Chance and vision on the road to pedometrics

Richard Webster

A first degree in pure chemistry at Sheffield University might seem an unlikely start to a career in statistical pedology. A *Pattern of Islands*, Arthur Grimble's account of his life as a young officer in the British colonies, might seem an even less likely lead into the subject. But I was fascinated by it. So when by chance I learned that the Colonial Office's 'recruiting sergeant' was in the University to lure chemists into tropical agricultural science I was easy meat.

I was awarded a post-graduate scholarship so that I could study soil science and statistics at London University and Rothamsted, and on completion of this 'conversion' to agricultural science I was posted to Northern Rhodesia (now Zambia) as Soil Chemist in 1957. The Land Use Survey of the Copperbelt, an 8000-km^2 region in the north of the country containing several large copper mines, had recently been completed. The original intention of my posting was to evaluate the agricultural worth of the soils mapped by the survey. The fiasco of the East African Groundnut Scheme was fresh in people's minds, and both the Colonial Office and the local Department of Agriculture were keen that future agricultural development would be based on sound information about the suitability of land derived from farming experience, insofar as it existed, field experiments and a proper understanding of the soil's chemistry.

In the event, when I arrived there were more pressing problems. The mine-workers and their families on the Copperbelt could afford to buy food imported by rail from the richer lands in the south. The rural communities in the northeast, cultivating the poor soil there, could barely subsist, and with the population's increasing there was need to identify land that could be developed. It became my job to survey soil for that purpose.

Don Mackney, who would later become head of the Soil Survey of England and Wales, had taught me all I knew about soil mapping—how to classify the soil into series, how to draw boundaries between the series in the field, and so on. Classification was somewhat arbitrary, but not especially difficult in the post-glacial English landscape. Mapping soil on the ancient African plateaux was a very different matter. There were no obvious boundaries between one kind of soil and another. Instead, the soil changed gradually in response to the gentle rise and fall of the land, though with local fluctuation superimposed.

Figure 1 The author at work in the miombo woodland of Zambia.

I was dealing with classic catenas, but that recognition did not solve my problem of dividing them into relatively homogeneous patches for agricultural management and predicting what I should be likely to find at any place in the landscape. I pondered the situation long and hard.

I was no nearer to solving my problem when Philip Beckett and Frank White of Oxford University landed in the Country. The Royal Society of London had sponsored them to study the soil and vegetation in relation to the physiography of the African plateaux. I joined them in the field for part of their stay, and in the evenings we would sit around the table in the rest house, lit by a single paraffin lamp, and discuss the day's observations. We would return to my unsolved problem of prediction time and again, only to determine to find a solution.

Some months later I received a letter from Philip: he had obtained a grant from the British Government to solve the problem in a military context. He wanted a soil scientist to work alongside the Royal Engineers and their civilian counterparts. Was I still interested in predicting soil conditions at unvisited places? Of course I was, and as soon as the various documents had been signed and sealed I headed for Oxford.

The year was 1961. A few engineers had begun to realise that the problem was essentially statistical and were toying with a combination of classical soil maps and prediction statistics based on stratified random sampling in which the classes of the maps were the strata. As far as we knew at the time none of them had tested their own maps in these terms. We should make maps ourselves. Because the context was military we should do so largely by air-photo interpretation and we should call the classes 'land facets' rather than soil series or soil types.

We should then test our maps for their effectiveness in (a) diminishing the variance of soil properties within classes and (b) predicting the values of those properties with acceptably small variances. We also tested maps made and sampled by several of our collaborators. We had mixed success. Our map of the Oxford region enabled us to predict the mechanical properties of the soil reasonably well. It predicted relatively poorly the soil's pH and organic matter content, and it was useless for predicting the plant nutrient status of the soil.

Even in the most favourable situations there were substantial residuals for which we could not account. Also, we were still wedded to classification as a way of describing the variation we could see. We had not solved the problem of the catena or any other form of gradual change or trend. If we simply drew boundaries in those situations then the residuals would contain trend. We also recognized that trend surface analysis, then becoming fashionable in geography and petroleum exploration, was unsatisfactory because the residuals were correlated. Further, if there were neither evident trends nor clear boundaries then how were we to describe the variation?

Figure 2 Philip Beckett (right) leading the army's pool of engineering geologists on English farmland.

Enter Heriberto Cuanalo de la Cerda from Mexico. He brought new ideas: time-series analysts have similar problems, and they treat actuality as realizations of stochastic processes to describe fluctations quantitatively in time. Could we not do the same for soil? So we switched our thinking from classical mode and took the wild leap of imagination; we would treat the soil as if it were random against all the tenets of the day! To test the feasibility of this approach he painstakingly described 321 pits at 10-m intervals on a transect across north Oxfordshire, and we wrote programs to compute correlograms from his data and to plot them. We found strong spatial correlation extending to 250 m or more. There were evident boundaries between classes of soil. We removed the class means from the data and discovered that there was still spatial correlation in the residuals. Where were we to go from there? How could this form of analysis lead to prediction? Cuanalo returned to Mexico and I, by now a member of the Soil Survey of England and Wales on the Rothamsted staff, turned my attention to another problem that was taxing my colleagues, namely multivariate classification.

However, one morning in 1972 I received a telephone call from Australia. Gordon Hallsworth, then chief of the CSIRO Division of Soils, was on the line. A young Englishman, John Norris, whom he had recruited to help Bruce Butler quantify soil survey, had died in a road accident, and Gordon wanted someone with the necessary skills to take over the project as soon as possible. Was I free? I negotiated a year's leave of absence from Rothamsted and joined Bruce Butler in Canberra the following January. I re-analysed Norris's soil data on the Southern Tablelands of the Australian Capital Territory but could make little sense of them. Correlation between variables was in general weak. Soil classification scarcely improved the ability to predict.

There seemed to be no common spatial pattern among the many variables that Norris and Butler had recorded, and some variables had no evident spatial pattern at all. This was not how things were supposed to be; it was not what CSIRO's pedologists expected when they first attempted, and failed, to map the soil there. Butler and I asked the questions: (a) on what spatial scales are the individual properties of the soil varying? and (b) can we discover economically what those scales are? This led to our adapting nested random sampling in which each hierarchical level was a fixed distance separating sampling points on the ground. By analysis of variance we could estimate the variance associated with each distance. We later realised that we had rediscovered an innovation of Youden and Mehlich but whose publication in 1937 in the house journal of their research institute had lain unheralded for more than 30 years.

Figure 3 Bruce Butler scanning the Riverina in south east Australia with expert eye.

Nevertheless, we did take the analysis a step further in that we accumulated the components of variance from the shortest distance to the largest and thereby formed crude variograms, the first of any soil properties. We also discovered that the properties we analysed in that exercise did indeed vary on disparate scales and that it was small wonder that the pedologists had largely failed to map the soil by conventional means. At the same time I pressed the analogy with time series into the spectral domain and surveyed a transect across gilgai landscape as a case study. Now the gilgais, typically gentle depressions in plains and widespread in eastern Australia, appear in patterns that seem to have some degree of regularity. These manifested themselves as peaks in the spectra.

These two studies occupied much of my sabbatical year in Australia, but I had still not worked out how to move from the variogram or spectrum to prediction. I was about to depart and was tidying up my office when a total stranger marched in unannounced. The intruder came straight to his point without a moment's delay or even introducing himself: 'They tell me that you are some kind of statistician. Well, what's this kriging?' I had never heard the word, and his brutal introduction put me on the defensive; I played for time. I asked the newcomer who he was and to explain the context, which he did. He was Daniel Sampey, a mining geologist. He told me that a certain Professor Krige had discovered how to optimize the estimation of gold reserves on the Reef in South Africa. He also told me of Georges Matheron, of the theory of regionalized variables and of its application in geostatistics. I let him continue, which he did for about 15 minutes with only the occasional 'mmm' and 'yes I see' in the way of encouragement from me. Then, clearly disappointed that I knew even less than he did, he left as abruptly as he had arrived. His parting shot was that as I was about to return to Britain I should visit Leeds University where mining engineers knew a thing or two.

I shall not forget Daniel Sampey. In those 15 minutes I realised that my problem of spatial prediction of soil conditions at unvisited places had been solved, at least in principle and by miners, and in general terms I understood how. On my return to the Soil Survey I lost no time in contacting Anthony Royle, lecturer in the subject at Leeds, and a week or two later we met. He amplified what Daniel Sampey had told me, and he generously gave me a copy of his lecture notes on the subject and a few references to the literature including Matheron's seminal thesis—in French. My next student was Trevor Burgess, an Oxford mathematician who was looking for a postgraduate position in which he could apply his talents. He was the ideal person, keen and swift to appreciate the problem and its solution. We turned Matheron's equations into algorithms and algorithms into computer code. And in 1980 our first scientific papers appeared, the first to describe for soil scientists the variogram as we know it today and the first to display maps of soil properties made by kriging.

Trevor obtained his doctorate and was followed by Alex McBratney and Margaret Oliver, both of whom made their careers by extending the applications ever more widely and helping to establish geostatistics as an essential strand in modern quantitative soil survey through their writing and teaching.

Someone some day was bound to see in geostatistics the tools that are now proving so effective in soil survey. It was largely by chance that that someone happened to be me. What might I have done had I not read Arthur Grimble's book, had I not learned of the Colonial Office's recruiting drive, had Philip Beckett not ventured into Zambia, had John Norris not met such an early death, and had Daniel Sampey breezed into my office the week after I had left Australia instead of the week before? We shall never know. All we know is how it happened by a combination of opportunity, good fortune and more than a little persistence. Let's give thanks for it.