G172/172L Introduction to Geographical Data Analysis

Final Project

A comparison of subjective and objective measures of mapped landscapes using GIS, MDS, and correlation analysis.

By Marco Ruocco

Summary

A set of 10 DEMs has been assembled and analyzed using GIS software to extract statistical measures for elevation and slope. A pairwise subjective comparison of the DEMs represented in the form of isopleth maps allowed the generation of a 3D configuration of dissimilarity using Multi Dimensional Scaling software. A correlation between the location of the maps on the three dimensions of the MDS plot and the respective statistical measures of physical characteristics have been computed and tested for significance. The results show only Slope Mean significant against the first dimension of the plot, suggesting difficult interpretation of part of the MDS output and possibly the low explanatory power of the statistical measures in interpreting the subjective evaluations. The project has been useful in suggesting new ways to approach the subjective-objective landscape evaluation and analysis problem, namely extending the statistical analysis to other measures, including fractals, and improving the experiment design and the statistical techniques used.

Introduction

The perception of the physical landscape can be studied from a variety of perspectives. The approach used here is to consider the experience of landscape as a combination of physical and perceptual aspects that are related one another. By physical aspects I mean the characteristics of the terrain considered as a surface with statistical properties: elevation and slope are the variables that are considered here in terms of average value and standard deviation. The perceptual aspects are instead concerned with the subjective nature of the evaluation of landscape appearance when represented by a map. The interest is placed on the outcomes of a quick evaluation of similarity between maps. The objective of the study is to compare physical objective measures with perceptual subjective measures, and find if there is a correlation between, say, the mean value of altitude and an extracted axis of the perceptual evaluation of a given map.

The objective of this study is to practice with methods of analysis, and therefore not one of arriving at any conclusive statement concerning landscape evaluation. The procedures actually used are just one of the many possible available and not strictly the most suitable, but they have been helpful in familiarizing with tools and reasoning about the problem.

The data sources for this minimal "experiment" are a set of USGS DEMs, which provide the landscape maps on which the physical and perceptual analyses are based. ArcView GIS provided the tools for the simple terrain analysis (extraction of statistical descriptors) carried out on the maps. The subjective measures are instead originated from the evaluations of a single subject (myself). Those have been input into ALSCAL Multidimensional Scaling (MDS) software that produced a three dimensional configuration. The final correlation analysis was aimed at determining if there was a correlation between the position of the map along the dimensions of the MDS output and the respective statistical characteristics of the map.

DEM elaboration

The source of the 10 1-degree DEMs used in the study was a USGS repository located at http://edc.usgs.gov/glis/hyper/guide/guide/1_dgrdemfig/index1m.html. They were distributed over the entire territory of Arizona and together they covered more than half of the state surface (see figure 1 for the entire dataset and the attached map for the selection used in this study).

After importing the DEMs in the ESRI ArcView GIS environment, I extracted the statistical characteristics of mean elevation, elevation standard deviation, mean slope, and slope standard deviation using the built-in tools (the statistics were already available with the summary information of the coverage). Elevation and slope MIN-MAX ranges were also recorded but later were replaced by the standard deviations as more representative measures of dispersion. For slope measures, the software computed slope maps from elevation maps, and then produced statistical values using the same methods as those for elevation maps.

The actual interpretation of the physical measures in defining the landscape depicted by the DEM maps is not straightforward but nonetheless some principles can be defined, as presented in Hoffman & Pike (1995). Detailed explanations and the diagrams for interpretation are available in that paper. Mean Elevation and Mean Slope were outside the analyses of Hoffman & Pike (1995), but I decided to include them here for experimentation's sake. It will suffice to say that Mean Elevation is a central tendency measure that suggests a grouping of landscapes in high, medium and low elevation, influencing the actual landforms present (a high altitude environment in general looks different than a low altitude one). Elevation Standard Deviation is a surrogate of the roughness of the landscape (high standard deviation means high dispersion, that is, extensive areas with low altitude and extensive areas with high altitude, indicating an unbalance that is visible on the terrain in the form of alternating extreme valleys and peaks). Mean Slope indicates the average characteristics of the landscape in terms of elevation drop rate (a high Mean Slope means in general a dominance of steep slopes, while low Mean Slope suggests low or rolling terrain). Slope Standard Deviation resembles the dispersion measure of Elevation Standard Deviation, in that a high Standard Deviation means dispersion (a wide spread of values across low, medium or high slopes, resulting in rough, irregular landscapes) while low Standard Deviation indicates a landscape that is either dominated by a compact range of slopes, whether high, medium or low (resulting in more regular landscapes)

The maps were displayed as isopleths with elevation tints using the standard coloring of dark (low altitude) to bright (high altitude), as seen on the attached sheet. I produced a summary layout of all maps for quick reference (see attached sheet), but all the subsequent subjective evaluations were carried out on screen, with two separate windows open on the two maps being compared.

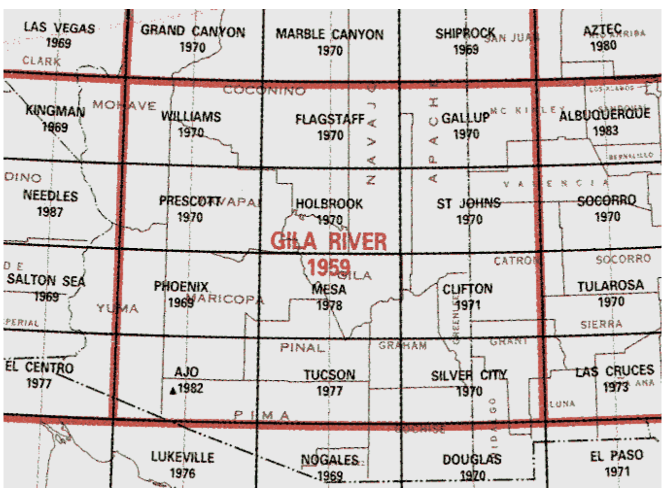


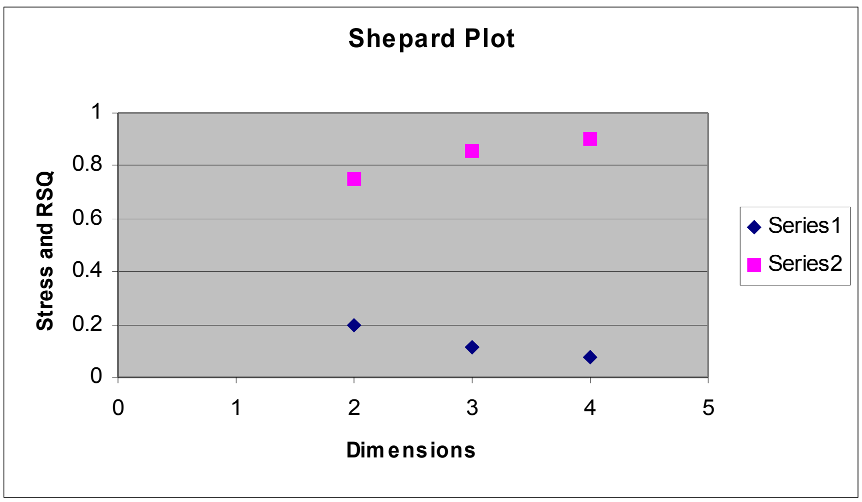
Figure 1 - Source area of USGS DEMs

Subjective measures

The 10 DEMs have then been compared in a pairwise fashion. For each of 45 pairs of DEMs I determined a subjective rating of dissimilarity from 0 (identical) to 9 (very dissimilar), with the number 0 used only for the case of comparison between the same DEMs. The data matrix is visible in the MDS printout. Dissimilarity was a result of a subjective judgment following the exposure for ten seconds to the pair of maps. The assumption is that a range of judgmental processes can take place in those ten seconds and no particular criterion was imposed to guide that judgment. The objective of the MDS analysis was exactly the one of extracting the judgmental processes underlying the scoring of dissimilarity.

MDS analysis

The subjective dissimilarity matrix has been input to ALSCAL (an MDS software) after setting the parameters for the analysis in the ALSCAL header file. The details are reported on the attached sheets. The analysis, considering the limited number (10) of observations, could be carried out up to a maximum of 4 dimensions. The following Shepard plot shows the pattern of decrease of stress values in the configuration with an increasing number of dimensions (Series 1), and a correspondent decrease of RSQ values (Series 2). The breakpoint, even if difficult to identify, seemed located at dimension 3, and therefore the chosen number of dimension for the analysis is 3.



The MDS computation resulted in the plots for three dimensions presented in the attached sheets. The interpretation of the dimensions is complex but a pattern emerges. Dimension 1 presents a clear succession of maps from the most regular and with limited variation in form (H Saint Johns, J Williams) through those moderately varied with distinct alternation of valleys and peaks (B Flagstaff, I Tucson) and finally to the most extreme examples of very rugged terrain with many narrow valleys and peaks for an overall complex texture (A Clifton, G Prescott). Dimension 2 is more complicated: it seems to follow roughly the pattern of dimension 1, but it seems more concerned with form than with variability. In other words it displays a range of maps from those with elongated fluvial features (I Tucson, C Gallup) to those showing landscapes less demarcated by distinct valleys (H Saint Johns, A Clifton). Notice that A Clifton is clearly a fluvial landscape, in contrast to what just said, but the valleys are not as wide and clearly distinguished as for I Tucson. Dimension 3 is more compact on the plot, and the least interpretable. The criteria that separate the extreme landscapes of this dimension don’t seem to comprise those of Dimensions 1 or 2 (G Prescott and A Clifton, previously considered similar, are here nearly opposite). No simple interpretation seems possible at this stage.

Correlation analysis

The next step has been to compare the physical measures of altitude and slope with the subjective measures as expressed by the MDS 3D plots. A correlation analysis was set up to compare, for each dimension considered separately, the coordinates of the maps on the MDS plot, and the respective physical measures. A total of 12 correlations were computed (three dimensions by four physical measures). The Pearson product moment coefficient measure was chosen as for the test statistics (S-PLUS command: r <- cor (X,Y)). The correlation matrix below reports the obtained values.

	Elevation Mean	Elev. Std Dev.	Slope Mean	Slope Std Dev.
Dimension 1	0.265	0.437	0.758	0.562
Dimension 2	0.335	-0.234	-0.220	-0.426
Dimension 3	0.114	0.382	0.289	0.228

The correlation coefficients were tested for significance (S-PLUS command: x <- cor.test (x, y, alternative = “two.sided”, method = “pearson”)). The table reports the results:

	Elevation Mean	Elev. Std Dev.	Slope Mean	Slope Std Dev.
Dimension 1	t = 0.7783 df=8 p-value=0.4588	t = 1.3748 df=8 p-value=0.2065	t = 3.2885 df=8 p-value=0.011	t = 1.9208 df=8 p-value=0.091
Dimension 2	t = 1.0053 df=8 p-value=0.3442	t = -0.6816df=8 p-value=0.5147	t =-0.6406 df=8 p-value=0.5397	t =-1.3316 df=8 p-value=0.2197
Dimension 3	t =0.3234 df=8 p-value=0.7547	t =1.1679 df=8 p-value=0.2765	t =0.8531 df=8 p-value=0.4184	t =0.6624 df=8 p-value=0.5263

With a significance level of ◎=0.05 (two tailed test), ◎ > p-value only in the case of Slope Mean in Dimension 1. In all other cases the correlation coefficient r is not significantly different from zero.

Final results

The tests of significance show poor correlation between the physical measures and the subjective judgment, suggesting that the problem of connecting these two elements cannot be approached using these simple comparisons. On the other hand the significance of Slope Mean in dimension 1 suggests that the progression of landscape characteristics as extracted from the MDS can be explained in terms of increasing Slope Mean from regular terrain to very rugged terrain; even if not attaining a significant level, Slope Standard Deviation shows an identical pattern. Dimension 2, considering the structural interpretation given above of macroscopic fluvial versus non-macroscopic fluvial landscapes, is hardly captured by the four measures, which function as surrogates that individually cannot attain significance levels. Dimension 3, hardest to interpret, is also removed from the explanation offered by the four measures. The four measures are therefore either not representative of the actual appearance of the landscape or have not been used directly in the judgment process. As far as the first point is concerned, Elevation Mean is a loose measure of landscape characteristics that can capture only broad patterns, while it is surprising that Elevation Standard Deviation scored lower than Slope while should be able to represent the first order variability of the landscape.

Conclusions

The project was aimed only at practicing with the several software tools for statistical and terrain analysis (ArcView, ALSCAL, S-Plus) and having a first exposure to the task of comparing subjective with objective measures of the landscape. The process of conducting this project has been helpful in identifying possible improvements in the design of the experiment and new possibilities for research. The following aspects in particular emerge from the experience:

- The interpretation of statistical measures (Elevation Mean, Elevation Standard Deviation, etc.) applied to the landscape require a thorough interpretation that goes beyond the tentative criteria used for carrying out this mock project. Also other measures, like Skewness, Kurtosis, and the consideration of Curvature, can be included in the analysis and more powerful descriptors can be assembled to describe statistically a landscape. An explorative device would be helpful in generating surfaces with varying characteristics, so that we can consider how the different parameters used in varying quantities interact to characterize real world landscapes. An interesting direction of research would be to transcend the statistical measures and explore the contribution of more complex measures like fractals in extracting and representing the characteristics of a landscape. It all contributes to an exciting field of terrain analysis closely tied with simulation of landscape forms.
- During the evaluation a greater degree of control on the type of exposure to the map is advisable. The exposure to a map can be substituted by an exposure to a 3D representation of the landscape, for testing the appearance of the landscape and not the appearance of a map indirectly representing a landscape.
- The statistical analysis results in a sequencing of tools (from MDS to correlation) that might decrease the explanatory power available from accessing directly the original data. Also, other statistical techniques might be more helpful in these respects than a simple correlation carried out directly on MDS output of internal coordinates.

In conclusion this project identifies a field of inquiry and a methodological approach to the subjective-objective distinction even if not strictly a set of methods of investigation. Certainly it would be interesting to proceed in the several directions emerging from the project: 1) landscape physical analysis using statistics, simulation and fractal theory 2) techniques of subjective measures analysis like MDS and 3) consideration of alternative views on landscapes other than maps, such as 3D reconstructions.

References

Hoffman, R.R., Pike, R.J. (1995) "*On the Specification of the Information Available for the Perception and Description of the Natural Terrain*". In **P.A. Hancock, J. Flach, J.K. Caird, & K. Vicente** (Eds.) *Local applications of the ecological approach to human machine systems* (Vol. 2, pp 285-323) Hillsdale, NJ: Lawrence Erlbaum Associates.