# Satellite Image Restoration using the VMCA Model

Sharolyn Anderson* Mark A. Fonstad† Claudio Delrieux‡

**Abstract**

One of the most common patterns of the geographic landscape is the fractal or near-fractal form. Unfortunately, most traditional methods of spatial interpolation assume some type of continuous and regionalizeable variation of the underlying geographic form, an assumption at odds with the observed fractal properties of many landscapes. An extremely simple iterative algorithm, the voter model cellular automata (CA), produces discontinuous fractal patterns useful for interpolation while at the same preserving a realistic amount of spatial autocorrelation, extracted from neighboring existing data, also found in these landscapes. This adaptive algorithm is based on the principle of iteratively interpolating a missing data point using the value of a randomly selected neighbor cell. The model can also be extended to interpolate field-like variables by adding random deviations from the randomly chosen neighbor cell value. In this paper we explore the effect of satellite image restoration using a simple VMCA over obscured by clouds areas. This model is computationally advantageous, given its localty and restricted underlying computational model. Thus, an adequate computer implementation may perform significantly faster than other restoration methods, with roughly similar overall results. Also the local/scalable/parallelizable nature of CAs allows hardware FPGA implementation that might be embedded within the imager devices in satellites and remote sensors. On the other end, a GPU implementation might take advantage of highly specialized parallel processors capablde of restoring huge images in real time.

## 1 Introduction

Interpolation of geographic information is an important part of geographic analysis. Interpolation can be needed to fill obscured areas, to fill a spatial continuum between known point values, or to downscale data. Many landscapes exhibit a near-fractal spatial structure, and users wishing to interpolate missing data in such landscapes commonly face poor interpolation results. We wish to propose an iterative voter model based on cellular automata theory that can produces a near-fractal interpolation surface for just such landscapes, as well as a more continuous surface in less fractal landscapes. This model removes pixel values classed as clouds or other obscured features and replaces them with a "best guess" of what should actually exist in that location based on the neighboring, unobscured pixel values. Such a removal would be useful not only in visualization but also in time series analysis and other statistical analysis; interpolation error due to the method —if reasonable— can be added to the normal errors of classification rather than just omitting scenes that have obscured areas.

*Department of Geography - University of Denver - Denver, CO 80210 - Sander24@du.edu

†Department of Geography - Texas State University at San Marcos - San Marcos, TX 78666 - mfonstad@txstate.edu

‡Departamento de Ing. Eléctrica y Computadoras - Universidad Nacional del Sur, Av. Alem 1253 - (8000) Bahía Blanca, ARGENTINA - claudio@acm.org

One difficulty in producing a suitable interpolation method for most GIScience needs is that all interpolation procedures assume some conditions about the nature of the landscape they interpolate. For example, Kriging assumes that the variables interpolated are regionalizable and that the local autocorrelation is based on a continuum change structure. While many field variables (such as upper-level air temperatures) follow this assumption extremely well, many landscapes are not well approximated by the continuum assumption. Instead, many landscape elements follow fractal or near-fractal behavior patterns in their spatial scaling (Lam 2004). In cases with such a near-fractal structure, a non-continuum basis for interpolation is more appropriate.

Bak et al.'s (1987) sandpile (BTW) cellular automata is a simple yet elegant example of a discrete system with local interactions but does not have an assumption of a spatial continuum. Instead, changes from one site to the next follow a discrete, rule-based system that can have noncontinuous jumps. The BTW cellular automata can be based on deterministic or stochastic rules, but must be solved in an iterative fashion. Recently, landscape patterns of vegetation have been simulated using a variant of the BTW model known as the voter model (Bolliger et al. 2003). Sprott (2004) has extended this voter model cellular automata to digital image enhancement problems, specifically for interpolating missing data pixels.

The purpose of this paper is to assess the validity of the voter model cellular automata to the interpolation of typical geographic information such as classified or ratio-level remotely sensed imagery. Future improvements to this basic model include multiscale kernels for better control of the spatial autocorrelation of the interpolated pixels, multidimensional CAs in thematic images, adaptive CAs rules, and better validation techniques.

## 2 The Voter Model

The original voter model is an extremely simple, archetypal cellular automata. The model consists of an impressionable political electorate, where voters are represented as pixels in a grid (Holley and Liggett 1975, Sprott 2004). The pixels are initially assigned (randomly) one of two parties that they will vote for. However, during the next instance of time, these square voters are influenced by (i.e. take on the party of) a random one of their eight neighbors. After many iterated time steps the pattern that results is not a random hodgepodge of voters of each party, but rather a highly complex fractal spatial pattern, with small scale-dependent clumps of like-minded voters that is highly reminiscent of real party vote landscapes. Bolliger et al. (2003) have used the voter model idea to produce realistic near-fractal patterns of presettlement vegetation distributions in the absence of available presettlement vegetation data. An alternative construction of the voter model used the rule that each pixel would take on the party of the majority of neighbors, with a random coin flip determining the outcome of a tie. As with any cellular automata, a large number of model different transition rules will provide slightly different results, but with many structural similarities in the emergent patterns through time.

Sprott (2004) modified the original voter model CA for application to image enhancement, particularly interpolation of missing data pixels in an image. Unlike the original voter model, Sprott's rules stipulated that pixels that are initially fine (i.e. not missing or corrupted) should not be subject to alteration by neighboring pixels. This requires recoding the image to explicitly note which cells should be interpolated and which cells are to be left unchanged. During each CA time step, the cells are updated asynchronously (randomly or one at a time),

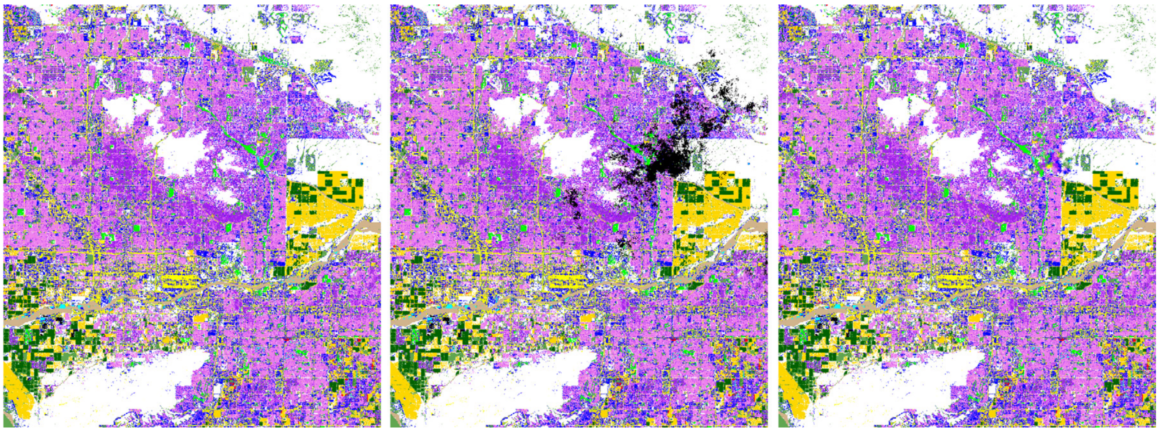

Figure 1: (a) Classified urban area (Phoenix, AZ Landsat image), (b) added clouds, and (c) interpolated with VMCA model.

although the more traditional synchronous update should produce almost exactly the same interpolation patterns.

In images with a low number of unranked classes (such as a classified remotely-sensed landcover map), the original voter model CA is immediately applicable, with only minor modification. Sprott's (2004) voter model did not explicitly deal with the possibility of missing data on the image boundaries, and the chosen neighborhood for that model was the 8-neighbor Moore neighborhood. In our application of the voter model to geographical data interpolation, we use an absorbing boundary condition, where the values along the edge of the image do not change their value during the iterative interpolation procedure. Also, we broaden the original Sprott approach to allow inclusion of the Von Neumann (4-cell rook's case) or Moore (8-cell queen's case) neighborhoods.

Sprott applied the original procedure to both black and white imagery as well as 256-level grayscale images. However, a more realistic approach to ordered/ranked data with a large number of classes or with ratio-level data is the modification of the updated pixel value during each time step by allowing a small amount of deviation from selected neighbor value. The deviation is simply selected as a random number from within preset bounds whose mean is equal to zero, indicating no deviation. Allowing deviations of this sort should increase the realism of ratio-level image data by producing more subtle variations between pixels of known value.

## 3 Interpolation Tests

The first test of the model to interpolation of missing data is the "removal" of added synthetic clouds from two classified 30-meter pixel Landsat scenes, one of urban Phoenix, Arizona and one of the Raton, New Mexico region. The purpose of testing the algorithm in these two areas is to assess its validity in areas of vary different spatial heterogeneity. We have used previously classified images that were produced to achieve high classification accuracy. The Phoenix image (Fig. 1(a)) was produced as part of the CAP-LTER program in central Arizona, and the New Mexico map (Fig. 2(a)) was produced by the USGS NALC program. Both of these programs worked to produce maps that were cloud-free. In order to have both

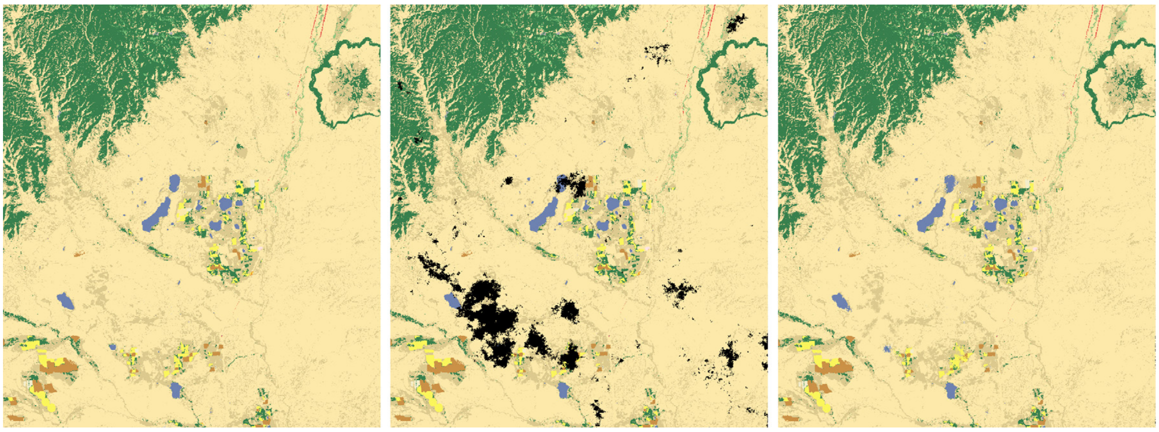

Figure 2: (a) Classified rural area (Ratón, NM, Landsat image), (b) added clouds, and (c) interpolated with VMCA model.

high-quality "true" data as well as areas suitable for interpolation, we have added synthetic fractal clouds to these "cloud-free" scenes, using a fractal cloud generator and placing them into the areas of interest. This procedure produces both realistic missing data patterns and high-quality truth data. We chose clouds and scene areas that would cover approximately 4% of the clipped classified Landsat scenes, a typical proportion for many satellite images. The cloud generator allows for the generation of typical cloud shapes by means of calibrating the fractal dimension of the resulting cloud.

The second test of the (altered) voter model CA is on two 8-bit, unclassified grayscale aerial photos of portions of south-central Texas, USA. The source of these aerial images is standard digital orthophotoquads produced from NAPP program photos. Clouds covering approximately 4% of the images were added in the same manner as was used for the Landsat scenes. The first photo (Fig. 3(a)) is of a portion of a residential neighborhood in Austin, TX. At the 1-meter resolution of these images, there is strong surface heterogeneity between roads, housetops, grassy areas, cars, and other cover types. The other image (Fig. 4(a))is from an agricultural area east of the Austin metropolitan area. It has the same pixel resolution as the first aerial photo, but the spatial structure of the image is considerably more homogenous.

## 4 Discussion and Conclusion

The results generated so far are very promising. As shown in the images, the interpolation is quite good for visualization purposes. Furthermore, in all the studied cases, the average error of the interpolated area is close to 0% (thus confirming the stationary nature of the VMCA model), and the RMS error is below 25%, being quite lower in rural images and with clouds of high fractal dimension. Thus, the interpolated images are good enough for most statistical analysis purposes. Future improvements to this work include multiscale kernels for better control of the spatial autocorrelation, multidimensional CAs in thematic images, adaptive CAs rules, and better validation techniques.

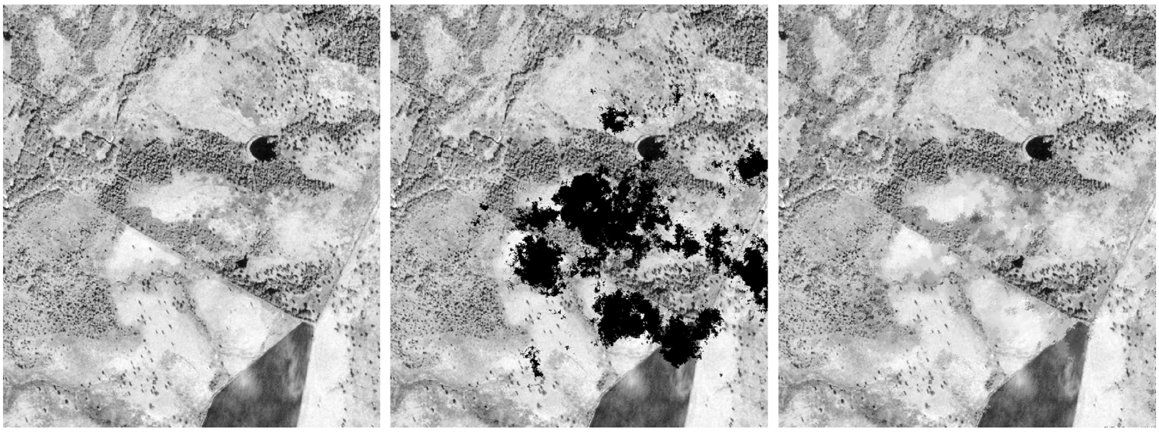

Figure 3: (a) Grayscale aerial urban image (Austin, TX, NAPP photo), (b) added clouds, and (c) interpolated with VMCA model.

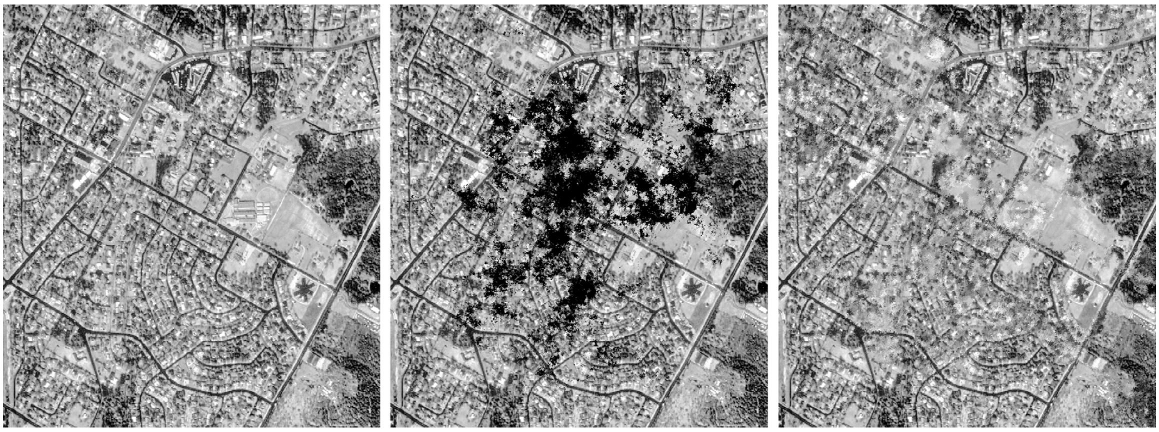

Figure 4: (a) Grayscale aerial rural image (agricultural area east of Austin, TX, NAPP photo), (b) added clouds, and (c) interpolated with VMCA model.

## References


[1] Bak, P., Tang, C. and Wiesenfeld, K., 1987. Self-organized criticality: an explanation of 1/f noise. Physical Review Letters, 59(4): 381-384.

[2] Bolliger, J. Sprott, J.C., and Mladenoff, D.J. 2003. Self-organized criticality and complexity in historical landscape patterns. Oikos 100, 541-553.

[3] Holley, R. and Liggett, T.M. 1975. Ergodic theorems for weakly interacting particle systems and the voter model. Annals of Probability 3, 643-663.

[4] Lam, N.S. 2004. Fractals and scale in environmental assessment and monitoring. In Sheppard, E. and McMaster, B. (eds), Scale & Geographic Inquiry: Nature, Society, and Method. Blackwell Publishing, Malden, 23-40.

[5] Sprott, J.C. 2004. A method for approximating missing data in spatial patterns. Computers and Graphics 28, 113-117.