" is defined as an area consists of two or more ecosystems in close proximity (Sanderson & Harris, 2000).

As a landscape usually spans over a kilometers-wide area, some types of remote sensing techniques that take a bird's-eye-view from the sky rather a worm's-eye-view from the ground has always been the essential tool in a landscape ecology study. From a bird's-eye-view, individual trees, rocks, houses, and so on in a land mosaic disappear and form conspicuous patches, striking corridors, and a background matrix (Forman, 1995). Take the composite satellite imagery of the Spearfish area for example (Figure 1 left), the landscape appears more like a mosaic with easily recognizable edges between underlying elements than a smooth gradient without discernible edges (GRASS Development Team, 2007a; Forman & Godron 1986). Forman (1995) defined the most discernible landscape element, a patch, that is fundamental to modern landscape ecology as "a relatively homogeneous nonlinear area that differs from its surroundings."

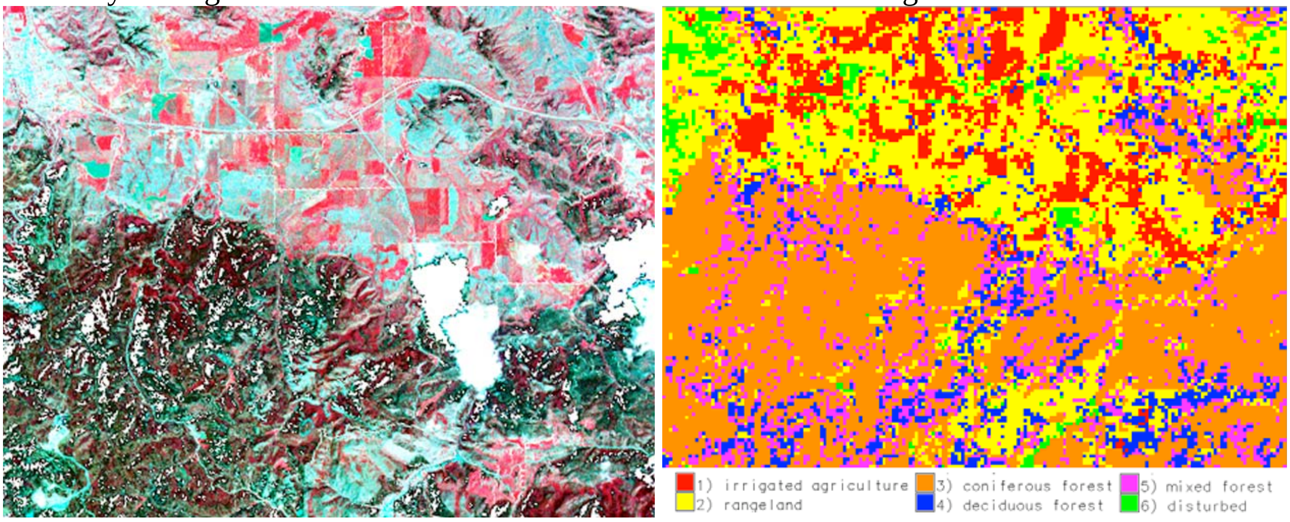

Figure 1. Left: a composite satellite imagery of Spearfish, South Dakota, USA. Right: a categorical map showing vegetation coverage of the same area.

2. 2 Landscape metrics

Due to wide adoptions of Forman's land mosaic analogy to landscape, categorial maps, which represent landscape elements as categories in the map (Figure 1 right), become the most suitable tool to depict landscape patterns. Accordingly, developments and applications of measurements for the categorial map pattern analysis has been popular in related literature and also become an important part in the evolution of landscape ecology as a discipline. McGarigal (2002) uses the term "landscape metrics" to exclusively refer to those measurements developed for categorical map patterns.

Given characteristics of all individual patches in a categorical map, three different levels of landscape metrics can be calculated (McGarigal *et al*., 2002). Patch-level metrics, of course, refer to the spatial character and context of individual patches, such as a patch's size, perimeter, shape, and edge. If the calculation aggregates individual patches into some types of class, for example according to their categories, they are called class-level metrics. When the calculation disregards the classes or categorical difference among individual patches and produces only summary values for the entire landscape in the map, they are called landscape-level metrics.

2. 3 Approaches of landscape metrics calculation solutions

There are two possible approaches to develop computerized solutions for landscape metrics calculation. The first approach is building on top of the existing capabilities of a GIS as a customized module or extension. In GIS world, such an approach is called tight-coupling with GIS.

The second approach is developing a stand-alone program that also provides some kind of data exchange capabilities with GIS. Such an approach is generally referred to as loose-coupling with GIS. The r.le set of programs inside GRASS is a well-known solution that takes the first approach (Baker, 2001). The current FRAGSTATS (version 3.3), on the other hand, is a popular solution that takes the second approach (McGarigal, 2006).

3. Methods and Procedure

The research approach of this study is based on the case study methods (Wikipedia contributors, 2008c). The cases examined and compared are the aforementioned r.le programs within GRASS, which will be referred to as the GRASS solution, and the FRAGSTATS coupled with a popular proprietary GIS, ArcView, which will be called A+F solution hereafter. The actual versions used for this study are ArcView GIS 3.3, Spatial Analyst 2.0a, FRAGSTATS 3.3, and GRASS 6.3. All are installed from the distributed pre-compiled binaries.

This case study will be divided into three parts. First, the standard operating procedure of the two solutions in conducting the landscape metrics calculation are outlined. Second, a comparison of the landscape metrics that the two solutions are capable of calculating is summarized in a table. Third, two spatial pattern analyses (Figure 2) using landscape metrics calculation are performed with both A+F and GRASS solutions. One analysis is evaluating the effects of clustered and scattered farmhouse development styles on land fragmentation (Lin & Wang, 2008) from the perspective of landscape ecology; the other is evaluating the effects of designating a protective buffer zone network along streams and/or roads as a green corridor system in the same regard (Yang & Wang, 2008).

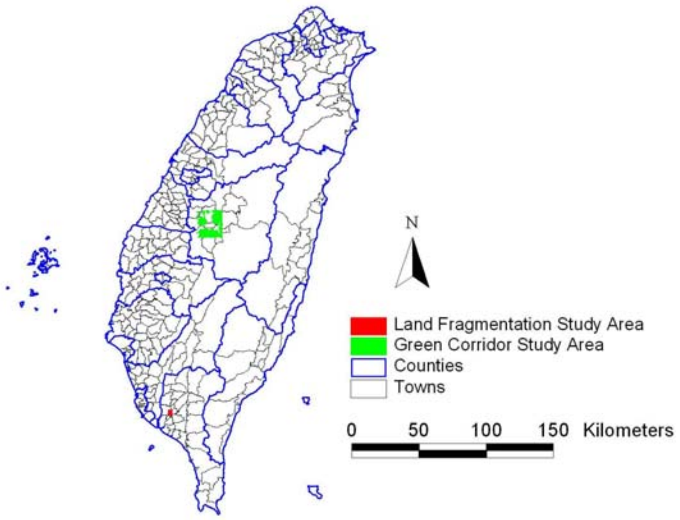

Figure 2. Locations of the two study areas in Taiwan.

4. Results and Discussion

4. 1 Standard operation procedures

For both solution the standard operating procedures are basically the same. They include the following steps.

1. Make sure all required components are properly installed and functional. The A+F solution needs three components, ArcView GIS, SpaticalAnalyst extension, and FRAGSTATS. For the GRASS solution, the GRASS GIS is all that required.
2. Obtain the categorical map in the suitable raster file format at an appropriate grid resolution and extent by either directly importing or rasterizing from a traced vector map. For the A+F solution, the file should be in the ArcGrid format, while for the GRASS solution, it should be a GRASS raster map.
3. Perform landscape metric calculations through a three-stage operation: setting parameters, executing calculation, and browsing the saved results.
4. Visualize and further analyze the results in GIS and other spreadsheet or statistical programs.

For the first step, the most significant difference is the cost of those required components. The cost of obtaining GRASS can be free, but the licensing fee for ArcView and SpatialAnalyst or the newer ArcGIS package ranges from several hundred dollars for academic users to several thousand dollars for business users. In fact, it is the licensing cost that keeps this study using the older ArcView instead of the newer ArcGIS.

For the second step, the two solutions show no major difference except that GRASS has a more flexible and precise way to obtain the desirable grid extent during the rasterization.

For the third step, the GRASS solution has the advantage of working in the same software environment, but in the A+F solution the user must switches from ArcView to FRAGSTATS. In addition, during FRAGSTATS execution, the same categorical map in ArcView must be closed to avoid file corruption caused by simultaneous accessing from both FRAGSTATS and ArcView.

For the final step, there are two differences. First, for examining the resulting maps A+F solution has to switch back to ArcView while the GRASS solution remains in the same environment. Second, r.le provide a convenient way to label the identified patch ID directly on the map.

4. 2 Metrics capable of being calculated

Table 1 shows the landscape metrics that the two solutions can calculate. The metrics acronyms in the table follow those in the FRAGSTATS 3.3 documentation (McGarigal *et al*., 2002). The metrics with both underline and yellow character background fill are those that r.le programs can calculate either directly or indirectly in a straightforward manner, *i.e.* not involving substantial scripting or programming efforts (Baker, 2001). There are a few metrics that are supposed to measure the same characteristics of the landscape but expressed differently in the mathematical formulas in r.le programs and FRAGSTATS. In short, the metrics that r.le programs can calculate seem to be only a subset of those FRAGSTATS can do. Besides, FRAGSTATS produces more detailed and better formatted outcomes.

Table 1. Landscape metrics that FRAGSTATS and r.le programs can calculate.

		Levels of Metrics		
		Patch	Class	Landscape
Aspects of Landscape Pattern Measured	**Area / Density / Edge**	AREA, GYRATE, PERIM	**AREA_dist**, GYRATE_dist, **NP**, **PD**, **TE**, **ED**, **LSI**, LPI, **CA**, PLAND, nLSI	**AREA_dist**, GYRATE_dist, **NP**, **PD**, **TE**, **ED**, **LSI**, LPI, **TA**
	Shape	PARA, SHAPE, FRAC, LINEAR, CIRCLE, CONTIG	PARA_dist, SHAPE_dist, **FRAC_dist**, LINEAR_dist, CIRCLE_dist, CONTIG_dist, **PAFRAC**	PARA_dist, SHAPE_dist, **FRAC_dist**, LINEAR_dist, CIRCLE_dist, CONTIG_dist, **PAFRAC**
	Core Area	CORE, CAI, NCA, ADEPTH, MDEPTH	CORE_dist, DCORE_dist, CAI_dist, TCA, NDCA, DCAD, CPLAND	CORE_dist, DCORE_dist, CAI_dist, TCA, NDCA, DCAD
	Isolation / Proximity	PROX, SIMI, ENN, FNN	PROX_dist, SIMI_dist, **ENN_dist**, FNN_dist	PROX_dist, SIMI_dist, **ENN_dist**, FNN_dist
	Contrast	ECON	ECON_dist, CWED, TECI	ECON_dist, CWED, TECI
	Contagion / Interspersion		PLADJ, AI, IJI, DIVISION, SPLIT, MESH, CLUMPY, MFRAC	PLADJ, AI, IJI, DIVISION, SPLIT, MESH, CONTAG
	Connectivity		COHESION, **CONNECT**, TRAVERSE	COHESION, **CONNECT**, TRAVERSE
	Diversity			PR, PRD, RPR, SHDI, SIDI, MSIDI, SHEI, SIEI, MSIEI

4. 3 Landscape metrics calculation tests

The metrics in blue bold text in Table 1 are those calculated in both spatial pattern analysis tests. Those in red and green bold texts are calculated only in the land fragmentation and green corridor tests, respectively. Table 2 summarizes those metrics with full names and acronyms.

Table 2. Summary of the metrics calculated in the two tests

	Shared Metrics	Unique Metrics
Land Fragmentation Test	Patch area distribution (AREA_dist) Number of patches (NP) Patch density (PD) Edge density (ED) Landscape shape index (LSI)	Fractal index distribution (FRAC_dist) Euclidean nearest neighbor distance distribution (ENN_dist)
Green Corridor Test		Total area (TA) Total edge (TE) Perimeter-area fractal dimension (PAFRAC) Connectance index (CONNECT)

For the land fragmentation test, FRAGSTATS can calculate all directly, but the comparable r.le.patch module can calculate only the first five (some indirectly). Presumably the Euclidean nearest neighbor distance distribution could be calculated by one of the r.le programs called r.le.dist. But for unknown reasons, it is not available in GRASS version 6.3. For those metrics both FRAGSTATS and r.le.patch can calculate, though may show in different units, the results are equivalent. Figure 3 shows two resulting maps of this test.

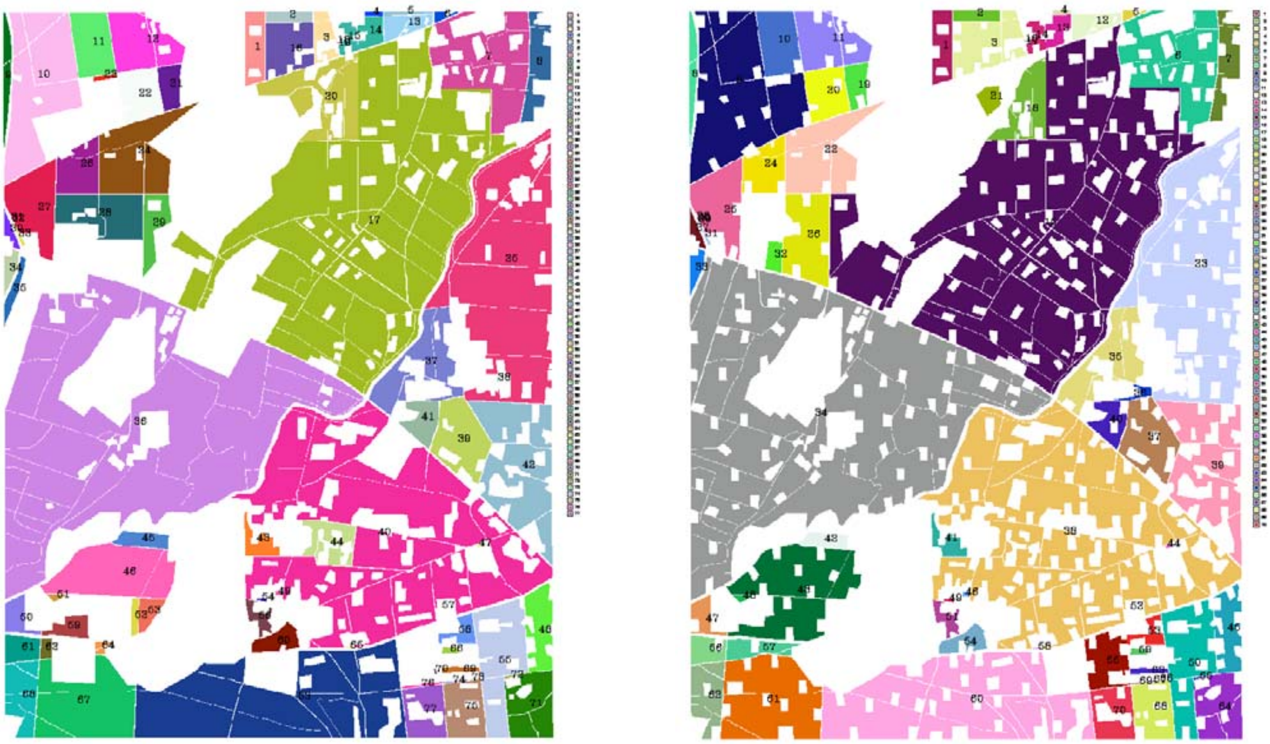

Figure 3. Left: individual patches identified in the clustered farmhouse development scenario. Right: individual patches identified in the scattered farmhouse development scenario.

For the green corridor test, FRAGSTATS can calculate all directly, but the comparable r.le.patch module can calculate only the first six (some indirectly). According to Baker (2001), the perimeter-area fractal dimension that was available in previous versions of the r.le programs are replaced with the supposedly more appropriate twist number statistics. For those metrics both FRAGSTATS and r.le.patch can calculate, though may show in different units, the results are equivalent. Figure 4 shows two resulting maps of this test.

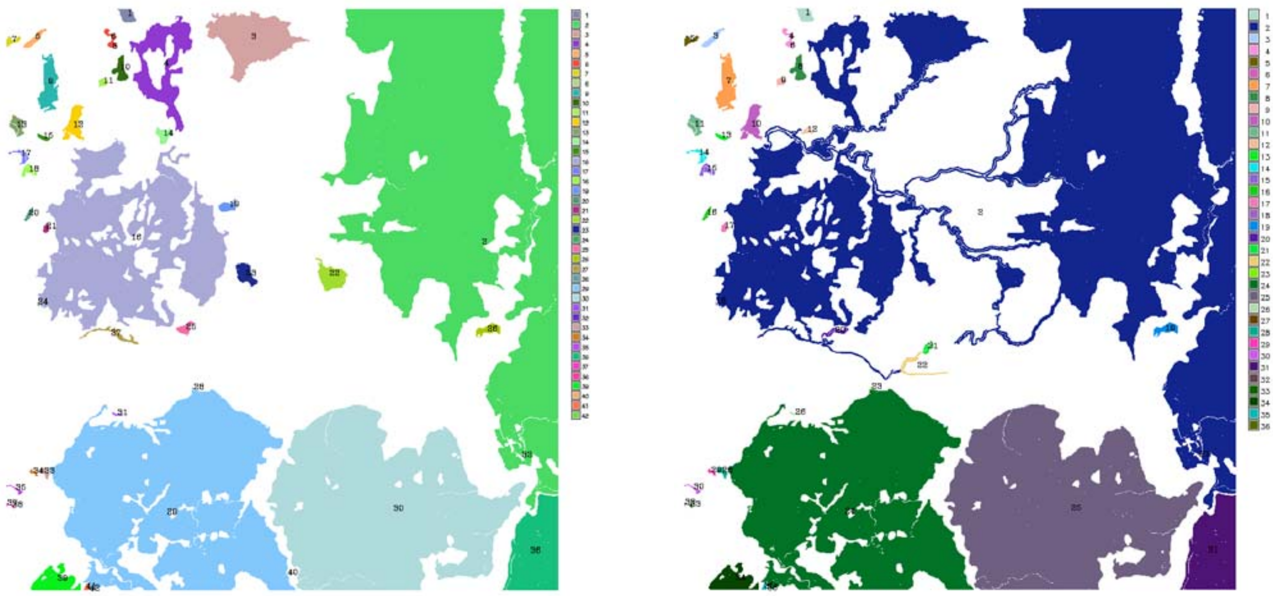

Figure 4. Left: individual patches identified in the areas currently covered by woodland. Right: individual patches identified in the scenario with proposed green corridors along streams.

4. 4 Discussion

Since GRASS 6.3, there appears the r.li suite of GRASS raster commands that claim "[it] aims at replacing the r.le suite of modules through a client-server, multiprocess implementation." (GRASS Development Team, 2007b) Among 152 raster commands available in GRASS 6.3, the r.li group takes up 17 commands, while the second and third largest groups, r.out and r.in, have only 16 and 11 commands respectively. This makes the r.li group become the most noticeable commands in the raster-command part of the on-line reference manual.

However, a quick Google Scholar (Google Scholar, 2008) returns not even a single publication, which means up until now the r.li solution has never been utilized by its intended landscape ecology community in the English speaking world at all. Judging from the fact that none published academic work have actually used the r.li suite let us doubt its importance ever matches its visibility given in the GRASS online reference manual. Initial tests conducted by the authors reveal that the most basic setting of assigning the whole map layer as the sampling area in the r.li doest not even work! This is why it was ruled out as one of the viable solutions in this study at the beginning.

From purely programming point of view, the internal structure, coding style, and execution efficiency of r.li might be much better than those of r.le, which were developed 15 years earlier. However, if a program is not well received, thoroughly tested, regularly maintained, and continuously improved by both users and developers in the target community, the value of such a program's existence is in question. In spite of this, nowadays devoting both time and resources to "reinvent the wheel" seems quite common in the open-source world.

Before FRAGSTATS version 3, one major advantage of the r.le programs over FRAGSTATS is its capability to perform the moving window analysis. This makes conducting multi-scale landscape pattern analyses a more feasible option in r.le than in FRAGSTATS version 2. Since version 3, however, FRAGSTATS has almost all the functionality of r.le and much more in terms of how many different metrics that can be calculated. In addition, FRAGSTATS is a single program with well-designed GUI but r.le programs divide functionality into five modules. Form end users' point of view, for the landscape metrics calculation per se currently FRAGSTATS is a far more superior option than r.le, if they do not mind FRAGSTATS' loose-coupling with GIS.

Some people may be concerned with relying exclusively on FRAGSTATS simply because it is "written in Microsoft Visual C++ for use in the Windows Operating environment" only (McGarigal, 2006). With the help of an open-source x86 virtualization product called VirtualBox, such a concern might be alleviated to some extent (Sun Microsystems, 2008). In fact, all calculation involving A+F solution in this study is executed on a virtual machine with Windows XP guest OS through VirtualBox.

Given FRAGSTATS is free, working well alone and with other GIS, and superior to r.le (and r.li of course), what should GRASS developers do next if they have the ambition and resources to further develop landscape metrics calculation solution in GRASS? With the improved vector processing capability and to avoid "reinventing the wheel," we would suggest to developing a vector-based landscape metrics calculation solution. There exists the Patch Analyst, which is an extension to the ArcGIS GIS system, that intends to facilitate the spatial analysis of landscape patches, and modeling of attributes associated with patches (Rempel, 2008). The Patch Analyst is an example showing there is a niche market for such a vector-based solution. Since Patch Analyst is an extension to the proprietary ArcGIS GIS system, an open-source landscape metrics calculation solution that is well-integrated inside an open-source GIS should be of interest to those who can't afford or would not be willing to use proprietary GIS.

5. Conclusion

5. 1 Summary

In the three parts of the case study, both solutions seem to have its own share of pros and cons. Not surprisingly, the most significant disadvantage of the A+F solution is its capital cost. Otherwise, the proprietary but free FRAGSTATS itself currently turns out to be a more capable, mature, and updated program than r.le. However, with the help of VirtualBox, FRAGSTATS can work with virtually all GIS. Furthermore, if users are willing to endure slight inconvenience in exchange of enormous monetary saving, linking FRAGSTATS with GRASS through a common file format works as well as the aforementioned two solutions. This makes devoting more efforts to update the veteran r.le or replace it with a newer one do not seem to be necessary.

5. 2 Further research

With the vastly improved vector-handling capability of GRASS 6, it might be worthwhile to develop a vector-based landscape metrics calculation solution that is integrated inside a well-developed open-source GIS in order to differentiates itself from the already excellent raster-based FRAGSTATS.

Given the experience that a rarely used and seemingly immature group of commands like r.li that floods the already over-crowed and confusing raster commands of GRASS 6, the GRASS development team should seriously consider conducting a thorough survey of how those hundreds of commands are actually used by users before they consolidate those commands into a more system manner in the GRASS 7 or beyond.

References

Baker, WL (University of Wyoming), 2001, *The r.le Programs: A Set of GRASS Programs for the Quantitative Analysis of Landscape Structure*, GRASS GIS 6.3.svn Reference Manual: raster: r.le.setup, viewed 16 August 2008, <http://grass.itc.it/grass63/manuals/html63_user/r.le.setup.html>.

Baker, WL and Cai Y, 1992, 'The r.le programs for multiscale analysis of landscape structure using the GRASS geographical information system,' *Landscape Ecology*, vol. 7, no. 4, pp. 291-302.

Forman, RTT, 1995, *Land Mosiacs: The Ecology of Landscapes and Regions*, Cambridge University Press, New York, New York.

Forman, RTT and Godron M, 1986, *Landscape Ecology*, John Wiley & Sons, New York, New York.

FOSS4G 2008, 2008, Open Source Geospatial Foundation, viewed 16 August 2008, <http://conference.osgeo.org/index.php/foss4g/2008>.

Google Scholar (BETA), 2008, *“r.li” - Google Scholar*, viewed 16 August 2008, <http://scholar.google.com/scholar?q=%22r.li%22&hl=en&lr=&btnG=Search>.

GRASS Development Team, 2007a, *GRASS GIS: GRASS 6 Sample Data Download*, GRASS GIS: The World Leading Free Software GIS, viewed 16 August 2008, <http://grass.itc.it/download/data6.php>.

GRASS Development Team, 2007b, *GRASS 6.3 Users Manual*, GRASS GIS: The World Leading Free Software GIS, viewed 16 August 2008, <http://grass.osgeo.org/grass63/manuals/html63_user/>.

Lin, YY and Wang WCJ, 2008, ‘Applying landscape ecology theories to assess the development of clustered farmhouses: A case study of Wantan township, Pingtung County,' *Proceedings of the Biannual Architecture Research Symposium 2008-1*, Architectural Institute of the Republic of China, <http://www.airoc.org.tw>.

McGarigal, K, 2006, *FRAGSTATS Home Page*, University of Massachusetts, Amherst, Massachusetts, viewed 16 August 2008, <http://www.umass.edu/landeco/research/fragstats/fragstats.html>.

McGarigal, K and Marks BJ, 1995, *FRAGSTATS: Spatial Pattern Analysis Program for Quantifying Landscape Structure*, USDA Forest Service General Technical Report PNW-GTR-351.

McGarigal, K, Cushman SA, Neel MC, and Ene E, 2002, *FRAGSTATS: Spatial Pattern Analysis Program for Categorical Maps*, Computer software program produced by the authors at the University of Massachusetts, Amherst, Available at <http://www.umass.edu/landeco/research/fragstats/fragstats.html>.

Rempel, R, 2008, *Patch Analyst Home Page*, Centre for Northern Forest Ecosystem Research (Ontario Ministry of Natural Resources), Lakehead University Campus, Thunder Bay, Ontario, viewed 16 August 2008, <http://flash.lakeheadu.ca/~rrempel/patch/index.html>.

Sanderson, J and Harris LD (eds.), 2000, *Landscape Ecology: A Top-Down Approach*, Lewis Publishers, Boca Raton, Florida.

Sun Microsystems, Inc, 2008, VirtualBox.org Home Page, viewed 16 August 2008, <http://www.virtualbox.org>.

Turner, MG and Gardner RH (eds.), 1991, *Quantitative Methods in Landscape Ecology: The Analysis and Interpretation of Landscape Heterogeneity*, Springer-Verlag, New York, New Year.

Wikipedia contributors, 2008a, *Ecology*, Wikipedia: The Free Encyclopedia, viewed 16 August 2008, <http://en.wikipedia.org/w/index.php?title=Ecology&oldid=231331408>.

Wikipedia contributors, 2008b, *Landscape Ecology*, Wikipedia: The Free Encyclopedia, viewed 16 August 2008, <http://en.wikipedia.org/w/index.php?title=Landscape_ecology&oldid=231714340>.

Wikipedia contributors, 2008c, *Case Study*, Wikipedia: The Free Encyclopedia, viewed 16 August 2008, <http://en.wikipedia.org/w/index.php?title=Case_study&oldid=225313829>.

Yang, JL and Wang WCJ, 2008, ‘Analyzing the effects of ecological green networks on landscape patterns of a rural area,' *Proceedings of the Biannual Architecture Research Symposium 2008-1*, Architectural Institute of the Republic of China, <http://www.airoc.org.tw>.