Percolation Theory

PR King[1], SV Buldyrev[2], NV Dokholyan[2,4], S Havlin[5], E Lopez[2], G Paul[2], HE Stanley[2]

1 – Department of Earth Science & Engineering, Imperial College, SW7 2BP, London, UK
2 - Department of Engineering, Cambridge University, Cambridge, CB2 1PZ, UK
3 - Center for Polymer Studies, Boston University, Boston, MA 02215, USA
4 – Department of Chemistry & Chemical Biology, Harvard, Cambridge MA02138, USA
5 - Minerva Center & Department of Physics, Bar-Ilan University, Ramat Gan, Israel

Draft 27/5/02

Introduction:

Oil reservoirs are very complex having heterogeneities on all length scales from microns to tens of kilometres. These heterogeneities affect all aspects of the flow and have to be modelled to make reliable predictions of future performance. However, we have very few direct measurements of the flow properties. Core plugs directly measure the permeability but they represent a volume of roughly 10^{-13} of a typical reservoir. Well logs and well tests measure large volumes (10^{-4} and 10^{-7} respectively) but the results have to be interpreted to infer flow properties. The flow itself takes place on the scale of the pores which are typically around 10^{-21} of the volume of the reservoir. So there is a great deal of uncertainty about the spatial distribution of the heterogeneities which influence the flow.

The conventional approach to this is to build "detailed" reservoir models (note that the largest of these has around 10^7 grid blocks so they fall very short of the actual level heterogeneity that we know about), "upscale" or "coarse grain" them to around 10^4 or 10^5 grid blocks and then run flow simulations. These models need to be taken from a whole range of possible models with a suitable probability attached to each to determine the uncertainty in performance. The problem with this approach is that it is computationally very expensive. Therefore, there is a great incentive to produce much simpler models which can predict the uncertainty in performance. These models must be based on the dominant physics that control the displacement process.

It has long been understood that flow in heterogeneous porous media is largely controlled by continuity of permeability contrasts – flow barriers e.g. shales or high permeability streaks or faults. Although there are other influences these are the predominant features affecting flow. With this in mind we look to ways of modelling reservoir flow which concentrate on the connectivity of permeability contrasts. The basic mathematical model of connectivity is called percolation theory. Whilst there is a very extensive literature on percolation theory in both the mathematical and physics literature it is mostly not very accessible to the general geoscientist. The aim of this article is to attempt to redress this balance.

Percolation theory

Percolation theory is a general mathematical theory of connectivity and transport in geometrically complex systems. The remarkable thing is that many results can often be encapsulated in a small number of simple algebraic relationships. First let us describe what percolation theory is. There are many different variants which turn out to be identical in almost all important aspects so we shall describe the simplest version.

Take a square grid and occupy sites on this grid with a probability p. For small values of this probability we see mostly isolated occupied sites with occasional pairs of neighbouring sites that are both occupied. If neighbouring sites are both occupied we call it a cluster. As the occupancy probability increases we get more isolated clusters, some clusters grow and some merge. So the clusters on the whole get larger (see Figs). Then at a particular value of the occupancy probability one cluster dominates and becomes (infinitely) large. Above this the other clusters become absorbed into this one until at $p=1$ every site is occupied.

Series of Figures. ->

Note that a very peculiar thing happens at one particular value of the occupancy probability. Suddenly one cluster becomes infinitely large (for these purposes we are discussing infinitely big lattices, we shall discuss what happens on finite size lattices later). This is called the spanning cluster as it spans the entire lattice. This sudden onset of a spanning cluster occurs at a particular value of the occupancy probability known as the *percolation threshold* (p_c) and is the fundamental characteristic of percolation theory. The exact value of the threshold depends on which kind of grid is used and strongly on the dimensionality of the grid. A table of values is given below.

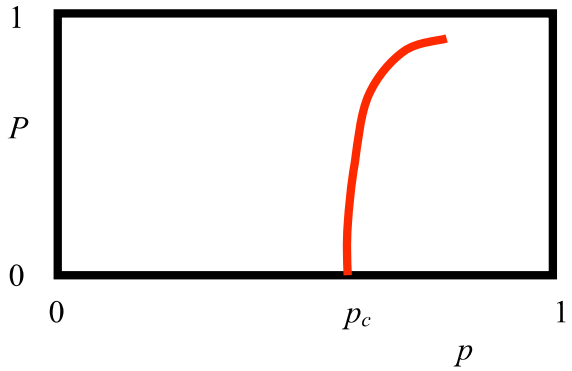
Lattice	Site	Bond
Hexagonal	0.692	0.6527
Square	0.592746	0.5
Triangular	0.5	0.34729
Diamond	0.43	0.388
Cubic	0.3116	0.2488
BCC	0.246	0.1803
FCC	0.198	0.119

Note that we have described the connectivity of sites on the lattice, so this is known as *site percolation*. Instead we could have occupied the edges of the sites (or the bonds), this problem is known as *bond percolation*. Again this affects the percolation threshold, but not the other fundamental properties, we shall return to this point later.

Sudden changes like this are common in other branches of physics. For example a magnet, when heated, looses its magnetisation at a particular temperature (the Curie Temperature). In general these are known as *phase transitions* (or *critical phenomena*) and the percolation threshold is just another of these. It turns out that many of the properties close to this transition can be described in very simple terms.

Not all occupied sites are in the infinite (or spanning) cluster. If we look at the probability that an occupied site is in the infinite cluster ($P(p)$) then clearly this must be zero (since there is no spanning cluster) below the percolation threshold. Above the threshold it can be described in very simple analytical terms

$$P(p) \sim (p - p_c)^{\beta} \qquad p > p_c$$



This is known as a *power law* or *scaling law* and the exponent β is known as a *critical exponent*. This has the remarkable feature that it is entirely independent of the kind of lattice being studied or whether it is bond or site percolation, it only depends on the dimensionality of space (i.e. 2D or 3D). This is known as *universality* and is an important aspect of percolation theory (and indeed critical phenomena in general). Broadly speaking it means that the large scale behaviour of these systems can be described by (relatively) simple mathematical relationships which are entirely independent of the small scale construction. Clearly this is a very powerful concept as it enables us to study and understand the behaviour of a very wide range of systems without needing to worry too much over much of the detail. One key factor which is not universal however is the percolation threshold. But the scaling laws and critical exponents are. A table of values for this exponent β is given below.

	2D	3D
β	5/36~0.139	0.41

Note that in two dimensions it is often possible to determine exact values for the exponents, whereas in three dimensions there are only approximate results or numerical estimates. There are many other critical exponents that can be defined which describe the properties of the percolating system at or near the threshold. There are too many to describe in an introductory article like this and the literature should be consulted for a more complete description (ref to Stauffer & Aharony). Here we shall describe only the most useful for the application described above. Consider first the size of the clusters. First we have to be a bit careful by what we mean by cluster size. For these purposes we start with the *two point correlation function* $g(r)$. This is the probability that if one point is in a cluster then another point a distance r away is in the same cluster. Then this typically has an exponential decay given by a *correlation length* (ξ)

$$g(r) \sim e^{-r/\xi}$$

Clearly at a low value of p these are small (typically clusters of size one or two), this increases until the threshold when the spanning cluster dominates and is infinite in size, so the cluster size diverges. What about above the threshold? Well we have to remove the infinite cluster from our calculation otherwise it will always dominate and only consider the other clusters. As the clusters get absorbed into the spanning cluster

the "typical" size of those left goes back down again. So we have a cluster size which increases, diverges at the threshold and then decreases again. This can be described in a mathematical form as

$$\xi \sim |p - p_c|^{-\nu}$$

Here ν is another critical exponent. As with the connectivity exponent (β) it is universal (independent of details of the lattice) but does depend on the dimension of the system. A table of values is given below.

	2D	3D
ν	4/3~1.333	0.88

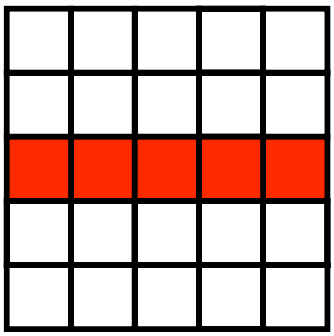
There is a huge literature on percolation theory which defines and calculates a large number of critical exponents. The intention (even were it possible) is not to review all of this material but to focus on the issues pertinent to the questions asked in the introduction. Before we do this there are two issues that must be covered. Everything so far has been defined for an infinite lattice. What happens if a) the lattice is finite and b) there is no lattice at all.

Finite size scaling

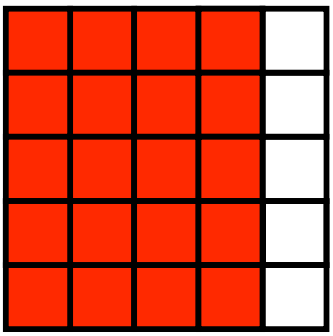
The problem of how to deal with finite size lattices is known as *finite size scaling*. It is a useful introduction to the style of theoretical argument that is often used in percolation theory. We shall look at the connectivity, P.

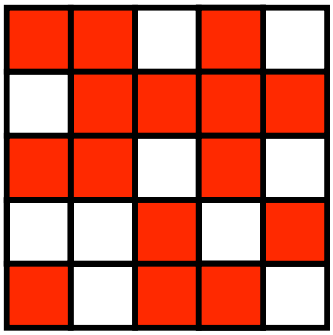
Consider a lattice of size L. For the sake of argument we shall assume it is square so the number of cells is L^d in d dimensions (physicists like to consider the general case wherever possible, although many of our illustrations will be in 2D where it easier to visualise things). The main thing to notice is that things are less clear cut. Consider the following configurations on a 5x5 grid (notice that this is very small and one doesn't really expect these kind of scaling arguments to apply but these give useful insights into the real problems).

a)

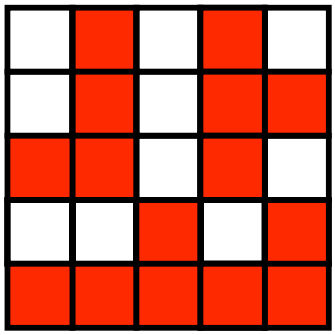


b)

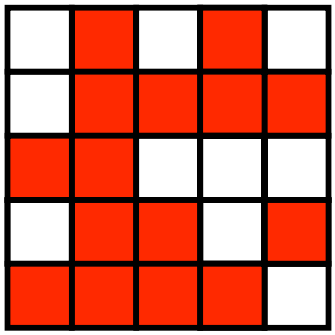




p=0.6 P=0.67

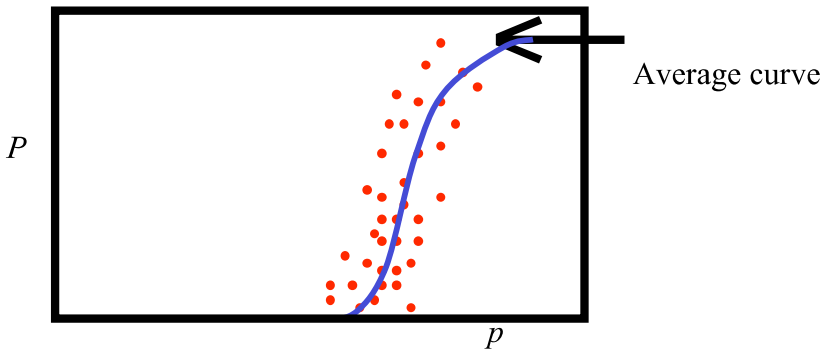


p=0.6 P=0.47

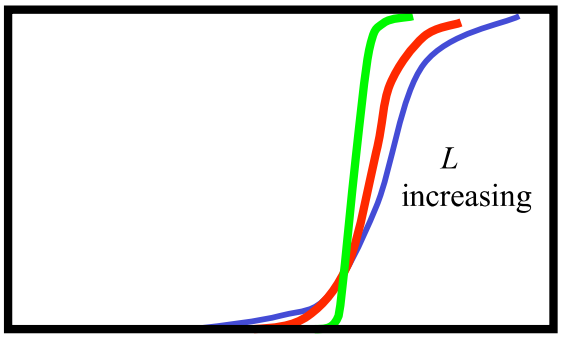


p=0.6 P=0.93

The thing to notice is that you can get connectivity (here defined as connectivity from left to right) at very much less than the percolation threshold (p_c~0.592... for a square lattice in 2D infinite system) (0.2 for Figure a) but not get it at a much higher occupancy (0.8 for Figure b) for a different realisation. Also if we look at some realisations at exactly the same occupancy we can find very different connected fractions (fraction of occupied sites belonging to the connected cluster). This is a consequence of the fact that a finite size system can only sample from the entire distribution of possible configurations and so there is a sample size uncertainty. If we plotted the connected fraction as a function of the occupancy sampled over a large number of configurations we would get a scatter of points. As the size of the system got larger the scatter would reduce until we return to the plot for the infinite system shown earlier.



Instead of considering the whole scatter we can look at the average connectivity (that is averaged over all realisations at the same occupancy fraction). Now we get the following curve. Notice that there is no longer a sharp transition, it gets “smeared out”. As we increase the size of the system the smearing gets less.



This phenomenon is again familiar from other thermodynamic phase transitions where small size systems have "smeared" transitions. We can describe this smearing in simple mathematical terms. First we look at the length scales in the problem. There are only two the system size and the correlation length. Clearly if the system is much larger than the correlation length (which we recall represents the "typical" size of the clusters) then the clusters don't really notice the finite boundaries and the system must behave like the infinite system. On the other hand when the cluster size "sees" the boundaries a new behaviour must be introduced. So the important parameter must be the dimensionless ratio of these two lengths, ξ/L (in fact for later convenience we use an equivalent dimensionless parameter $z=(\xi/L)^{1/\nu}=(p-p_c)L^{1/\nu}$). Then at a given value of the system size the connected fraction must be a function of the form

$$P(p,L) = f_L\left[(p-p_c)L^{1/\nu}\right]$$

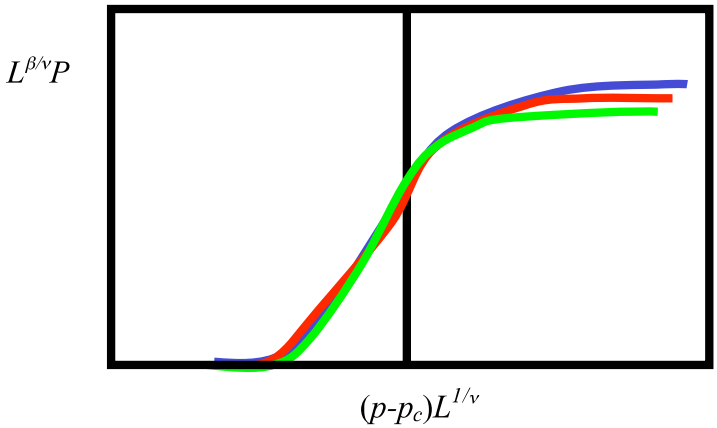
where f_L is some function. Now consider how the behaviour of the system changes under an arbitrary change in size. Let $L \to \kappa L$. Under this change of scale we expect the essential percolation behaviour to be unaltered, that is $P(p,\kappa L) \to c(\kappa)P(p,L)$. The only function that has this property is the power so we can write

$$P(p,L) = L^A F\left[(p-p_c)L^{1/\upsilon}\right]$$

where F is a universal function and A is a power to be determined. To do this we consider the asymptotic behaviour of P. As $L \to \infty$ we must recover the infinite system critical law. Hence $F[z] \to z^\beta$ for large L. To cancel the L dependence we must have $A=-\beta/\nu$. So the finite size scaling law for the connectivity can be written.

$$P(p,L) = L^{-\beta/\nu} F\left[(p-p_c)L^{1/\nu}\right]$$

If we plot $L^{\beta/\nu}P$ against $(p-p_c)L^{1/\nu}$ all the curves found previously should lie on top of each other to form a single universal curve. We can then use this to great effect because if we want to know the finite size connectivity for any system that we haven't performed simulations of we can "unscramble" the result we require from this universal curve.



Continuum percolation

This is very straightforward because of the universality principle. There is no reason why we need to use a grid at all. We can just geometrical objects randomly and independently (formally this is called a *Poisson process*) in a continuum space. Connectivity is defined as the overlap of the objects. In place of the occupancy probability p we have the volume fraction of objects (or the probability that a point chosen at random lies within one of the objects). For the sake of clarity we give this the same letter, p. We get the same threshold phenomenon of a single cluster growing and dominating the system. The percolation threshold depends only on the shape of the objects, but for circles it is 0.678 and for squares it is 0.668 (similarly in 3D for spheres it is 0.28 and for cubes 0.276) so the difference is not very large and numerical experiments indicate that for reasonable convex (i.e. not very spiky) objects the threshold is around the same value. This is known as *continuum percolation.*

Then the principle of universality applies (and has been extensively demonstrated by numerical experiments) that the same scaling laws (e.g. finite size scaling) with the same critical exponents (with the same numerical values). This is a remarkable result and a very powerful one that we can now use.

Application to reservoir modelling

Imagine a typical reservoir model constructed with an object based technique. That is geometrical objects (representing geological entities, e.g. shales, fractures, sand bodies etc.) are placed randomly in space. Then the connectivity can be estimated directly by percolation theory. Take a concrete example of sand bodies in an otherwise impermeable (or low permeability) background. The net to gross ratio is the volume fraction of the good sand and is, therefore, identical to the occupancy probability p. Suppose we have a reservoir of size L and a pair of wells separated by a Euclidean distance r. We can ask question about the probability that the two wells are connected, or in percolation language, in the same cluster. This is just the two point correlation function defined previously. Suppose we want to know what fraction of the sand in contact with the wells is connected to both wells. This is just the connectivity function P defined earlier. We can use finite size scaling to estimate this fraction. Also we can use related scaling laws to estimate the uncertainty. Note that these are algebraic laws with no spare parameters. The percolation threshold is defined by the shape of the objects, but it is largely unimportant whether we model the sand units as rectangles or ellipsoids or other shapes (provided they are not too exotic). The scaling laws and exponents are determined from lattice models (and this has been done very extensively in the literature) and can be straightforwardly applied. As the expressions are simple algebraic they are very rapid to calculate. Compare this with the building of a typical 3D reservoir model which can be computationally very expensive. Here we show results for a North Sea fluvial system of intermediate net to gross. A conventional cross sectional model was built to determine the connected sand fraction. The sand bodies were very long which is why a cross sectional rather than a full 3D model was built. The sand bodies had typical dimensions of 300mx2m and the reservoir interval was 100m thick. Well spacing was taken as 1km. Sensitivities to the net to gross ration were considered as this was uncertain. These results were compared with the predictions from percolation theory. It can be seen that the two are in good agreement. In particular the percolation approach is able to estimate the uncertainty which is not possible for the single realisation reservoir models. The percolation calculations were done in a fraction of a second in a spreadsheet.

Net to Gross ratio	Connected sand fraction (conventional model)	Connected sand fraction (percolation prediction)
25%	6%	2.7 ±4 %
35%	10%	10 ±8 %

Percolation can do more than predict static connectivity. There are scaling laws for the effective permeability. It may be noted that percolation clusters have dead ends which cannot be swept. Only oil in the so called "*backbone*" can be swept. This is a *fractal* object of known dimensions and so obeys a scaling law and can be estimated. We shall describe in more detail how breakthrough time can be estimated.

Estimation of breakthrough time

For simplicity we consider a pair of wells (one injector, one producer) separated by a distance r. we shall also only consider transport of a passive tracer (or a unit mobility ration miscible flood). We can then determine the probability distribution (over different realisations of the sandbody locations). This is the conditional probability that the breakthrough time is t_{br} given that the reservoir size (measured in dimensionless units of sandbody length) is L and the net to gross is p, i.e. $P(t_{br}|r,L,p)$. In previous studies (Dokholyan et al., 1998; Lee et al., 1999) we have shown that this distribution obeys the following scaling

$$P(t_{br}|r,L,p) \sim$$

$$\frac{1}{r^{d_t}}\left(\frac{t_{br}}{r^{d_t}}\right)^{-g_t} f_1\left(\frac{t_{br}}{r^{d_t}}\right) f_2\left(\frac{t_{br}}{L^{d_t}}\right) f_3\left(\frac{t_{br}}{\xi^{d_t}}\right) \qquad (1)$$

$$f_1(x) = \exp(-a\,x^{-\phi})$$
$$f_2(x) = \exp(-b\,x^{\psi})$$
$$f_3(x) = \exp(-c\,x^{\pi})$$

Currently the best estimates of the various coefficients and powers (as found from detailed numerical experiments on lattices and theory, see Andrade et al. 2000) in this are:

$d_t = 1.33\pm0.05$; $g_t = 1.90\pm0.03$; $a = 1.1$; $b = 5.0$; $c = 1.6 (p<p_c)$ $2.6 (p>p_c)$

$\phi = 3.0$; $\psi = 3.0$; $\pi = 1.0$ and $\xi = |p - p_c|^{-\nu}$ $\nu = 4/3$; $p_c = 0.668\pm0.003$ (for continuum percolation).

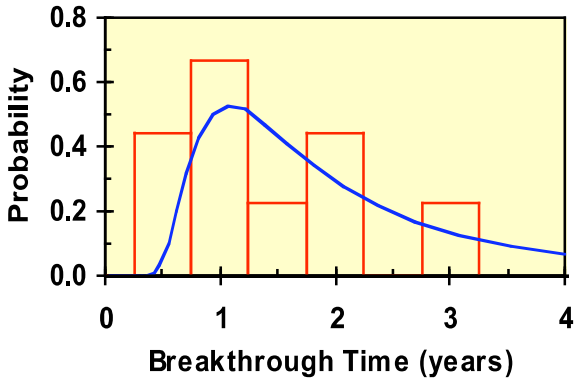
In this paper we will not discuss the background to this scaling relationship but concentrate on how well it succeeds in predicting the breakthrough time for a realistic permeability field. However, it is sufficient to mention that this is typical of the kind of scaling result that percolation is able to provide. However, it is worth spending some time describing the motivation behind the form of the various functions. The first expression (f_1) is an extension to the expression developed by others (see Havlin & Ben-Avrahim, 1987 for a detailed discussion) for the shortest path length in a percolating cluster between two points. The breakthrough time is strongly correlated with the shortest path length (or *chemical path*).

To this there are some corrections for real systems. In a finite size system very large excursions of the streamlines are not permitted because of the boundaries so there is a maximum length permitted (and also a maximum to the minimum transit time). This cut-off is given by the expression f_2. Away from the percolation threshold the clusters of connected bodies have a "typical" size (given by the percolation correlation length,

ξ) which also truncates the excursion of the streamlines. This leads to the cut-off given by the expression f_3. The multiplication together of these three expressions is an assumption that has been tested by Dokholyan et al. (1998). Also a more detailed derivation of this form is given there and the references therein. Here we shall concentrate on using this scaling form to make predictions about the distribution of breakthrough times for a realistic data set.

Rather than considering the theoretical aspects we shall look at an application to real field data. We took as an example a deep water turbidite reservoir. The field is approximately 10km long by 1.5km wide by 150m thick. The turbidite channels, which make up most of the net pay (permeable sand) in the reservoir, are typically 8km long by 200m wide by 15m thick. These channels have their long axes aligned with that of the reservoir. The net to gross ratio (percolation occupancy probability, p) is 50%. The typical well spacing was around 1.5km either aligned or perpendicular to the long axis of the field. In order to account for the anisotropy in the shape of the sand bodies and the field we first make all length units dimensionless by scaling with the dimension of the sand body in the appropriate direction (so the field dimensions are then L_x, L_y and L_z in the appropriate directions). Then scaling law for the breakthrough time can be applied with the minimum of these three values (L=min(L_x, L_y, L_z)). The validity of using just the minimum length has been previously tested (Andrade et al., 2000).

The real field is rather more complex than this, and a more realistic reservoir description was made and put into a conventional flow simulator. We could then enter these dimensions into the scaling formula. It should be noted that first the dimensionless units were converted into real field units to compare with the conventional simulation results. Using these data we find breakthrough times of around one year. The full probability distribution of breakthrough times from the scaling law is given by the solid curve in the figure below.



In addition conventional numerical simulations were carried out for the field. We could then collect the statistics for breakthrough times for the various well pairs to compare with this theoretical prediction. Not all pairs exhibited breakthrough in the timescale over which the simulations were run and there were only three injectors so there were only 9 samples. The histogram of breakthrough times is also shown on the figure. Clearly with such a small sample these results cannot be taken as conclusive however, certainly they are indicative that the percolation prediction from the simple model is consistent with the results of the numerical simulation of the more complex reservoir model. The agreement with the predictions is certainly good enough for engineering purposes. We would hope that if the simulation had been run for longer and more well pairs had broken through that better statistics could have been

collected. The main point being that the scaling predictions took a fraction of a second of cpu time (and could be carried out on a simple spreadsheet) compared with the hours required for the conventional simulation approach. This makes this a practical tool to be used for making engineering and management decisions.

Conclusions

In this article we have introduced many of the concepts of percolation theory and described the simplest of the scaling laws used to describe the behaviour. We have also demonstrated that it can be a practical tool for answering many of the questions that arise when considering geometrically complex systems where connectivity is the primary issue, such as hydrocarbon recovery from oil reservoirs**.**

There are many issues which we have not considered. suppose the bodies are not placed independently of each other (so called *correlated percolation*) such as may be found for stacking patterns of particular sand units, or the distribution of faults. Does this alter the percolation properties? Well essentially no, the same scaling formalism applies. If the spatial correlation between the bodies is of a finite range then this doesn't alter any of the previous discussion. If the spatial correlation has no finite range (such as a power law correlation function) then some of the critical exponents may change their numerical value.

If the permeability cannot be simply split into good and bad then again the approach has to be modified slightly but if the permeability distribution is very broad we can apply a cutoff to the permeability such that we are sitting just at the percolation threshold. This cutoff value then dominates the flow.

These issues, and other, extensions to the simple predictions described here (such as post breakthrough behaviour) are the subjects of our current research into applying percolation theory to oil recovery problems.