Modelling stochastic structures in soil

G.W. Horgan[1] & I.M.Young[2]

It is well appreciated by environmental scientists and ecologists that the apparently uniform, appearance of an area of ground hides a rich complex structure of organic and inorganic material, and porosity, that is variable in its properties at almost every scale at which it can be examined. Statistics, which is the science of variability, can play an important role in modelling and understanding this variation. There are several ways in which a model can be applied. We may use it to *describe* the variability, and to compare aspects of it in different situations. We can use a model to *predict* what will happen when parameters of the model are varied. Comparing predictions with observations provides a *test* of the validity of the model. Perhaps the most satisfying use of a model is when it gives insight that helps to *understand* aspects of a system that cannot be revealed readily by observation alone.

The analysis of pores and cracks in soil provides good examples of the way in which stochastic models can complement experimental work. As the medium through which water, gases and micro-organisms move through soil, their importance cannot be overstated. The structure of the pore space is complex and, being three-dimensional, is difficult to observe. Descriptions in terms of fairly simple concepts such as pore size distributions and indices of tortuosity will be inadequate. Although all models of pore space will fall short of reality, more complex models will be more realistic and useful, but more demanding to work with. Fortunately, continuing improvements in the power of affordable computing allows us to be ever more ambitious in this regard.

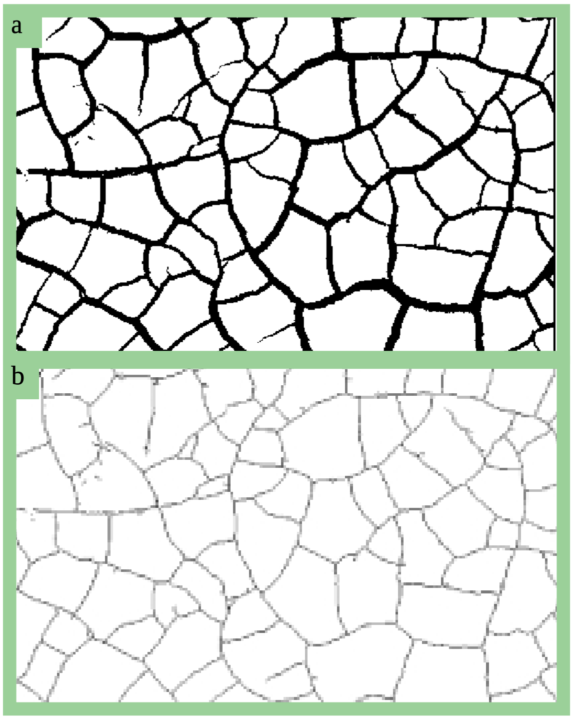

Figure 1 (a) Image of cracks in soil collected from Cruden Bay, sieved, and dried in a 10cm dish at 20°C. (b) Skeletonisation operation to reduce the cracks to single pixel thickness before estimating length density.

Image Analysis

Image Analysis is the extraction of information from digital pictures. It sits somewhere between Image Processing, which is the modification of a picture to produce another picture, and Image Understanding, which tries to build a complete representation of all the relevant content of an image - a type of Artificial Intelligence. In practice, these three disciplines overlap and interact.

We can illustrate image analysis by considering the image in Figure 1(a). It shows cracks in soil collected from Cruden bay, sieved and dried at 20°C. We are interested in quantifying aspects of the crack pattern. For example, the length density is relevant. Measuring the length requires an initial image processing step: we must reduce the cracks to single pixel thickness, a process termed skeletonisation. It is done by iteratively removing pixels, providing crack connectivity is unaffected, until no more can be removed. The result is shown in Figure 1(b). Some further tidying will be needed, either by removing unwanted artefacts from the skeleton, or by some initial smoothing of the image. Crack length can be measured by simply counting pixels, with allowance made for the number of pixels per unit length for lines at different angles. Assuming a lack of any directional preferences in crack growth, which seems reasonable, allows adjustment by a simple correcting factor.

The curvature of the cracks is thought possibly to be related to soil composition. This is trickier to measure. It requires more than looking at patterns of pixels and their neighbourhood. We need to trace each crack across the image as it passes many crack junctions.

[1] Biomathematics and Statistics Scotland, Rowett Research Institute, Aberdeen
[2] Scottish Informatics, Mathematics, Biology and Statistics Centre, University of Abertay Dundee

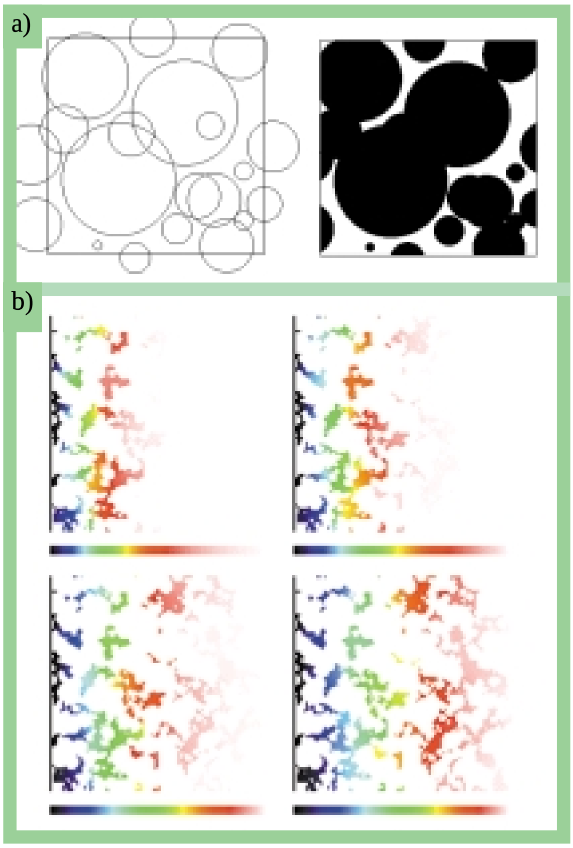

Figure 2 (a) Illustration of a two-dimensional slice of the Boolean model of soil pores: Spheres with a random radius are placed at random positions in space and allowed to overlap. The remaining spaces, shown here in white, are the pores. (b) Gas diffusion at four time points through a Boolean pore space. Gas diffuses from a source on the left. Intensity is shown by colours from white (low) through to black (high). Diffusion through an empty cylinder is shown underneath each image for comparison.

This was done by identifying segments of cracks (the lengths of crack between junctions) in the skeletonised image and deciding which segments of cracks were continuations of which other cracks. Segment length and angle were the main criteria used for this. Once whole cracks had been listed, some measure of curvature could be readily calculated.

Pore space diffusion

Although observation of the three-dimensional nature of soil pores is difficult, it is possible to construct models of the pore space that are simple to define, yet give rise to complex realisations. We have used the Boolean model for this purpose. It is defined by placing random sets at random positions in space and allowing them to overlap. Because of their simplicity, spheres with some distribution of radius are a common choice for the random sets. This is illustrated in Figure 2(a). The pore space is the complement of the spheres - the space that remains. We have found that an exponential distribution of sphere radius gives realisations which, in two-dimensional sections, resemble those of real pores.

We have used this model to explore the effect of model parameters (mean sphere radius, percent of pore space) on the properties of gas diffusion. Figure 2(b) shows the density of a tracer gas diffusing from left to right through a piece of soil, at four time points. Diffusion through an open cylinder is shown underneath each image for comparison. Properties of the diffusion process, such as steady-state flux, and the time to establishment of this steady-state, can be estimated. These properties can be compared with those obtained from experiment. We have found good agreement, which helps confirm the validity of the Boolean model. For example, steady-state flux appears approximately proportional to the square of the porosity. At very low porosities, no diffusion occurs. This can be understood in terms of percolation theory, whereby long-range connectivity of pores only occurs above some critical porosity threshold.

What is important for diffusion?

The modelling just described enables us to summarise and predict the effect of pore characteristics on diffusion. However, it does not tell us what properties of the pore space are important for diffusion, or what their influence is. Why does the diffusivity vary for

Geometric measure	Effect on flux	Effect on delay
1. p: the porosity		
2. p_l: left boundary porosity		
3. p_r: right boundary porosity		+
4. L_l: mean path length from left		
5. L_A: mean of all path lengths		+
6. p_p: proportion of pixels on paths		−
7. p_5: area within 5 pixels of paths	+	+
8: p_{10}: area within 10 pixels of paths		
9. W_G: Geometric mean path width.	+	
10. W_N: mean path bottleneck width		
11. D_1: Area of deadend pixels		−
12. D_2: Square root of deadend area		+
13. D_M: Maximum deadend length		
14. C_1: Porosity near sample centre	+	+
15. C_2: Harmonic of measure 14.		

Table 1 Effect on diffusion of pore geometry.

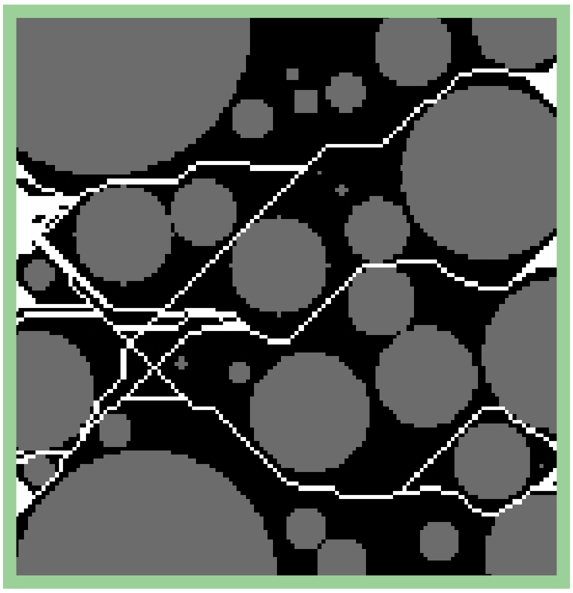

Figure 3 Paths through a simulated random pore space. A path is the shortest route from a point on the left to a point on the right of the pore space. It is the proportion of pore space near these paths, rather than their length, which has the greatest influence on the amount of steady state diffusion.

several samples with the same porosity? What geometric properties of the pore space can explain this?

We have investigated this issue by simulating diffusion in a number of random pore spaces, calculating a range of geometric characteristics, and looking at the associations between them. Two-dimensional models were used for simplicity, while permitting a substantial range of geometric structures. With gas diffusing from a source on one side of a sample, as described above, we were not surprised when principal component analysis of the diffusion curves revealed most variability to be accounted for by the amount of steady state flux, and a delay effect in the establishment of the steady state.

There are many ways in which the pore space geometry may be summarised. We calculated 15 summaries (see Table 1). Those based on path lengths are illustrated in Figure 3: a path is the shortest route from a point on the left to a point on the right of the pore space.

Multiple regression of the principal components of the diffusion curves on pore geometry measures showed that the main influence on steady state flux was the area within 5 pixels of a path. Diffusion will be highest when much of the pore space is concentrated near these paths. The main influence on delay is the proportion of pore space near the centre of the sample. The measures which contributed significantly to variability in diffusion, and the sign of their marginal effect, are shown in Table 3.

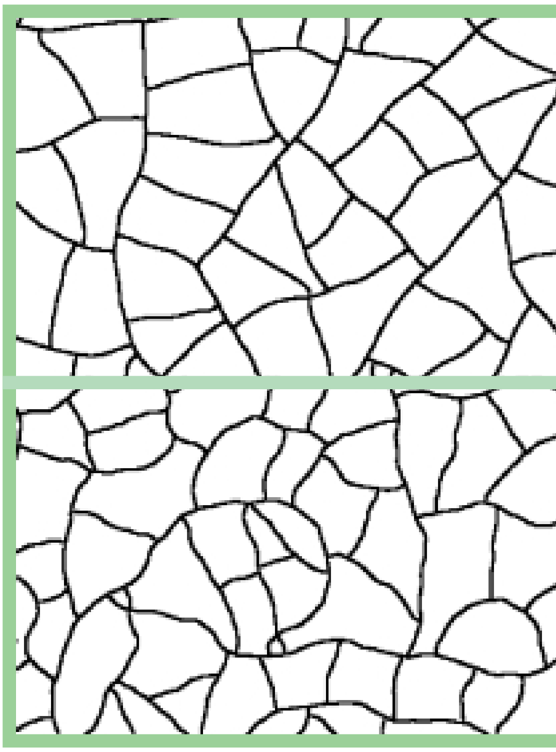

Figure 4 Simulated soil cracks. Crack direction changes more rapidly in the lower image.

The geometry of soil cracks

We saw earlier how image analysis allows us to measure features of soil crack geometry. A challenging question these images pose is to say what processes lead to these observed patterns. Modelling may be able to help understanding of processes that cannot easily be observed.

We found no models in the literature of stochastic geometry which in any way resembled real cracks. Those which had been proposed had sacrificed realism for mathematical tractability. After much scrutiny of images like Fig 1(a), and many attempts to devise algorithms which generated something resembling them, we found that five processes were necessary:

1. Random walk crack growth which continually makes small changes in direction.
2. New cracks forming on existing cracks and initially growing perpendicularly.
3. Crack attraction at small distances.

4. A tendency for large aggregates to split.
5. Stability below a certain aggregate size.

Each of these processes requires a parameter to describe it (process 3 needed two), leading to six parameters in total. Elegance is lost, but the result is realistic. Some realisations are shown in Figure 4. An important conclusion is that all five processes are necessary. For example, if process 4 is weak, the resulting simulations look nothing like cracks. The modelling has demonstrated the need for this process in understanding how cracks form.

It can be seen that the real cracks have even richer structure: varying thickness and a possibility of stopping before reaching other cracks. The model could be extended to incorporate this. Heterogeneity of crack patterns can also be seen in some crack images. This too could be handled by processes a step higher in a hierarchy. As we said at the start, soil structure is rich and complex at many scales.