ΠΕΔΟΜΕΤRΟΝ

The Newsletter of the Pedometrics Commission of the IUSS

Issue 23

December 2007

Chair: Murray Lark

Vice Chair & Editor: Budiman Minasny

From the Chair

Dear Pedometricians,

Welcome to the final Pedometron of 2007, I hope that you will find it a worthwhile read. There are reports on Pedometrics 2007, Markov Chain random fields, geostatistical bibliometrics, pyrometrie, profiles, Alex's preferred papers and a new Pedomathemagica quiz to keep your brain cells working over the Christmas holiday."

In September this year I attended the 2nd International Soil Forensics conference in Edinburgh. I had not given much thought to forensic soil science before I was asked to go and give a talk on soil geostatistics, but I am glad that I went as it is always interesting to try to get up to speed in a new area of science.

Soil forensics is not new. In 1856 a barrel of silver coins was stolen from a train on the Prussian railway. Or rather the coins were stolen, the barrel was filled with sand and left in position to avoid arousing suspicion. Professor Christian Ehrenberg of Berlin University was asked to help the investigators. He examined the sand from the barrel under a microscope, and then compared it to samples that he collected from every station that the barrel had passed through. He found a match, and when the police visited the corresponding station they soon recovered the coins and made an arrest. Today soil scientists are using various technologies, including microscopy and spectrometry, to measure a wide range of geochemical, biochemical and physical properties of the soil in the hope that soil material from an exhibit (such as a suspect's footwear or vehicle) can be linked to a crime scene for evidence or intelligence purposes.

Much effort is being devoted to making measurements, but very little to the problem of how appropriate inferences can be made from those measurements. In fact this is a gaping hole in soil forensics from what I could see (and possibly in related environmental forensic sciences as well). This is despite some very advanced work on forensic inference which has been undertaken by statisticians, largely motivated by the problems of interpreting evidence from DNA or other molecular measurements. We all know that the soil is enormously variable. I think that most pedometricians would deduce from this that it is no small problem to prove 'beyond reasonable doubt', as English criminal law puts it, that soil on a suspect's trainer comes from the site where the body was buried. These problems are exacerbated by the fact that a soil sample will have a defined support, while the support of soil taken from a shoe, if treated as a sample of soil at some unknown site, is hard or impossible to define. The soil forensics community, as a whole, does not seem to pay much attention to these issues. They conclude from the enormous variation of soil that any site will, in consequence, have unique soil conditions. The questions that statisticians raise were treated with some suspicion by speakers at the meeting. In fact I started collecting questionable quotations such as 'statistics just opens a can of worms', 'statistics just confuses a jury and makes it harder to win a case' and 'statistics is a bit dangerous'.

Christmas Issue

From the Chair

The English Court of Appeal recently overturned the conviction of a woman for murdering her child. The child had been the second of two to die inexplicably in its sleep, cot-death as it is called. An eminent paediatrician was called as a prosecution witness to the original trial, and gave evidence that the probability of two cot deaths occurring in one family is so small that this could be ruled out as a cause. His evidence was based on squaring an estimate of the probability of one child dying in this way over a certain period. That is to say he assumed that occurrences of cot death within a family could be treated as independent random events. Of course, he made no such *explicit* assumption. He was simply not competent to do the statistical and probabilistic inference, but the lawyers and jury failed to recognize that his ability in medicine did not stretch to statistics. My concern about soil forensics is that forensic inference from soil evidence could sometimes be similarly flawed, and might bring soil science into disrepute, unless pedometricians get involved in questioning and improving the way in which forensic soil scientists sample crime scenes and go on to make inferences. This is one new area that pedometricians could engage in, and I invited delegates to the Forensics meeting to contact me if they were interested in holding a joint workshop. No takers yet, but if any interested Pedometricians were to contact me then perhaps we could start something up.

With all best wishes for the New Year.

Murray

www.digitalsoilmapping.org

Dear Pedometricians,

It is our pleasure to announce forthcoming events relating to Digital Soil Mapping for 2008:

1. The High Resolution Digital Soil Sensing and Mapping Workshop (University of Sydney, Australia, 5-8 February 2008)
2. The Digital Soil Mapping Session (SSS25) of the Soil System Science Group to be held at the European Geosciences Union General Assembly 2008 (Vienna, Austria, 13 – 18 April 2008).
3. The Session on Pedometrics and Digital Soil Mapping (S26) at the Eurosoil Congress 2008 (Vienna, Austria, 25-29 August 2008)
4. The Digital Soil Mapping Workshop 2008 (Utah State University, Logan, Utah, USA 30th September to 3rd October 3, 2008).

You can find further information on each event at the updated DSM Website: www.digitalsoilmapping.org

There is growing interest in DSM and this is very encouraging. This interest is likely to increase further because of improved links to related disciplines (e.g. remote sensing, applied instrumentation, Earth system science) and the pressing need for better soil information (e.g. to support critical studies into global carbon dynamics).

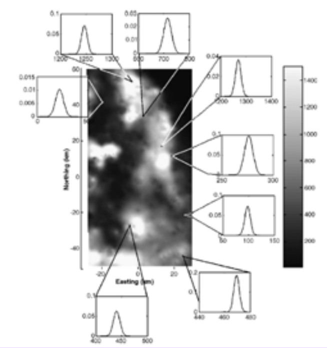
We have to develop a much better basis for comparing the efficacy of different methods for DSM. In particular, we need a clear basis for estimating the uncertainties of predictions for soil properties in space and time. The measurement programs, pedological models and statistical methods for data analysis have to satisfy the standard tests of scientific peer review and be sufficiently explicitly for our colleagues in related disciplines to understand.

To this point, we wish you a Merry Christmas and a Happy New year and we hope to see you soon!

Florence and Neil

Best Papers 2005 & 2006

The Best Papers in Pedometrics for 2005 and 2006 were announced at the Pedometrics 2007 conference dinner. The best paper for 2005 was by E. Savelieva, V. Demyanov, M. Kanevski, M. Serre, & G. Christakos, entitled *BME-based uncertainty assessment of the Chernobyl fallout*. (*Geoderma*, **128**, 312–324). This paper described how the Bayesian maximum entropy method was used to predict the distribution of Caesium 137 in soil affected by the Chernobyl incident. The BME methodology allowed different sources of information to be integrated for prediction, and some new ideas for presenting the resulting uncertainty were explored.



The best paper for 2006 was by G.B.M. Heuvelink J.M. Schoorl, A. Veldkamp & D.J. Pennock. This was another paper to present and exemplify some emerging methodology in pedometrics. It was entitled *Space-time Kalman filtering of soil redistribution* (*Geoderma*, **133**, 124–137). It was shown how the state-space approach of Kalman filtering could be used to combine empirical and process-based methods to predict soil redistribution. Another link between the papers is that this study also considered Caesium 137, although this time the isotope was used as a tracer for soil redistribution.

It is sometimes argued that pedometrics substitutes powerful statistical methods for scientific understanding of the soil. These two papers are good evidence against that tired old canard (I restrict myself to polite language in these august pages); they show how very powerful methodology is precisely what enables us to use process-based information for real-world prediction.

Call for Best Paper 2007

It has been agreed that we will change the way in which the Best Paper award is managed. We shall now hold a vote each year on papers published in the previous year, and the awards will be presented at the biennial Pedometrics meeting.

We now want to invite members of the commission to submit papers that they think should be considered. The rules are as follows.

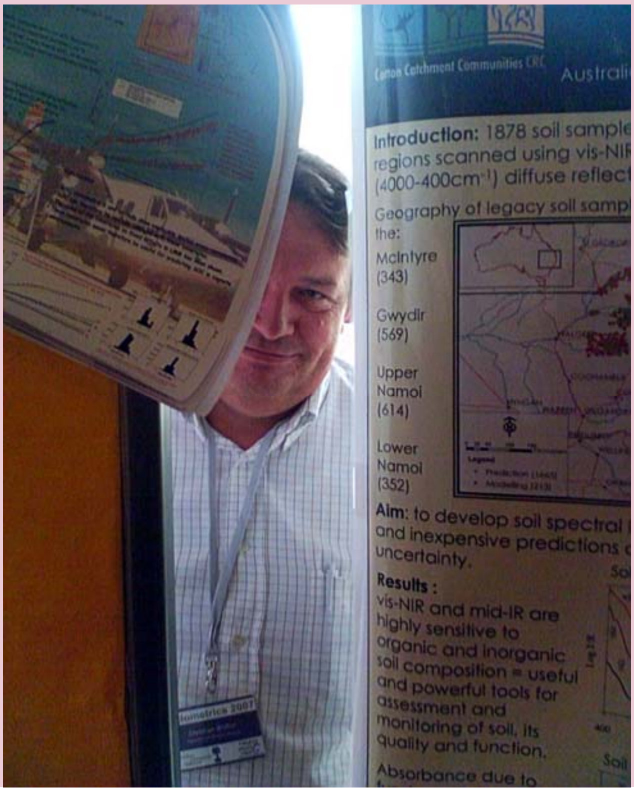
1. Any member of IUSS may nominate a paper, provided that it is not a paper on which they are author or co-author.

2. One member of the Commission will be invited to produce a shortlist of five papers from all that have been nominated, or other eligible papers.

3. A vote will then be held, in which all IUSS members may participate.

4. To be eligible for consideration a paper must be published in the year for which the award is given (this does not include publication in an 'Online Early' or 'Articles in Press' section). The paper must be published in a peer-reviewed, international and accessible journal. Conference proceedings and book chapters are not eligible. Papers must be on Pedometrics, showing how mathematical and statistical methods can advance the study of soil.

All nominations for Best Paper in Pedometrics 2007 should be sent to Murray Lark (murray.lark@bbsrc.ac.uk) by 1st April 2008.



OPINIONS

Dear Pedometricians,

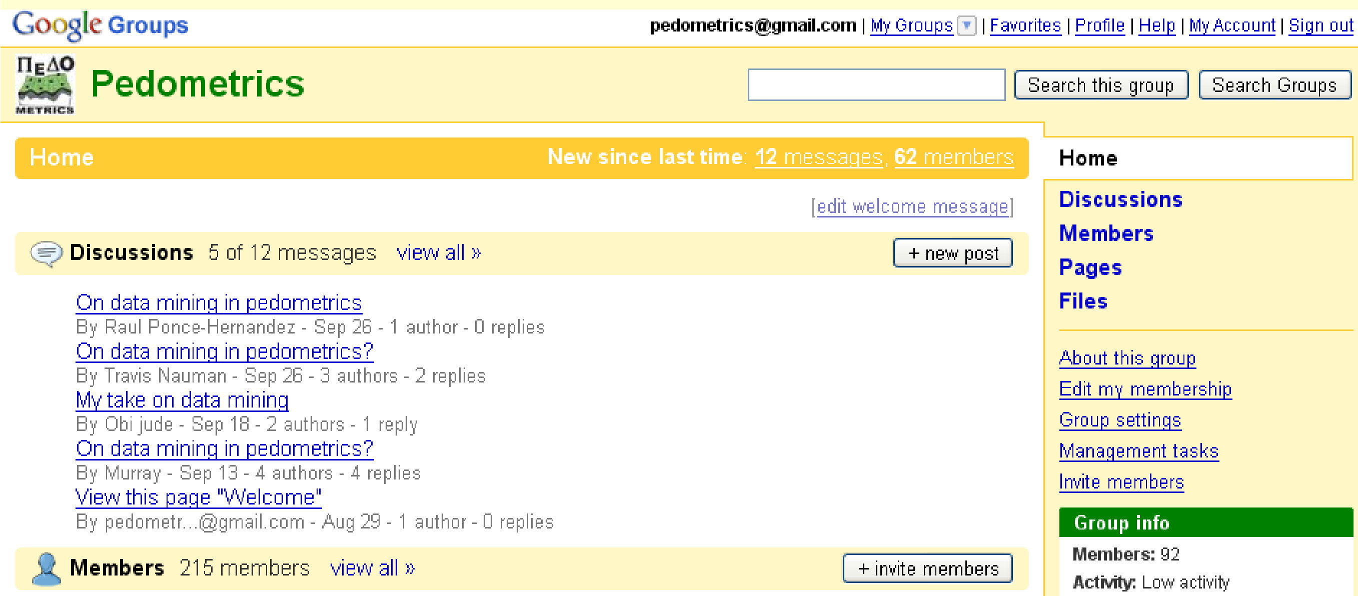
I would like to discuss the multiplication of DSM events and also some recent views. On one hand it is a positive point since it proves that DSM attracts more and more scientists, maybe due also to the facilitated access to satellite images, to the development of new tools which are also giving access to soil information (particularly proxy sensors) and of course, to DSM workshops which have been organized (thanks again to INRA Montpellier, France and EMBRAPA Rio, Brazil).

On the other hand, it is somehow less than ideal since too many workshops means unnecessarily duplication of scientific presentations. Innovations are then more difficult to distinguish from applications. This can even lead to misinterpretation of the knowledge due to lack of experts giving presentations since demand is outstripping supply.

It is not surprising then to hear: “now, with DSM you can produce a soil map from only few samples”. This propaganda is based on real applications (we let you find these applications). In that case, you will rarely find the ‘real’ uncertainties (calculated properly) associated with the maps. Sorry, but DSM can not be reduced to a click on a button. DSM is a concept with a scientific framework involving soil expertise and specific scientific steps to be performed as well explained in peer-reviewed papers!

Florence

Pedometrics Discussion Group



In a meeting in Tubingen, some suggested that we should have a “Discussion” group, which is useful as users can post messages, and they are forwarded directly to our email, which we read daily. So the Pedometrics Discussion Groups has been set-up in Google Groups. There is now 215 members, we started a discussion on “Data Mining” and got several good feedbacks. But now it is quite dormant. As you can see above, it has a **Low activity**. We need to keep us the discussions, if you have some burning issues or questions related to Pedometrics fire them to the group. If you haven’t joined the group, please do so at : http://groups.google.com/group/pedometrics

Pedometrics 07, Tübingen

Photos by Brian Slater http://homepage.mac.com/bslater/Pedometrics2007/

Report from Pedometrics 07, Tübingen

by Johan Van de Wauw

After a long trip, we finally arrived in Tübingen. The sun was shining and luckily we still had time to visit the Biergarten on the banks of the Neckar river. When we felt strong enough, we headed for the Welcome Social. After registration, we caught the first glimpse of German hospitality at the welcome BBQ in the Geographic institute of the Eberhard Karls Universität. Apart from the wealthy food we could also choose a range of exclusive beers. My first pedometrics conference couldn't start better! I was happy to find out that behind many people whose names I previously only encountered in articles were such a pleasant company!

The first day of the pedometrics conference started with a keynote speech by the recipient of the first Richard Webster medal: Alex McBratney. He focused on his own achievements and those of his team. The major point of his speech seemed to be that to develop pedometrics, we need new creative ideas. After stating that this is incompatible with iPods, some conference participants craftily hided their white devices. The morning continued with talks on sampling and digital soil mapping.

The afternoon sessions started with an interesting keynote of Zhu A-Xing, who discussed the effects of neighbourhood size on digital terrain derivates. I believe this was a very interesting talk, proving that the fixed link between grid size and neighbourhood in most GIS software should be abandoned in favour of algorithms where neighbourhood size can be adjusted independently. The session continued with mixed topics and -even though the chair tried to declare the decline of geostatistics- some interesting geostatistics papers. The long day ended with the first poster session. Unfortunately, due to the limited space and the long preceding day, many people dropped out during this presentation.

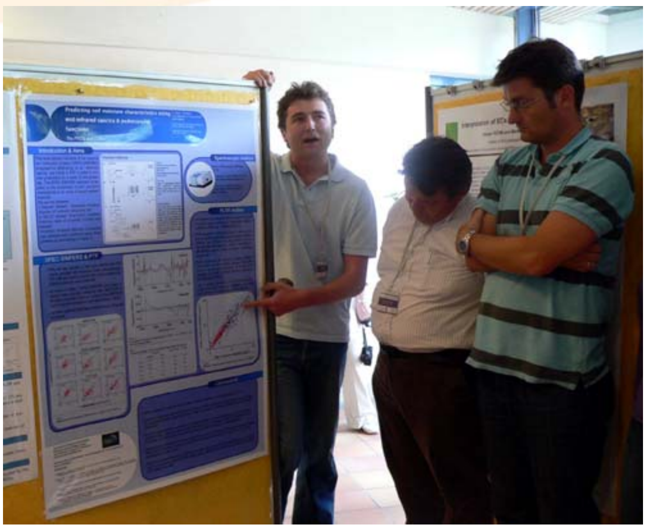
The day ended with a guided tour in the university-town of Tübingen. The highlight for most of us was probably the old university prison.

The keynote talk on this second day of pedometrics was done by Murray Lark who's topic was *"not data mining in pedometrics"*. One of the conclusions was that expert knowledge is still very useful for developing methods instead of relying only on statistical tools. The session continued with talks on digital soil mapping. After the break the session theme was uncertainty. During the breaks more and more pastry and sweets appeared (I definitely gained a few kilos there!).

The afternoon had sessions on Soil Sensing, the second poster session and finally some mixed topics. After those presentations Tomislav Hengl presented the new pedometrics website where the true nature of some pedometricians was revealed. I skipped the rest of the business meeting and rushed to the hotel to prepare for the conference dinner.

The eagerly awaited conference dinner was hosted in the "Casino" near the Neckar river. Those who came armed with VISA-cards and fake dollars were deceived to find out that this was not a gambling place, but an old French officers quarter. During the dinner three prices were handed out: the best paper award for 2005 and 2006 and the maybe even more prestigious "pedomathemagica quiz" award.

The best paper for 2005 was awarded to Savelieva et al. for their BME-based uncertainty assessment of the Chernobyl fallout. Unfortunately none of the authors was present to receive the price. For 2006 this award was given to Heuvelink et al. for their article on



Space-time Kalman filtering of soil redistribution. Gerard was very pleased to win the price after being nominated for several years.

The answers of the "pedomathemagica" quiz that was running during the first two days of the conference were given and explained. The winner received an original copy of *Jenny's Factors of soil formation*, and that winner was .... Gerard Heuvelink!

Waking up the third day of pedometrics was becoming more difficult, but luckily I managed to be in time for the interesting keynote of Dick Brus about Sampling design. After this keynote different presentations on fuzzy logic were given. The best poster award was handed over to Raphael Viscarra after the break, and the last session on digital soil mapping was started. The afternoon concluded the conference with presentations on different topics. Due to the rain the planned punting activity was replaced by a very successful pub strawl in Tübingen.

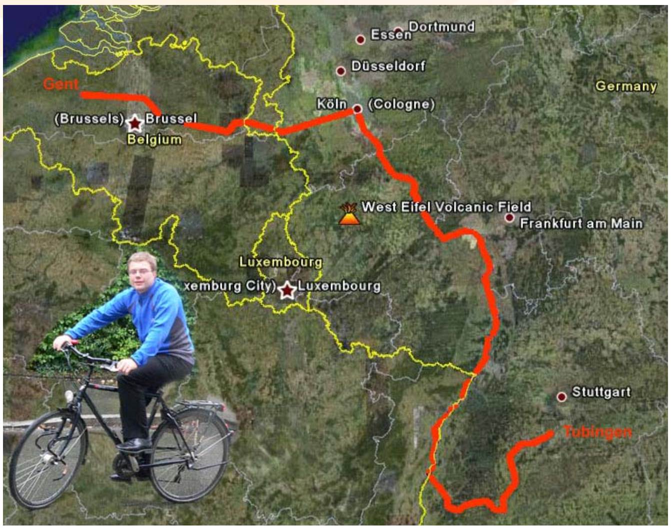
The next morning everyone was on time for the excursion. After a first stop where we had a general introduction and we headed for the castle of Hohenzollern. At least, this was what we were told because we only saw clouds were the castle was supposed to be. When we arrived the clouds started to disappear and by the time our guided tour was finished we could actually see where we were standing! Our tour continued along some soil profiles and the very impressive landslide of Mössingen. Already behind schedule we headed for lunch in a wood side barbecue site. After lunch we visited a few soil profiles that formed on limestone in the experimental farm of the university of Hohenheim. all of us learned to pronounce the word "*Schichtstufe*" and visited one of the highest

escarpments of this region in Germany. Just before sunset we also visited the soil trail of Beuren. By sunset we arrived at the vineyard where we received a guided tour and had the opportunity to taste their wines. I was glad to find some bread and finger food on the table as it was already getting late and we didn't have dinner.

August 31

Those who booked their plane-tickets based on the official programme (or those who woke up too late) were unlucky, because they missed the very relaxing sun-drenched post-conference punting tour hosted by the students who helped to organise the conference.

To conclude I would like to thank the organisation and the speakers for the very interesting programme (inside and outside the conference hall) and the very smooth organisation. Furthermore I would also like to thank all other participants for the interesting discussions we had during those six days.



Johan is a a PhD student at the Department of Geology and Soil Science of the University of Gent, Belgium. Johan rode his bike from Belgium to Tübingen taking 10 days for 791 km.

by Rosina Grimm

In 2005 I started my Master's project in soil science. Soon after, I found out about the international working group on pedometrics. While examining the broad range of publications on pedometrics, I was increasingly fascinated by the extremely outstanding quality of these studies. The same year, I had the chance to attend the pedometrics conference in Naples, Florida. From that moment on, I was sure that pedometrics exactly matched what I had been looking for. Now, two years later, I was very pleased to give a talk in this community.

Except for those already attending the pre-conference workshop, for most participants the conference started off with a very cordial welcoming reception held in the garden of the Institute of Geography at the Eberhard-Karls-University of Tübingen. Since cold beer from local breweries and traditional sausages were served, it was also a cultural welcome to Germany for all international attendees. Immediately after arriving, everyone was involved in a social and comfortable atmosphere.

One of the major aspects about conferences is communicating and networking. The small size of the pedometrics conference with approximately 100 participants is a great advantage concerning these aspects. During the whole conference especially in the breaks and evenings, participants always gathered together discussing ideas, problems, future projects etc. – as well as cultivating old and gaining new friendships.

For my part, I can give an extremely valuable example concerning communication and networking: Two years ago, I met Thorsten Behrens on the pedometrics conference in Florida. Together we formulated new ideas on potential research subjects and submitted an abstract for the latest conference in Tübingen. When our abstract got accepted for oral presentation, this teamwork was my first time of successful scientific collaboration beyond the borders of institutions and universities – which became possible through the pedometrics conference.

Equally important, the conference was educationally at a high level. The program covered both theoretical and applied pedometrics research, and nearly every contributing presentation or poster was scientifically very sophisticated and clearly presented. The four keynote speakers also covered a broad range of pedometrics research, and emphasized critical points in

Communication & Networking among participants

a comprehensible format. The first keynote talk by Alex McBratney, who is a celebrity in pedometrics, was a broad review on pedometrics research and the ideas and methods he developed together with his excellent working group and others. On the other hand, the following three keynotes also given by luminaries in pedometrics were very critical to my ongoing thesis with special regards to DEM derived terrain attributes and their neighbourhood sizes (A-Xing Zhu), on restrictions of data mining techniques (Murray Lark), and on deciding soil sampling strategies for different research aims (Dick Brus).

In addition to the scientific aspects on the conference that everyone contributed to, I was fully enthused by the whole organization of the conference. Besides the welcome social at the institute's garden, the conference dinner in the historical French casino building with its delicious European cuisine and the guided tour through the old town of Tübingen were also set up very nicely and gave the conference an unforgettable social element. Unfortunately, I did not attend the post-conference field trip. However, people who attended ensured me that it was an interesting trip with special program points on both cultural parts like visiting the famous castle of Hohenzollern and the wine-tasting of a local wine grower, as well as scientific aspects like examining soil profiles.

In closing, I want to - and at this point I think I can speak for all of the participants - thank the organizers Thorsten Behrens, Thomas Scholten and Volker Hennings as well as all of the students who worked so hard on making this conference a nice and pleasant as well as scientifically successful and profitable time for everyone.

by Brian Slater

Pedometrics 2007, the biennial conference of Commission 1.5 of the IUSS, was held in Tübingen, Germany, in late August 2007. I believe that this was the seventh international Pedometrics meeting. Judging by the vitality of the meeting, the marriage of quantitative methods and Pedology is not yet suffering from the "7-year itch". In large part, the success of the meeting resulted from exemplary organization provided by Thorsten Behrens and his local team in Tübingen. Also, Tübingen is a gem of a University town, and was a perfect venue.

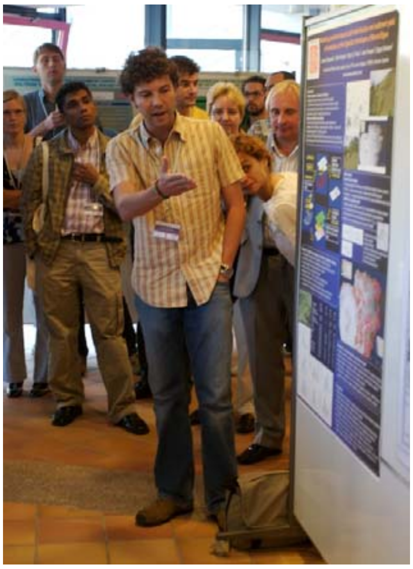
On the evening of Sunday August 26, the Tübingen folks prepared a social gathering at the Institute of Geography and welcomed participants with plenty of local food and liquid refreshments. A number of people remarked on the success of the pre-conference workshop on Spatial Uncertainty Propagation, which had been held over the weekend under the leadership of Gerard Heuvelink and James Brown. (Bas Kempen provided a report of the workshop at http://www.pedometrics.org).

The conference began on August 27 at the Theologikum, just a short distance from the Institute, and a pleasant walk from various hotels in Tübingen. Part of the charm of the town is that the University and the town intermingle, as do students, residents and visitors. Thomas Scholten from the Institute of Geography provided a warm welcome to more than 80 participants. The format of the conference consisted of keynote talks framing general sessions, and poster sessions. The initial keynote from Alex McBratney provided a typically encyclopedic overview of the history of Pedometrics and made clear how seeds Alex and his colleagues have sown have germinated, emerged, and in many cases matured to bear fruit. Many of the areas of discourse and memes that Alex mentioned emerged as themes throughout the conference.

Subsequent keynote talks were given by A-Xing Zhu (on the effects of neighborhood size on terrain derivatives for DSM), Murray Lark (on the dangers of the unanalyzed life in general and unanalyzed data mining in particular), and Dick Brus (on rational soil sampling strategies particularly as applied to environmental monitoring).

The general sessions included 42 presentations generally well-grouped into topical areas including sampling, digital soil mapping, geostatistics, fuzzy logic, uncertainty, soil sensing, and "mixed topics". Two poster sessions involved 44 posters.

Did an identifiable message or motif emerge during Pedometrics 2007? The presentations and approaches were diverse. Clearly, many of the themes and many of the presenters have been consistent throughout the history of Pedometrics and the biennial meetings. I think this is evidence of the maturity of the discipline, rather than an indicator that nothing is new. I think the brightest area of development in Pedometrics is the attention to uncertainty analysis as evidenced by the workshop and a number of the presentations and posters. Digital Soil Mapping has to some extent revitalized Pedometrics and has certainly opened the door



to new participants and some older acquaintances. Pedometricians remain among the most accomplished practitioners of geostatistical methods and continue to provide leadership in the application of innovative quantitative methods for addressing real issues in environmental management.

I think the quality of presentations and content of the general and poster sessions was excellent. The poster sessions were greatly enhanced by the short presentations given by the authors at their posters. Perhaps the most disappointing aspect of the conference for me was the lack of adequate time for discussion during the general presentation sessions, and the lack of any plenary session. There were many moments where discussion was warranted, and protagonists barely had time to put on their gloves for a fight before they had to sit on their hands again. This will be a challenge for future conferences if there are as many as 40 oral presentations in less than three days.

Fortunately, there were many other opportunities for professional and unprofessional interaction. Murray Lark chaired a open business meeting, during which Tom Hengl talked about the Pedometrics web site, and when it was announced that the next biennial meeting will be hosted in Beijing, China. Our hosts arranged a walking tour of Tübingen, giving us a great opportunity to learn about the history and character of the the town and the University.

There was also the challenge of Pedomathemagica, a quiz published in Pedometron 22, distributed at the-

conference. There was feverish activity to solve the quiz and some unholy alliances formed to develop consortia entries, with the goal of winning the coveted prize - an original copy of Jenny's book "Factors of Soil Formation". The conference dinner was held at the "Casino" restaurant, on the banks of the Neckar River. Following indulgence in excellent local food and wine, Murray Lark announced various prizes which were all taken home by Gerard Heuvelink: a share of the best paper award, and the best score on the quiz.

When rain ended any chance of a punting expedition on the Neckar River following the final day of the conference, we commiserated with quality time spent in several excellent watering holes in Tübingen.

On August 30, a post-conference field trip introduced the soilscapes and culture of Baden-Württemberg. We visited sites near Tübingen to gain an understanding of the local geology and geomorphology. After observing several soil pits, we explored the Hohenzollern Castle and a landslide area (Mössingen). A barbecue lunch was provided, as well as further opportunities to discuss the diversity of soils in the area. Our final field stop was a visit to the Beuren Soil Information Trail, an interpreted trail including soil pits provided for the public to learn about the soil resource. The day was perfectly concluded with a visit to Fellbacher Weingärtner, a winery near Stuttgart, where we toured the operation and enjoyed an excellent tasting session. Many thanks to Thomas Scholten, Thorsten Behrens and Karsten Schmidt for organizing an excellent field trip.

Traditional or Digital Soil Mapping?

Maria de Lourdes Mendonça-Santos

Soil Researcher at Embrapa Solos - Rua Jardim Botânico, Rio de Janeiro, RJ - Brazil

Quantitative Pedology

Pedometrics is about quantitative Pedology or quantitative methods applied to Pedology. Pedometry in itself, is not new, because mathematical and statistical methods have been applied to Soil Science since the sixties (Yaalon, 1975). Its origin is far back in Jenny (1941), that in a such simplified way (seemingly), got to express the complex relationships involved in soil formation processes established by Dokuchaev in the beginnings of Soil Science: the soil is the result of the interaction among five factors: climate, organisms, relief, parent material and time. Jenny expressed the complex relationships among the factors of soil formation in a quantitative way through the equation ***S* = *f* (*cl.o.r.p.t...*)**.

That equation, apparently so simple, has for each one of its parts, an explanation and complex correlations, that he details in his pioneer book "**Factors of Soil Formation - A System of Quantitative Pedology**". Through that equation, Jenny admits that the soil properties can be quantitatively measured and expressed numerically, through mathematical correlations. That theoretical equation allowed Jenny to predict the distribution of soil properties based on external variables and prior to field observations. Since then, pedologists all over the world took the quantitative road trying to formulate the famous Jenny equation through case studies where one of the factors varies and the others are kept constant, originating quantitative functions known as climo-functions , topo-functions, etc. Most of that kind of work developed before the existence of sophisticated numerical methods and availability of computational resources.

A Novelty?

Then, why to consider Quantitative Pedology or Pedometrics as a novelty or counterpoint to Pedology purely qualitative or traditional? Why not consider, in the current days, the accelerated development of Sciences and correlated Technologies, Computer Power, Remote Sensing, Geographical Information Systems, GPS, Databases, Geostatistics, Data-mining, and other tools that can enlarge the horizon of Soil Science? Pedology as a whole developed and expanded in the same pace with the evolution of the computational and technological resources as expected in any Science.

That evolution, illustrated in the Figure 1 (McBratney et al, 2000), shows that after certain time, the difference between the traditional approach and the quantitative one, is decreasing and even, nonexistent, when the technology and the formal knowledge are joined to produce important information for human society as a whole.

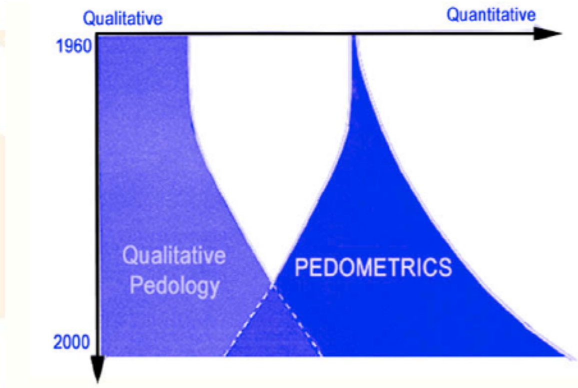


Figure 1 – Temporal evolution of the qualitative and quantitative pedological approaches

The knowledge of soils always benefits from the development of Science and Technology. The dichotomy, "traditional" versus quantitative pedology is just apparent and is not sustained, considering that the prediction of soil classes or properties need both, quantitative and qualitative approaches. That is to say, it doesn't work without qualitative knowledge of soils as well as incorporating it in the models. Nevertheless sound sense is always necessary ahead of everything!

Soil mapping

Perhaps the great change or innovation resides in the paradigm of soil mapping. In that sense, two different approaches exist: the discrete model and the continuous model, thoroughly discussed by several authors (Baize, 1986; Aubert and Boulaine, 1989; Lark and Beckett, 1998).

Thus, traditional soil survey and mapping, is also a model that is based in the concept of soils as a "natural body", the soil as an individual ", with its own characteristics, complete and indivisible. A "natural body" constitutes separable parts of the soil as a "continuum" in the Earth surface (Cline, 1949). These individual parts, or "natural" bodies, composing the "continuum " are called reference units, idealized to support taxonomic systems and soil mapping units. Units of reference depend on limits and concepts imposed by man, to assist to the several outlines of taxonomic soil classification. It is necessary to remark here that we refer to concepts, not real soils. In soil taxonomy, the model used by the pedologist to create soil classes and to determine its limits is an intuitive mental model, deterministic, that assumes the existence of a strong correlation among soil types and their environments. It is an essentially qualitative approach, of empiric-deterministic nature, accomplished on the basis of physiographic analysis of the landscape and interpretation of air photos.

This approach based on the discrete model, describes characteristics of the soils of a certain area, it classifies them in agreement with an effective taxonomic system, establishes the limits among classes defined in the map and it allows to do inferences about the behaviour of the soils with relationship to the use and management. The discrete model assumes that factors of soil formation (*clorpt*) control the distribution of the different soils in the landscape.

Based on these assumed concepts, and, mainly on field experience, the local topographic characteristics and vegetation are interpreted qualitatively as indications of combinations of factors of soil formation and abrupt limits between classes and properties of the soils inside the limits of classes are defined. It is assumed that for those individuals separated by intuitive boundaries, the characteristics or the soil "type" vary abruptly from one to other limits. The space representation has always been discrete, in the form of "Chloropleth" maps, with abrupt limits among the mapping units. In general, each mapping unit is composed of 3-4 soil classes or "type", varying in its proportion of occurrence. The mapping unit, then, can be represented by one soil type only, or for the so called "representative profile" (Soil Survey Staff, 1983). So, the predicted value of an attribute of the soil in any point not sampled will be the value of the typical pedon or the medium value of the mapping unit. The precision of the predicted values will be in that way, subject to the homogeneity of the mapping units, that is to say, to the internal variance of the mapping unit and no estimate of error is presented with the map or the legend.

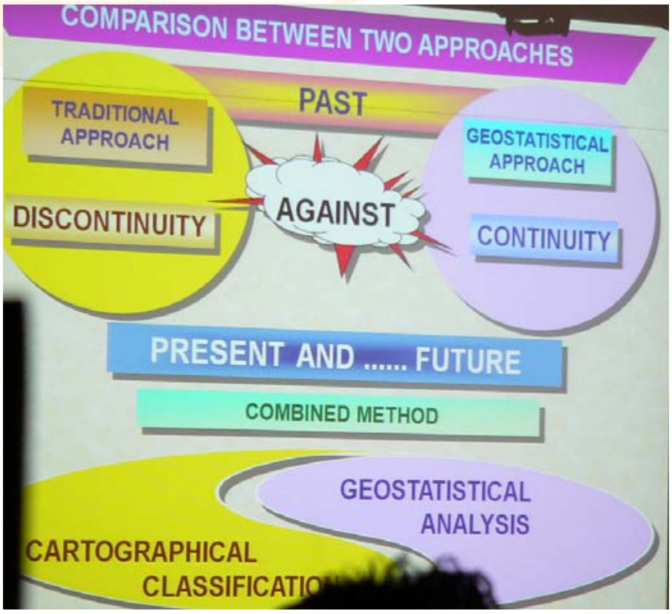
Soil individuals created through that mental model and expressed in discrete form model are artificial, but they are informative and were for a long time, the only option of space representation of a reality well known as continuous.

The Paradigm

This paradigm, however, ignores both the spatial variability of the factors of soil formation and of the own resulting soil, as demonstrated in several works (Burgess & Webster, 1980 and 1984; Burgess et al., 1981; McBratney & Webster, 1981). Therefore, the discrete model is not realistic, because the natural limits in the landscape are above all, gradual and not abrupt as emphasized by Jenny (1941), "Often it is not sufficiently realized that the boundary between soil and environment artificial and that in the two soil scientists have exactly the same enclosure of the soil system in mind".

In the continuous model it is assumed that the soil characteristics vary gradually in space and that it should be represented like this. The geostatistics models were then applied to Soil Science, in an attempt of representing the spatial variability of the soil and its continuous representation, as applied in geological studies on occurrences of minerals.

The search of solutions to the inherent uncertainties to the traditional method and at the same time, to express the complexity of the resulting combination of the factors of soil formation and the progress in the knowledge on the spatial variability and modelling of soil, impelled the development of the quantitative methods applied to Soil Science (or Pedometrics) in the last 30-40 years. A special volume of the journal Geoderma was entirely dedicated to the theme: vol.97, us. 3-4 of 2000, incorporating a selection of



works presented in the Symposium 17 of the XVI World Congress of Science of the Soil, entitled “Advances in Soil Survey using modern tools" and "Symposium Recent Advances in Soil Geostatistics", in Montpellier, France in 1998 and under the auspices of the International Union of Soil Sciences.

Now several quantitative approaches have been used in Pedometrics, such as Geostatistics, Fuzzy Logic and the Artificial Intelligence, in order to describe, to classify and to study the patterns of spatial variation of soils in the landscape, with the additional advantage of knowing the precision and the quality of the information (McBratney et al., 1981; Burgess et al., 1981; Volts & Webster, 1990; De Gruijter & McBratney, 1988; Voltz et al., 1997). Nevertheless, Geostatistics have assumptions that do not always correspond to the reality of the soil variability in the landscape.

McBratney et al. (2000) proposed to ally in a deterministic-stochastic model, pedological knowledge about the formation processes and distribution of the soils in the landscape (*clorpt* method), with quantitative techniques used in Pedometrics, with the purpose of predicting with larger speed and precision, and at a lower cost, the classes and/or properties of the soils for a certain area. This model combining the two methods through the derivation of the non linear correlation among the exogenous environmental factors that influence the processes of soil formation and the geostatiscal methods, both multivariate, developed today for what is known as "**Digital Soil Mapping - DSM**" of soil classes and properties (not to be mistaken with simple digitizing of existing soil maps).

The State of the Art

The state of the art in DSM, including the premises, the beginnings, the methods, the applications and the necessary data to execute the digital soil classes and properties mapping, was fully presented by McBratney et al. (2003) in Geoderma, constituting a reference on the subject. A formal definition, however, was only presented for the scientific community in 2004, during the First Global Workshop on Digital on Soil Mapping, held in Montpellier, France. The selected articles presented in the event were published recently as a book, entitled “Digital Soil Mapping - An introductory perspective “in the series Developments in Soil Science, volume 31 of Elsevier.

In the review on the state of the art in DSM, McBratney et al (2003) presented a generic proposal of protocol for DSM, adapting the theoretical model of Jenny (1941), not with the aim of explaining the variables responsible for the soil-forming process but rather for empirical quantitative descriptions of the relationships between soil and the other spatially-referenced factors so called *scorpan* factors (or environmental co-variates) which are used here as soil spatial prediction functions.

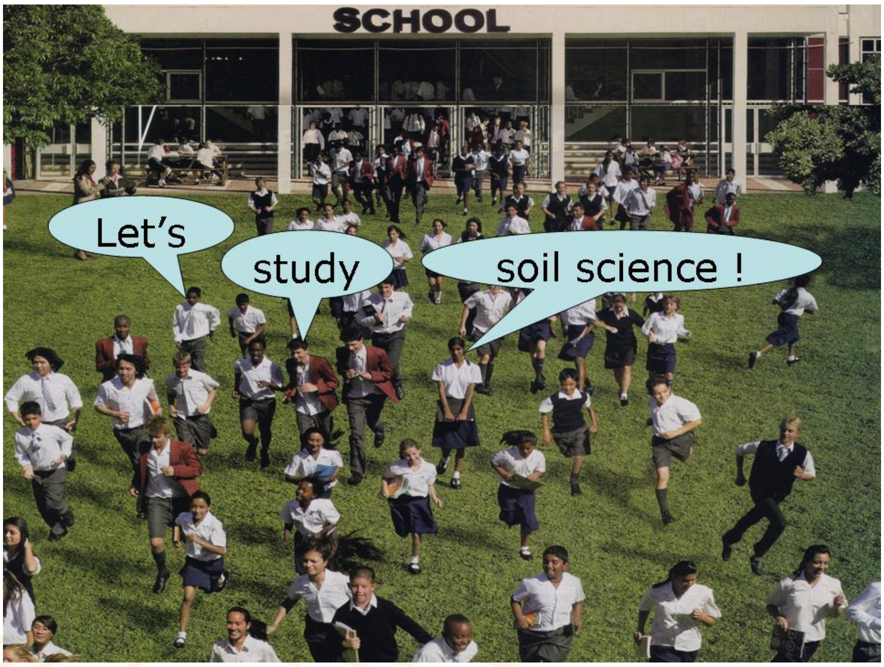
Another novelty, is the possibility to predict the soil starting from soil information (existing maps of soil classes or properties or even “expert knowledge” that is to say, the formal knowledge of the pedologist). The proposed *scorpan* model, together with the space functions for prediction of soils and errors have been showing as a particularly important method, mainly for areas where available data and information of soils are sparse, as it is the case of the Brazilian territory.

The Great Demand

In a global scale, soil research institutions, need precise and geo-referenced data and information for planning purposes, environmental administration and of productivity, as studies of potentialities, environmental risks and impacts, soil resilience and sustainability. They are great demands for pedological information to support studies on the global climatic changes and its impact in the use of the soil.

In general, there is a shortage of soil data in the world, they are limited and dispersed. This has been denominated as a limited spatial soil data infrastructure. That needs, and the growing demand for soil data and information, have put in action the development of digital soil mapping. However, the largest challenge will be to systematize and to understand the existing data and add to them the data produced by the new sensors, to meet the demands of society in relation to the complex threats of environmental changes, food safety, readiness of water, energy and environmental sustainability.

On the theme of shortage and limited infrastructure of spatial soil data, *Embrapa Solos* with the



support of the International Union of Soil Science and the Brazilian Soil Science Society organized The Second Global Workshop Digital on Soil Mapping in Rio de Janeiro. It gathered 75 researchers from 17 countries, to present and to discuss the advances in DSM. Results of that Workshop are being systematized and will be published by Elsevier in the series Developments in Soil Science next year. Another issue of that Workshop was the decision to organize a global consortium of data and information on soils of the world. This consortium is being consolidated in the form of a research project involving world institutions, with the objective of producing a global soil properties digital map, with a spatial resolution of 90 m x 90m, at the same time making available a global soil database in the Internet and some applied results to areas with such specific problems as degradation, contamination and productivity. More information on that project can be found at http://www.globalsoilmap.net.

The International Union of Soil Science was sensitive to this initiative and since 2004 has supported a permanent Working Group on Digital Soil Mapping, linked to the Commissions of Soil Geography (C1.2) and Pedometrics (C1.5). The Working Group (http://www.digitalsoilmapping.org) counts on participants of the whole world and will attempt to accomplish other conferences and training and above all, to implement the project of global digital mapping of soil properties as mentioned above.

In Brazil, several initiatives involving full or partial DSM methodology have been undertaken, involving applications to mapping soil classes and properties such as soil carbon storage, soil available water capacity, development of pedotransfer functions (PTFs) for estimates of soil bulk density, or simply for optimizing traditional soil mapping. Those works can be found in more recent literature or, in the Proceedings of the 2nd GWDSM (2006).

A Big Role

Soil Science has a fundamental role to perform, supplying information and knowledge for strategic decisions and for the establishment of public debate on land use planning and the sustainable use of the soil as a natural resource. For that purpose, it should harmonize and use in a strategic way its qualitative and quantitative attributes, as well as the available technological resources to supply information and requested knowledge to the society, in a fast and low-cost manner. It is necessary to not just supply maps, but databases, information and knowledge in a global extent and easily accessible. In Brazil, we have the exceptional advantage of the knowledge of the tropical soils and of pedologists of great experience that have a fundamental role to carry out the construction of the quantitative models of soil prediction.

Finally, one cannot forget that models are just a simplified representation of a complex reality, but that they are indispensable for understanding the real world and the evolution of the knowledge.

References

SECOND GLOBAL WORKSHOP ON DIGITAL SOIL MAPPING, 2006, Rio de Janeiro. Anais… Rio de Janeiro: Embrapa Solos, 2006. 1 CD-ROM.

AUBERT, G.; BOULAINE, J. Contributions de certains pédologues français à l'évolution des concepts pédologiques utilisés en cartographie. Science du Sol, Versailles, v. 27, n. 4, p. 395-411, 1989.

BAIZE, D. Couvertures pédologiques, cartographie et taxonomie. Science du Sol, Versailles, v. 24, n. 3, p. 227-243, 1986.

BURGESS, T. M.; WEBSTER, R. Optimal interpolation and isarithmic mapping of soil properties, I: the semivariogram and punctual kriging. Journal of Soil Science, Oxford, v. 31, p. 315-331, 1980.

BURGESS, T. M.; WEBSTER, R. Optimal sampling strategies for mapping soil types, I: distribution of boundary spacings. Journal of Soil Science, Oxford, v. 35, p. 641-654, 1984.

BURGESS, T. M.; WEBSTER, R.; MCBRATNEY, A. B. Optimal interpolation and isarithmic mapping, IV: sampling strategies. Journal of Soil Science, Oxford, v. 32, p.643-659, 1981.

CLINE, M. G. Basic principles of soil classification. Soil Science, Baltimore, v. 67, p. 81-91, 1949.

DE GRUIJTER, J. J.; MCBRATNEY, A. B. A modified fuzzy k-means method for predictive classification. In: BOCK, H. H. (Ed.). Classification and related methods of data analysis. Amsterdam: Elsevier, 1988. p. 97-104.

DEVELOPMENTS in quantitative soil resource assessment (Pedometrics´98). Geoderma, Amsterdam,.v. 97, n. 3-4, 2000. Special issue

ESTADOS UNIDOS. United States Department of Agriculture. Soil Survey Staff. Soil survey manual. Washington, DC: USDA, 1993.

JENNY, H. Factors of soil formation: a system of quantitative pedology. New York: McGraw-Hill, 1941.

LAGACHERIE, P.; MCBRATNEY, A. B. Spatial soil information and spatial soil inference system: perspectives for digital soil mapping. In: LAGACHERIE, P.; MCBRATNEY, A. B.; VOLTZ, M. (Ed.). Digital soil mapping: an introductory perspective. Amsterdam: Elsevier, 2007. p. 3-22.

LAGACHERIE, P.; MCBRATNEY, A. B.; VOLTZ, M. (Ed.). Digital soil mapping: an introductory perspective. Amsterdam: Elsevier, 2007. 636 p.

LARK, R. M.; BECKETT, P. H. T. A geostatistical descriptor of the spatial distribution of soil classes, and its use in predicting the purity of possible soil map units. Geoderma, Amsterdam, v. 83, n. 3-4, p. 243-267, 1998.

MCBRATNEY A. B. MENDONÇA-SANTOS, M. L.; MINASNY, B. On digital soil mapping. Geoderma, v. 117, p. 3-52, 2003.

MCBRATNEY, A. B.; ODEH, I. O. A.; BISHOP, T. F. A..; DUNBAR, M. S.; SHATAR, T. M. An overview of pedometric techniques for use in soil survey. Geoderma, Amsterdam, v. 97, n. 3-4, p. 293-327, 2000.

MCBRATNEY, A. B.; WEBSTER, R. Spatial dependence and classification of soil along a transect in northeast Scotland. Geoderma, Amsterdam, v. 26, p. 63-82, 1981.

MCBRATNEY, A. B.; WEBSTER, R.; BURGESS, T. M. The design of optimal schemes for local estimation and mapping of regionalized variables, I: theory and methods. Computers & Geosciences, Oxford, v. 7, p. 331-334, 1981.

MCBRATNEY, A.B., ODEH, I.O.A., BISHOP, T.F.A., DUNBAR, M.S., SHATAR, T. An overview of pedometric techniques for use in soil survey. Geoderma, Amsterdam, v. 97, p. 293–327, 2000.

VOLTZ, M., LAGACHERIE, P., LOUCHART, X.,. Predicting soil properties over a region using sample information from a mapped reference area. European Journal of Soil Science, v. 48, p. 19-30; 1997.

VOLTZ, M., WEBSTER, R.,. A comparison of kriging, cubic splines and classification for predicting soil properties from sample information. Journal of Soil Science, v. 41, p. 473-490, 1990.

WEBSTER, R. Quantitative spatial analysis of soil in the field. Advances in Soil Science, New York, v. 3, p. 1-70, 1984.

YAALON, D. H. Conceptual models in pedogenesis: can soil forming functions be solved? Geoderma, Amsterdam, v. 14, p. 189-205, 1975.

Soil spatial prediction with Markov random fields

Gerard Heuvelink,
Wageningen University and Research Centre, The Netherlands.
Email: gerard.heuvelink@wur.nl

Random cycling

More than 20 years ago, when I was still a student, on boring Sunday afternoons, my friend Theo and I sometimes went for an unusual bicycle tour. It worked as follows. At each crossing, we stopped and checked the valves of our front wheels. If both were up, we would turn right. If both were down, we went left. In all other cases we went straight on. What we did was a variant of the famous *random walk* (http://en.wikipedia.org/wiki/Random_walk). It could take us to strange places, sometimes it would force us to cycle the same street multiple times, and prevent us from escaping from an ugly neighbourhood, but the nice thing about it was that we never knew in advance where we would end up. Already then I was intrigued by uncertainty!

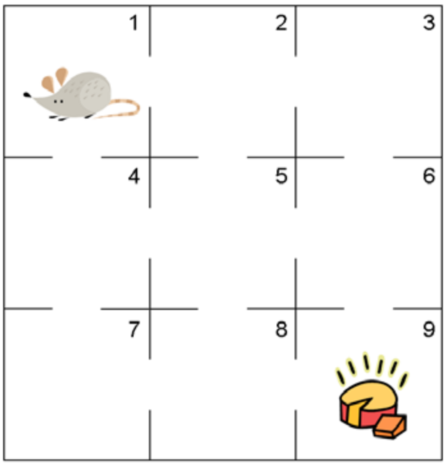


Figure 1. How soon will the mouse in room 1 find the cheese in room 9?

Another example of the random walk is presented in Figure 1. At time t_0 the mouse is in room 1 in the top left corner. At each time step it moves to a neighbouring room or stays in the room where it was, where each room has equal probability of being chosen. Thus, at time t_1 the probability about the position of the mouse is

$$p_1 = [1/3\ 1/3\ 0\ 1/3\ 0\ 0\ 0\ 0\ 0]^T.$$

At time t_2 it is

$$p_2 = [5/18\ 7/36\ 1/12\ 7/36\ 1/6\ 0\ 1/12\ 0\ 0]^T.$$

We can continue and calculate the probability for any future time by applying the relationship:

$$p_{k+1} = P \cdot p_k \qquad (1)$$

where:

$$P = \begin{bmatrix} \frac{1}{3} & \frac{1}{4} & 0 & \frac{1}{4} & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{3} & 0 & \frac{1}{5} & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{4} & \frac{1}{3} & 0 & 0 & \frac{1}{4} & 0 & 0 & 0 \\ \frac{1}{3} & 0 & 0 & \frac{1}{4} & \frac{1}{5} & 0 & \frac{1}{3} & 0 & 0 \\ 0 & \frac{1}{4} & 0 & \frac{1}{4} & \frac{1}{5} & \frac{1}{4} & 0 & \frac{1}{4} & 0 \\ 0 & 0 & \frac{1}{3} & 0 & \frac{1}{5} & \frac{1}{4} & 0 & 0 & \frac{1}{3} \\ 0 & 0 & 0 & \frac{1}{4} & 0 & 0 & \frac{1}{3} & \frac{1}{4} & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{5} & 0 & \frac{1}{3} & \frac{1}{4} & \frac{1}{3} \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{4} & 0 & \frac{1}{4} & \frac{1}{3} \end{bmatrix}$$

The element on row i column j of the transition matrix P represents the probability that the mouse moves from room j to room i in one time step. Note that each of the columns of P sums to one. The first time that the mouse can reach room 9 and finds the cheese is at time t_4, with probability 11/360 (a meagre 3 per cent). If we let time go to infinity then the initial position of the mouse becomes irrelevant and we arrive at the stationary probability distribution. We can compute it by running Equation (1) many times (25 runs already gets us quite close to the final solution), but a more elegant solution is to solve the identity $p_s = P \cdot p_s$ directly. If we rewrite it as $(I-P) \cdot p_s = 0$ then we can obtain p_s by taking the eigenvector associated with the zero eigenvalue of the matrix I−P, and standardising it such that its elements sum to one. This reveals that the stationary probability distribution is given by $p_s = [3/33\ 4/33\ 3/33\ 4/33\ 5/33\ 4/33\ 3/33\ 4/33\ 3/33]^T$, which assigns the largest probability to the centre room and the smallest to the corner rooms, as expected.

We can also simulate a random walk by the mouse, using a pseudo-random number generator. At time t_0 we begin in cell 1, and draw a random number between 0 and 1 from the uniform distribution (we can for example do this with the Excel function RAND()). If the number is smaller than 1/3, the mouse stays in cell 1. If it is greater than 1/3 but smaller than 2/3, it moves to cell 2. If the number drawn is greater than 2/3, the mouse moves to cell 4. In the new cell, we kind of repeat the procedure and carry on for as many time steps as we like. In this way many random walks of the mouse can be easily simulated. The sequence of rooms that the mouse visits is a realisation of a *Markov chain* (http://en.wikipedia.org/-wiki/-Markov_-chain), named after the Russian mathematician *Andrey Markov* (Figure 2).

Figure 2. Andrey Markov, Russian mathematician who lived from 1856 to 1922.

The Markov property

Now let us take a closer look at Equation (1). It says that the probability distribution at the next time step only depends on the transition probabilities and the state (or probability) at the current time. Apparently, it does not matter where the mouse was at previous times, it is only the current state that is important. In other words, the mouse has no memory. This is in essence the *Markov property*, which states that the "*conditional probability distribution of future states of the process, given the present state and all past states, depends only upon the present state and not on any past states, i.e. it is conditionally independent of the past states given the present state*".

It can be formulated as:

$$P(X_{t+1}|X_u=x_u, u=0,1,\ldots,t) = P(X_{t+1}|X_t=x_t) \quad (2)$$

where X_t is the state of the system at time t and where a vertical bar denotes conditional probability. The Markov property is often used to model real-world processes because it greatly simplifies the statistical analysis and is sufficiently realistic for many situations. Examples can be found in meteorology (rainfall, temperature), hydrology (flood forecasting), econometrics (interest rate, stock market) and computer science (waiting queues).

Most applications of the Markov property consider a process that varies over time. But it can also be used for variables that vary over space (or space and time). This brings us to so-called *Markov random fields*. The difference with Markov chains is that we have moved from one to two dimensions or higher, and that there is no sense of (causal) direction anymore. By this I mean that the value of the variable of interest at some location will be influenced by the variable at all of its neighbours, not by those in one direction only. The Markov property in a spatial setting thus states that:

$$P(X_s|X_u=x_u, \text{all } u\neq s) = P(X_s|X_u=x_u, u \text{ in neighbourhood of } s) \quad (3)$$

where s is a spatial coordinate. The smallest meaningful neighbourhood in the gridded 2D environment would be composed of the four immediate neighbours. Thus, in Figure 3 the value of the target variable in the dark grey cell would be conditionally independent of the values in the white cells, given the values in the light grey cells.

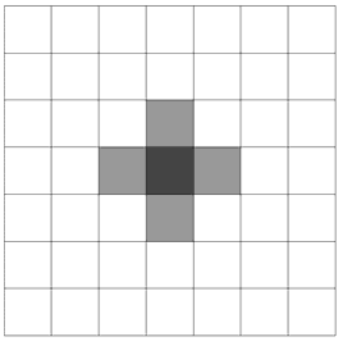


Figure 3. Illustration of the Markov property in space using the von Neumann neighbourhood. The value in the dark grey cell is conditionally independent of the values in the white cells, given the values in the grey cells.

Note that this does not imply that the variables in the dark grey and white cells are independent! They are only *conditionally* independent, given the variables in the light grey cells. Let us take the first order spatial autoregressive model as an example:

$$X[i,j] = \alpha \times \{X[i-1,j]+X[i+1,j]+X[i,j-1]+X[i,j+1]\} + e[i,j] \quad (4)$$

where i and j are row and column numbers and e is a white noise (pure nugget) process. This model, already presented in the classic 1954 Biometrika paper of Peter Whittle (which you may be able to download from http://biomet.oxfordjournals.-org/-content/-vol41/-issue3-4/index.dtl), satisfies the Markov property but will have a strong spatial dependence with variogram ranges that are many cell sizes large when the value for α is close to 0.25.

Simulating soil type maps

We can also use the Markov random field approach for categorical variables, such as soil type. This means that we assume that the soil type at some location is conditionally independent of the soil type at locations outside its neighbourhood, given the soil type in the neighbourhood. We must define the conditional probability distribution to characterise the dependence of soil type at some location on that in its neighbourhood. For example, suppose that there are only three soil types A, B and C, then it could read:

$$P(X[i,j]=A|X[i-1,j]=A,X[i+1,j]=B,X[i,j-1]=C,X[i,j+1]=A) = 0.84$$
$$P(X[i,j]=B|X[i-1,j]=A,X[i+1,j]=B,X[i,j-1]=C,X[i,j+1]=A) = 0.06$$
$$P(X[i,j]=C|X[i-1,j]=A,X[i+1,j]=B,X[i,j-1]=C,X[i,j+1]=A) = 0.10 \quad (5)$$

where X represents soil type in this case. Note that Eq. (5) only specifies the conditional probability for one particular neighbourhood combination (out of a total of 3^4=81). In practice, simplifying assumptions will therefore be needed to be able to estimate the many conditional probabilities from available calibration data. Once this is done, the Markov random field model can be used to predict soil type at unobserved locations and to simulate from its probability distribution. This is not a trivial problem, because to simulate (or predict) soil type at cell [i,j] using a conditional probability distribution such as Eq. (5) requires that we know its value at the neighbouring cells, but to simulate soil type at the neighbouring cells we need to know the soil type at the centre cell. The one needs the other and the other needs the one. How to solve this problem? One possibility is to remove all feedbacks and enforce a sense of direction using a Markov chain approach (Li et al. 2004, Li 2007). A similar approach is used in Wu et al. (2004) to model soil structure at the pore scale (spatial extent of 1 cm and spatial resolution of 20 µm). However, personally I am not in favour of that because the underlying model denies a fundamental property of spatial data, namely that influence goes in all directions (for some phenomena, such as water flow, the directional assumption may be valid). In fact, we also do not need to make such assumptions because we can use the *Gibbs sampler* to simulate from Markov random fields such as characterised by Eq. (5).

The Gibbs sampler (http://en.wikipedia.org/wiki/Gibbs_sampling) is a so-called Markov Chain Monte Carlo technique. MCMC techniques are much used in Bayesian statistics, because they help tackle problems that cannot be solved analytically (which holds for most Bayesian problems). The method is extremely simple, but computationally demanding:

1. Start with an (arbitrary) initial soil map (take some value for X[i,j], for all i and j);
2. Repeat the following steps M times:
   a. Visit all grid cells of the map one by one;
   b. For each cell [i,j], replace X[i,j] by a random draw from the conditional distribution P(X[i,j]|X[k,l], all k and l except for the case where k=i and j=l), where conditioning is done on the most recently simulated values;
3. Provided M is large enough, the final map will be a realisation from the joint probability distribution of X.

So far, all that I have done myself using these techniques is some experimenting, see Figure 4 below for one of the examples I have been working on. The figure reports a case where an existing soil map is used as soft information, which is incorporated by letting the conditional distributions

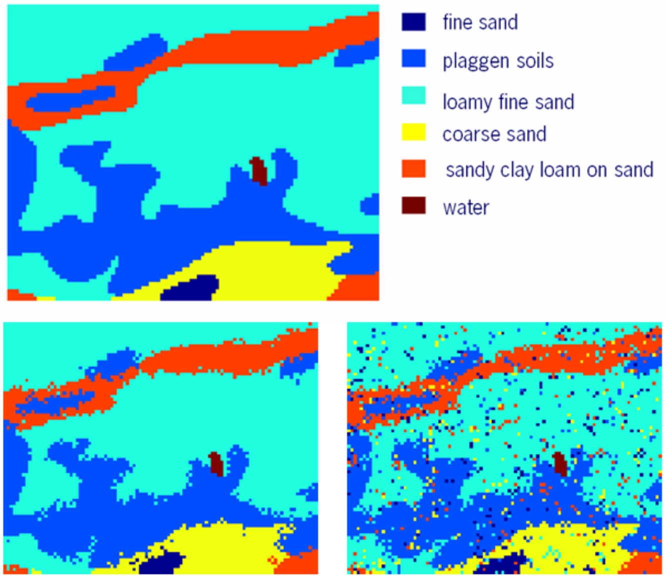


Figure 4. Soil type simulation using the Gibbs sampler. Top: 1:50,000 soil map for a 2.5 x 2.0 km area in the North-East of the Netherlands. Bottom: Example simulations generated using the Gibbs sampler, whereby conditional probabilities are influenced by neighbouring values and by the deterministic soil map. Probabilities for the right image were chosen such that inclusions could occur.

depend on the soft soil type. In this example the conditional probabilities are strongly controlled by the soft soil type, which explains why the simulated maps have a similar pattern as the existing map. Hard information (point data) could easily be incorporated by forcing the 'simulated' soil type at sampling sites to be equal to the observed soil type (similar to what is done in geostatistics when using conditional sequential simulation). The Gibbs sampler can also be used to interpolate soil type by 'averaging' many independent simulated realities. It is a flexible method that has a lot of potential and can compete with and perhaps even outperform alternative methods such as indicator geostatistics and Bayesian Maximum Entropy (Brus et al. 2007). Am I too optimistic? Perhaps I am, because there are two main problems: calibration (see above) and check on *statistical compatibility*.

Check on statistical compatibility

We all know that in geostatistics we cannot use just any function to fit a variogram to an experimental variogram. The variogram function must be negative semidefinite, to ensure that the variance of any linear combination of the variable is non-negative. A similar thing happens when we define conditional probability distributions as in Eq. (5). We must ensure that the individual conditional distributions are *compatible* with regard to determining the joint distribution (Banerjee et al., 2004, page 76). I did not realise this before and the annoying thing (perhaps not to everyone!) is that the Gibbs sampler will run and produce results regardless of whether a valid model was defined or not.

To see that indeed there may be a serious problem, let us consider a simple example involving only two binary variables A and B. Let the conditional probabilities be given by:

$$\begin{aligned} &P(A=0|B=0) = 0.1 \\ &P(A=0|B=1) = 0.9 \\ &P(B=0|A=0) = 0.9 \\ &P(B=0|A=1) = 0.1 \end{aligned} \qquad (6)$$

From this we can compute:

$$\begin{aligned} P(A=0) &= P(A=0, B=0) + P(A=0, B=1) = \\ &= P(A=0|B=0)\times P(B=0) + P(A=0|B=1)\times P(B=1) = \\ &= 0.1\times P(B=0) + 0.9\times P(B=1) \geq 0.9\times P(B=1) \end{aligned}$$

In the same way we obtain $P(B=1) \geq 0.9\times P(A=1)$, $P(A=1) \geq 0.9\times P(B=0)$ and $P(B=0) \geq 0.9\times P(A=0)$.

We also know that:

$$P(A=0, B=0) = P(A=0|B=0)\times P(B=0) = 0.1\times P(B=0)$$
$$P(A=0, B=0) = P(B=0|A=0)\times P(A=0) = 0.9\times P(A=0)$$

By combining all this, we get:

$$0.1\times P(B=0) = P(A=0,B=0) = 0.9\times P(A=0) \geq (0.9)^4\times P(B=0) = 0.6561\times P(B=0)$$

This cannot be true (unless $P(B=0) = 0$, but that cannot be true either), so apparently our initial model Eq. (6) was incompatible. We had defined a set of conditional distributions for which there exists no joint distribution. The Gibbs sampler would produce rubbish.

How can we make sure that we have defined a compatible model? One way is to stick to models which have been proven to be compatible (in geostatistics we always use authorised variogram models such as the spherical and exponential for the same reason), but unfortunately the existing models, such as the Potts model (Banerjee et al. 2004, page 84), are pretty restrictive. One other solution that I would like to pursue is to construct the joint distribution from the conditionals by using Brook's lemma (Banerjee et al. 2004, page 77):

$$P(X_{s_1},\dots,X_{s_n}) = \frac{P(X_{s_1} \mid X_{s_2},\dots,X_{s_n})}{P(X^*_{s_1} \mid X_{s_2},\dots,X_{s_n})}\cdot\frac{P(X_{s_2} \mid X^*_{s_1},X_{s_3},\dots,X_{s_n})}{P(X^*_{s_2} \mid X^*_{s_1},X_{s_3},\dots,X_{s_n})}\cdot \cdots\frac{P(X_{s_n} \mid X^*_{s_1},\dots,X^*_{s_{n-1}},X_{s_n})}{P(X^*_{s_n} \mid X^*_{s_1},\dots,X^*_{s_{n-1}},X_{s_n})}\cdot P(X^*_{s_1},\dots,X^*_{s_n}) \quad (7)$$

where X is the state of the system (soil type in our case) as before, the s_i are locations and where X^* is just an arbitrary outcome of the state. Using Eq. (7) we can express the joint probabilities for all possible values of X as a function of $P(X^*_{s1}, \dots, X^*_{sn})$, from which we then can derive the full distribution by introducing a normalising constant so that the probabilities add up to one. When this can all be done and yields a proper joint distribution, then we have proven by construction that we had defined a compatible system of conditional probabilities. Does this sound like a sensible approach?

Well, you can see that I am still in the exploratory phase, haven't found solutions to some of the fundamental problems yet. I can use all the help I can get!

Gerard

References

Banerjee, S., B.P. Carlin and A.E. Gelfand (2004). *Hierarchical Modeling and Analysis for Spatial Data*. Boca Raton: Chapman & Hall.

Brus, D.J., P. Bogaert and G.B.M. Heuvelink (2007), Bayesian Maximum Entropy prediction of soil categories using a traditional soil map as soft information. *European Journal of Soil Science* (doi: 10.1111/j.1365-2389.2007.00981.x).

Li, W. (2007), A fixed-path Markov chain algorithm for conditional simulation of discrete spatial variables. *Mathematical Geology* 39, 159-176.

Li, W., C.R. Zhang, J.E. Burt, A.X. Zhu and J. Feyen (2004), Two-dimensional Markov chain simulation of soil type spatial distribution. *Soil Science Society of America Journal* 68, 1479-1490.

Wu, K. N. Nunan, J.W. Crawford, I.M. Young and K. Ritz (2004), An Efficient Markov Chain Model for the Simulation of Heterogeneous Soil Structure. *Soil Science Society of America Journal* 68, 346-351.

Upcoming Events

Global Workshop on High Resolution Digital Soil Sensing & Mapping. 5-8 February 2008. Sydney, Australia. http://www.digitalsoilmapping.org

European Geophysical Union, EGU 2008. Vienna, Austria, 13 – 18 April 2008. The Digital Soil Mapping WG will have a session titled: **Digital soil mapping: novel approaches to the prediction of key soil properties for modelling physical processes** http://meetings.copernicus.org/egu2008/

Accuracy 2008, The eighth symposium on spatial accuracy assessment in natural resources and environmental sciences. 25-27 June 2008. Shanghai, China http://2008.spatial-accuracy.org/

International Symposium on Spatial Data Handling. 23-25 June 2008 - Montpellier, France. http://sdh-sageo.teledetection.fr

The First International Conference on HYDROPEDOLOGY. 28-31 July 2008. Penn State University, USA. hydropedology.psu.edu

EUROSOIL 2008, 25 - 29 August 2008, Vienna, Austria. www.ecsss.net

The **3rd Global Workshop on Digital Soil Mapping**. Utah State University, Logan, USA, 30 Sept—3 October 2008. http://www.digitalsoilmapping.org

International Geostatistics Congress Santiago, Chile 1-5 Dec 2008 http://www.geostats2008.com/

Declining impact of (soil) geostatistics?

Budiman Minasny, Alfred Hartemink & Alex. McBratney

At the Pedometrics conference in Tübingen Tom Hengl showed us a graph from a paper in *Scientometrics* on bibliometric analysis of geostatistical papers (Zhou et al., 2007). The multivariate analysis was based on 2866 publications in the ISI database. The articles spanned 107 subject categories and about 13% were on the soil geostatistics. Highest mean impact factor and annual citation per publication were R.M. Lark and P. Goovaerts whereas A. Stein was the most productive author. Geoderma ranked third in the top 10 of the most productive geostatistics journal and is together with the Soil Science Society of America Journal in the top 5 of the most cited journals. To us, this shows that pedometrics have greatly contributed to geostatistics.

The Zhou et al. paper was not about soil geostatistics but showed that the number of papers and citations is growing over time (1967-2005), but the "impact" of the papers is going down. The impact was measured by average citations per year and mean impact factors and the mean impact factor was calculated based on journals impact factors calculated in 2005. As we all know, there are some problems with these two measures: the first is that citation rate is not constant over time and mostly follows a sigmoidal function, usually it takes more than 1 for a paper to get cited. The second is that there is a large interannual fluctuation in the impact factor of a journal and that the journal's impact factor holds no relation to individual papers. In many journals the impact factor is determined by only a handful of papers.

The impact factor for a journal (let's say in 2006) is calculated as follows:

A = Number of citations in 2006 to articles published in 2004-2005

B = Number of papers published in 2004-2005.

The impact factor for year 2006 is = *A*/*B*.

The impact factor can be calculated for different periods e.g. with a two or four year window. Usually, the two-year impact factors are reported and the factors for the preceding year are mostly reported in June (that's when most publishers, some journal editors and perhaps even some authors get excited).

A better measure for the "impact" of a topic is the impact factor itself, which can be calculated for a topic in the same manner as a journal's impact factor. We conducted our own bibliometric analysis focusing on soil geostatistics. The ISI Web of Science was used searching for papers with keywords: soil* and [geostatistic* or kriging or variogram*]. This may not be a complete list, but will give a good indication of the bibliographic trend.

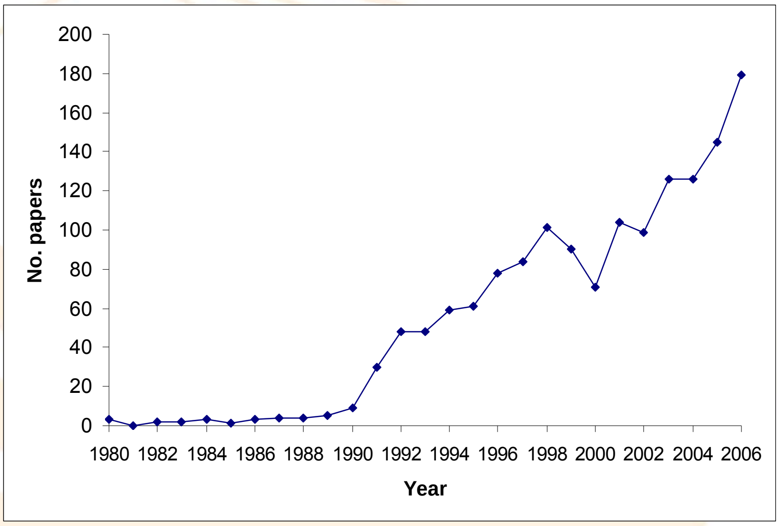


Figure 1. Number of publications in soil geostatistics over time.

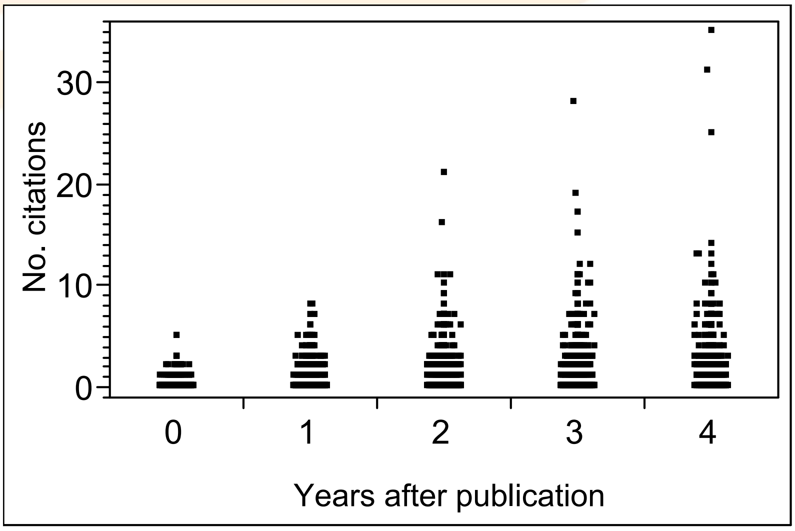


Figure 2. No. citations as a function of number of years after the papers were published.

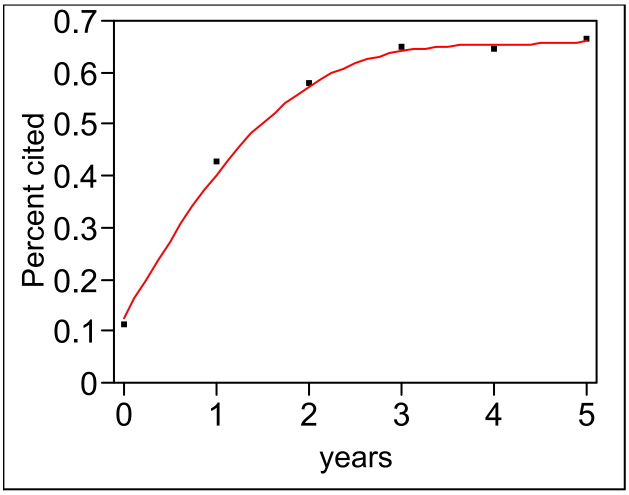


Figure 3. Distribution of percentage of self-citations in Pedometrics papers.

There are 1485 papers published between 1980 and 2006. Fig. 1 shows the increasing number of papers on soil geostatistics and since 1999 the rate of increase is about 13 papers per year. The number of papers drastically increased after 1990. The number before 1991 is underestimated because the ISI databases contain no abstracts prior to 1991. Therefore, we only present impact factor calculations for the years 1991 to 2006 (and we also wonder how Zhou et al. obtained their data).

Figures 2 and 3 show the number of citations as a function of the number of years after the paper was published. For soil geostatistics (and generally in soil science), only 10% of the paper will be cited in the same year as it is published, and after 1 year 42% will be cited. It takes at least 3 years to get two-third of the citations. About 35% of the papers are never cited (*nil desperandum* folks).

The paper with the highest immediacy factor of 4 (number of citations of an article in that year) is from pedometricians:

De Gruijter, J.J.; Walvoort, D.J.J.; van Gaans, P.F.M., 1997. Continuous soil maps - A fuzzy set approach to bridge the gap between aggregation levels of process and distribution models. Geoderma, 77: 169-195

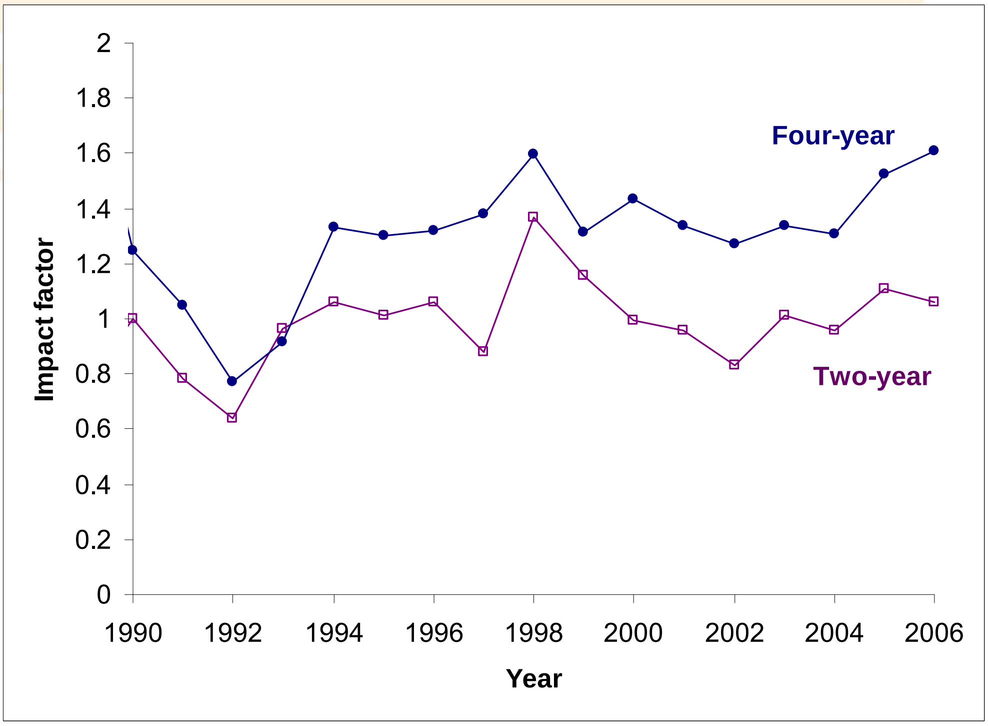


Figure 4. Two-year and Four-year impact factors for soil geostatistics papers.

We calculated the two-year and four-year impact factors (figure 4). For both there that there is a peak in the 1998, and a slight drop until 2002, where it starts to increase steadily again. The impact factor is around 1, which means on average expect only one citation per year, and for a soil science journals is not high.

Our results showed that soil geostatistics research and “impact” (as measured by citations) is at a steady rate, and seems to contradict the findings of Zhou et al. which showed a declining trend since 2002. That may be good news but overall the impact factor of soil geostatistics is not high.

Tomi has promised that he will do a more thorough analysis on geostatistical analysis of geostatistical publications. As always, we are keen to see that!

References

Zhou, F., Guo, H.C., Ho, Y.S., Wu, C.Z., 2007. Scientometric analysis of geostatistics using multivariate methods. Scientometrics 73, 265-279 . DOI: 10.1007/s11192-007-1798-5.

Something trivial on the topic of self-citation

In the last issue of Pedometron, Budiman et al. wrote an article on the topic of the h index and self-citation; the latter being a means of boosting one's own index. This led me to consider the perils of doing so, and the need to highlight these. I thought that a limerick might be well-suited to the task, and set about composing one. Much of the inspiration came to me on my cycle ride home, and interdisciplinary collaboration enabled its completion.

Here it is:

There was a researcher called Wendy

Who suffered from h-index envy

So to make it alright

All her papers she'd cite

And grant favours to others as necess'ry

Those wishing to cite this salutory tale should use the full, bibliographic reference below*.

Barry Rawlins & Ruth Parry

* My interdisciplinary colleague (RP) refers to this as ironic metacommentary.

Two Pedometrical Haiku (more or less, perhaps Mathemaku) to mark 70 years of nested sampling in soil science*

I The balanced case

Find the first variance;

The others unfold in order —

Model-free.

II The unbalanced case

No symmetry in Gower's ugly coefficients —

So,

Use REML.

*Youden, W.J. & Mehlich, A. 1937. Selection of efficient methods for soil sampling. *Contributions of the Boyce Thompson Institute for Plant Research*, **9**, 59-70.

Murray

Vacant Positions

For More info see www.pedometrics.org

Senior Scientist (tenure track) - Soil Landscape Modelling We are looking for a scientist with excellent skills in spatio-temporal soil landscape modelling. **Deadline to apply:** 31/01/2008. **Requirements:** (i) a PhD in soil science, other related geosciences, or agricultural sciences, with a focus on spatio-temporal soil modelling, (ii) more than 5 peer-reviewed papers in ISI-cited scientific journals, (iii) experience in fund raising, as well as personnel management. **For more info:** http://http://www.zalf.de/home_zalf/aktuelles/aktuelles_e **Contact:** Michael Sommer (sommer@zalf.de)

Postdoctoral Fellow - High Resolution Digital Soil Mapping (6 months). The pedometrics group at the University of Sydney is looking for a young early career researcher with a PhD in soil science or related areas. Skills in pedometrics, and on-the-go sensors will be beneficial. For more info contact: Alex McBratney (alex.mcbratney@usyd.edu.au)

Post-Doc - Geospatial Digital Soil Modeling / Soil Carbon Sequestration (University of Florida). A highly motivated applicant is sought to conduct research on "Rapid Assessment and Modeling of Changes in Soil Carbon Storage and Turnover in a Southeastern U.S. Landscape (Florida)". Deadline: 30/01/2008. For more info contact: Sabine Grunwald http://grunwald.ifas.ufl.edu

PhD-position (4 yrs) Modeling pedogenesis and hydrology of sandy Flanders over the last 15000 years for geoarchaeological reconstruction. **Deadline to apply:** 30/01/2008. **Requirements:** MSc-degree in bio-engineering sciences, soil science, geology, physical geography or equivalent. Knowledge of processes of pedogenesis, soil chemistry, geostatistical methods and computer programming is an asset. Willingness to move the Ghent, Belgium. **More info:** http://https://webster.ugent.be/vacatures/AAPWP/WE13Soil **Contact:** Peter Finke (peter.finke@ugent.be)

Help!

Looking for the best Pedometrics Haiku* 俳句

Submit your Haiku to vchair@pedometrics.org

* http://en.wikipedia.org/wiki/Haiku

Alex's Preferred Pedometrics Papers II

My top six most preferred pre-geostatistical pedometrical papers (2)

Studies in Soil Cultivation

I, The evolution of a reliable dynamometer technique for use in soil cultivation experiments. Bernard A. Keen and William B. Haines (1925) Journal of Agricultural Science, 15: 375-386.

II. Test of soil uniformity by means of dynamometer and plough. William B. Haines and Bernard A. Keen (1925) Journal of Agricultural Science, 15: 387-394.

III. Measurements on the Rothamsted classical plots by means of dynamometer and plough. William B. Haines and Bernard A. Keen (1925) Journal of Agricultural Science, 15: 395-406.

IV. A new form of traction dynamometer. W.B. Haines and B.A. Keen (1928) Journal of Agricultural Science, 18: 724-732.

In 1925, before the age of digital computer and electronic sensors, Bernard Keen (1890-1981) and William Haines conceived and built the first on-the-go soil strength sensor, made the measurements and made the first high resolution digital soil map, and discovered the reality of soil spatial variation. The papers can be downloaded here: http://www.usyd.edu.au/su/agric/acpa/papers/theclassics.htm For the biography of Keen, see Pereira (1982) http://tinyurl.com/2evp5x

The Rationale — *Cultivation and soil strength*

The rationale behind inventing the on-the-go soil strength meter was to measure the relative draughts of cultivation implements, to assess the impact of tillage/ cultivation on soil structure, and to quantify short-range soil variation. The rationale has not really changed since.

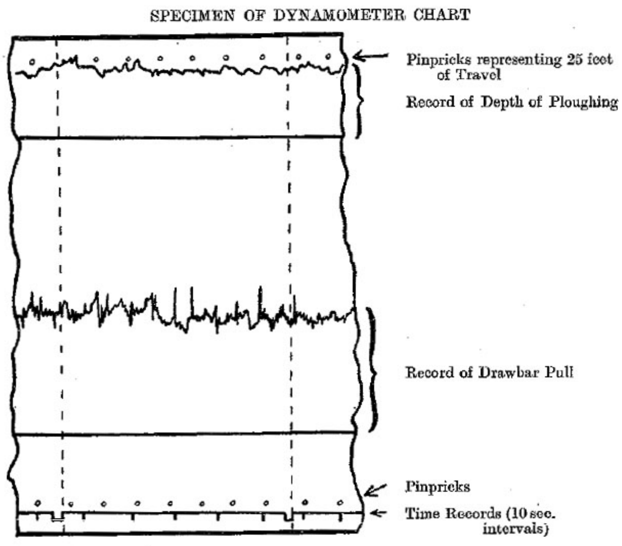
The Sensor— *A (pre-electronic) design for a soil draft force sensor*

Keen and Haines (Part I) designed an instrument that recorded the resistance that must be overcome by the applied force of drawing the ploughing implement through soil. The first prototype was a simple spring balance, the instrument is an automated recording pressure gauge, indicating the pressure produced in an oil system by the pull between the two joints to which the instrument is hitched.

Later (Part IV) Haines & Keen designed a portable version called the 'Rothamsted dynamometer' which was marketed by the Cambridge Instrument Company.

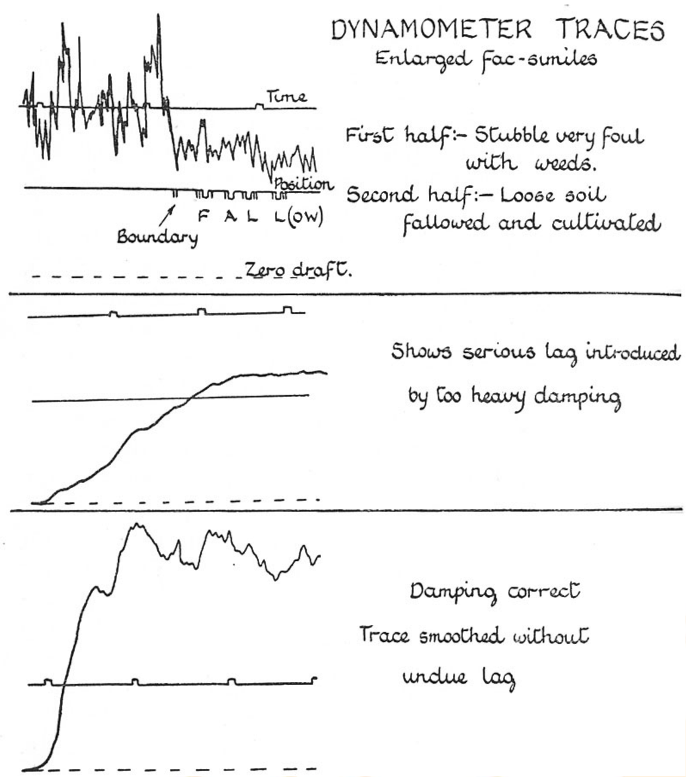
Data logging — *Analogue data recording*

Recording was on a chart of a continuous strip with parallel records of drawbar pull (related to soil strength), time interval, distance travelled, and depth of ploughing.



Data analysis — Spatial variation and data filtering

Haines & Keen (1925) was first to quantitatively document the considerable amount of short-range soil spatial variation, invalidating the assumption that soil in any sense was uniform. They also postulated the origin of soil variation, "the key is to be found in the observation that they remain constant from season to season. They are almost certainly the resultant of the age-long soil-forming processes."



To handle the data, rather than using sophisticated time-series analysis or Kalman filtering, Haines & Keen found that the data could be smoothed during recording by using oil in the sensing system with the appropriate viscosity. The amount of detail in the record could then be controlled by the viscosity of the oil used. A thin oil gave minute detail, while thick oil gave any desired degree of smoothing but with an undesirable time-lag. A compromise was needed to give a reasonable detail without an appreciable time-lag.

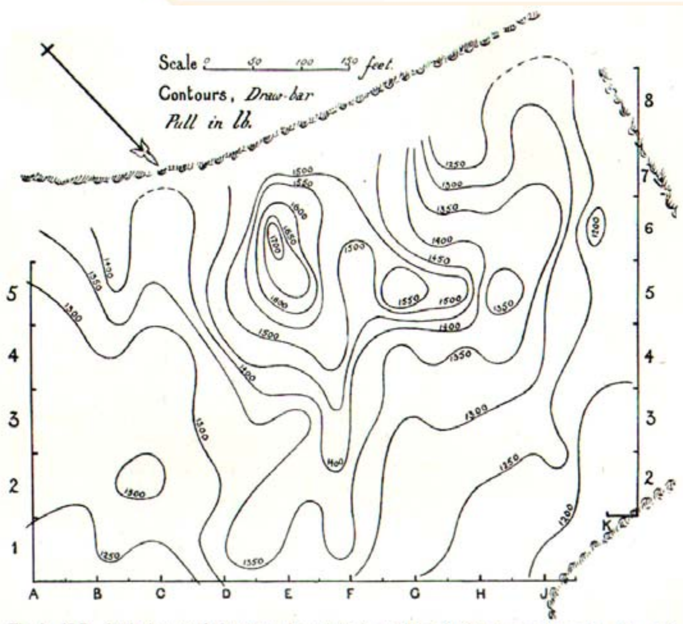


Fig. 1. N.B. Strip A extends between the points A and B, strip B between the points B and C, and so on. The left half of each strip is the one denoted in Table 1 as ploughed "down," and the right half "up."

The Product- *The first high resolution digital soil map*

The first experiment was on Sawyers field with an area of 2.4 hectares that had received uniform treatment for many years and was believed to be a uniform area. The result from Sawyers field is shown in Fig 1. Keen & Haines has made a contour map manually based on an average soil mechanical resistance values obtained for each plot. They called these “isodynes” for contour lines of equal force or soil mechanical resistance.

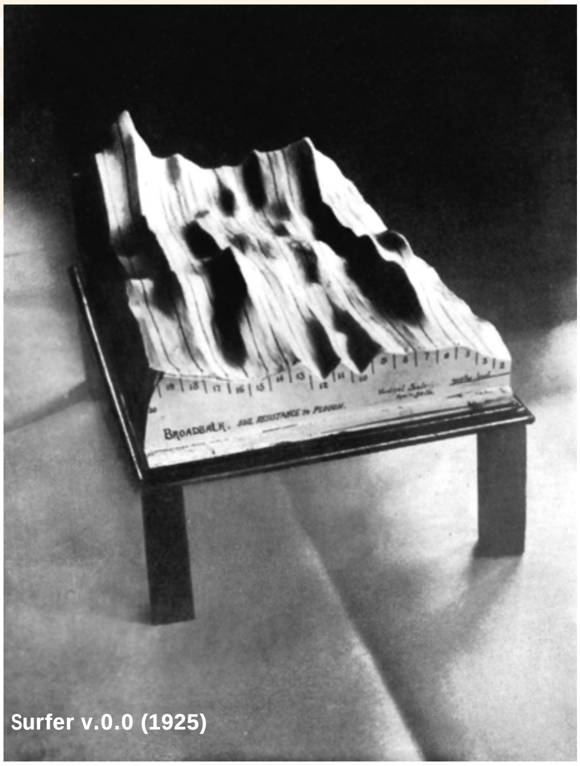
Keen and Haines (part III of the paper) used their dynamometer to map soil mechanical resistance in the classical Rothamsted experiment plots in 1925. They also built a real 3-D model (pre-GIS) for the soil mechanical resistance for one of the fields (the famous-Broadbalk permanent wheat plots), and showed that the predominant factor is natural soil variation and not ‘manuring’, in spite of the long period (since 1843) over which it has been applied. It seems that none of this work was referenced in the scientific literature until the Broadbalk measurements were repeated recently (Watts et al., 2006).

Reference

Keen, B.A., 1931. The Physical Properties of the Soil. Longmans, Green & Co, London.

Pereira, C., 1982. Bernard Augustus Keen. 5 September 1890-5 August 1981. Biographical Memoirs of Fellows of the Royal Society 28, 204-223. http://tinyurl.com/2evp5x

Watts, C.W., L. J. Clark, P. R. Poulton, D. S. Powlson & A. P. Whitmore. 2006. The role of clay, organic carbon and long-term management on mouldboard plough draught measured on the Broadbalk wheat experiment at Rothamsted. Soil Use & Management 22, 334-341.

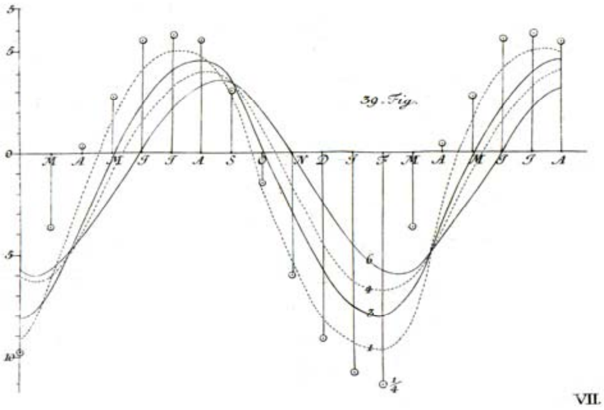


Surfer v.0.0 (1925)

Mehr über Pyrometrie

Uta Stockmann, Budiman & Alex

The graph below appears in the famous book **The Visual Display of Quantitative Information** by Edward Tufte. The graph is from a book by Johann Heinrich Lambert (1779) called Pyrometrie.



Tufte wrote (p. 29): "It was not until the late 1700s that time series charts began to appear in scientific writings. This drawing of Johann Heinrich Lambert, one of a long series, shows the periodic variation in soil temperature in relation to the depth under the surface. The greater the depth, the greater the time-lag in temperature responsiveness. Modern graphic designs showing time-series periodicities differ little from those of Lambert although the data bases are far larger."

This example shows one of the earliest time series in the scientific literature and is actually on soil data. As we wrote in the previous issue of Pedometron the variation of temperature with time and depth has attracted scientists to work on them using physical principles and other modern numerical tools such as Fourier transform and wavelet analysis.

And indeed analysing soil temperature attracted early scientists such as Saussure in 1785, Leslie in 1815 and James Forbes in 1846. Early studies were mainly interested in the theory of the conduction of solar heat to earth. See also the history of the inductive sciences http://tinyurl.com/2kue65

However until recently (e.g. Kuzyakova and Stahr, 2006), pedometricians have been more interested in applying modern mathematical/ statistical methods to the time series.


Johann Heinrich Lamberts

Pyrometrie

oder

vom Maaße des Feuers und der Wärme.

Mit acht Kupfertafeln.

Berlin
bey Haude und Spener.
1779.

Johann Heinrich Lambert (1728 –1777), was a German mathematician, physicist and astronomer. He is well known for his work on light and optics (Beer-Lambert law), was the first to introduce hyperbolic functions into trigonometry, and invented the Lambert cylindrical equal-area projection. Lambert also had a great interest in the theory of heat conduction. Pyrometrie was his last book completed in May 1777. He died on September 25, 1777 in Berlin at the age of only 49.

The Pyrometrie book can now be viewed at: http://num-scd-ulp.u-strasbg.fr:8080/64/

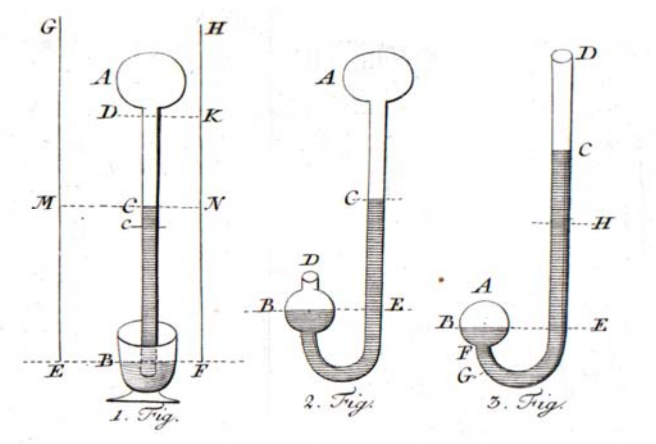
The temperature data of Lambert's famous graph came from measurements of differences in soil temperature with increasing soil depth. It was conducted by Mr. Ott, a salesman of Zurich, Switzerland beginning in 1762 for a period of 4 ½ years. The method is as follows:

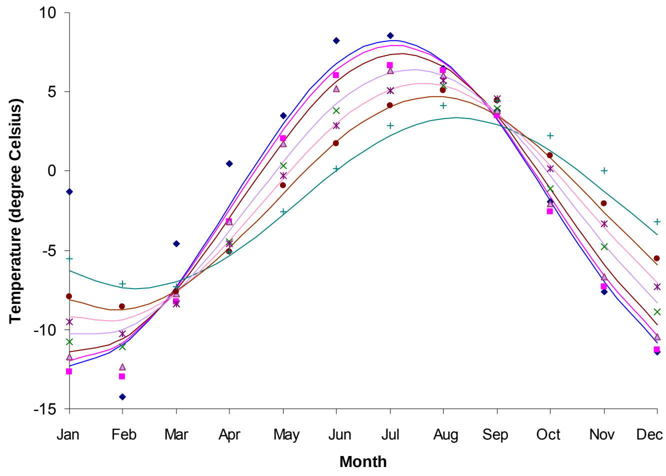
- The thermometers, surrounded by pipe, were inserted into soil at appropriate depths.
- The thermometers were filled with 'Weingeiste' (alcohol) and not mercury, with a scale from the freezing point of water (-10.4 degree) to boiling point of water (Fig. 3)
- Measurements were based on du Crest's temperature scale (after Micheli du Crest)
- The thermometer readings were at ¼, ½, 1, 2, 3, 4, 6 Fuß below soil surface (which is equivalent to the imperial foot, 1 Fuß = 0.31 m)
- The mean temperature for each month was calculated for the period of 4 years

Lambert used the mean temperature values to draw the famous figure 1. The curve is using the following equation:

$$v = a.f\,(b + x) + c.f\,(d + 2x) + \&c.$$

where v = stored heat. (The full equation is much more complicated than this.)





Soil temperature at Zurich at the depth of 0.08, 0.16, 0.31, 0.62, 0.9, 1.25, and 1.88 m below soil surface.

We present the data in Celsius in Table 1. We converted the temperature from the du Crest scale into Celsius by:

Degree C = 1.2 du Crest + 10.4

We also obtained the air temperature at Zurich. And we fitted the comtemporary model (see last issue of Pedometron):

$$T_{z,t} = T_{av} + A\exp\left(-\frac{z}{z_d}\right)\cdot\sin\left(\omega\left(t-t_0\right)-\frac{z}{z_d}\right)$$

The data and fitted curve are presented in the above figure. We obtained the following parameter values: average temperature T_{av} = -2.02 ^{o}C, amplitude A of 10.5^{o}C, phase constant t_0 = 0.33 year, with a damping depth z_d of 2.8 m and a thermal diffusivity of 24.43 m^2 $year^{-1}$. The amplitude is much higher compared with Forbes' data, but the damping depth is similar to the sand in Forbes' study. No information is given on the soil, but we expect that the soil where Mr. Ott conducted his experiment in Zurich was of a sandy nature.

Table 1. Mean soil temperature in degree Celsius (for 4 years)

Month	Depth of thermometer below soil surface (m)							
	0*	0.08	0.16	0.31	0.62	0.93	1.25	1.88
Jan.	-1.3	8.8	-2.1	-1.2	-0.3	0.9	2.4	4.7
Feb.	2.3	-3.6	-2.4	-1.8	-0.6	0.2	1.8	3.2
March	4.3	5.6	2.1	2.6	2.0	2.0	2.7	3.0
April	8.3	10.5	7.0	7.0	5.8	5.6	5.2	5.2
May	12.9	13.3	12.0	11.7	10.3	9.7	9.1	7.6
June	15.9	17.9	15.7	15.0	13.6	12.7	11.7	10.2
July	18.5	18.2	16.4	16.1	14.8	14.8	13.9	12.7
Aug.	17.1	16.2	16.1	15.7	15.1	15.4	14.8	13.9
Sept.	13.0	13.6	13.3	13.6	13.8	14.4	14.2	14.2
Oct.	7.7	8.2	7.6	8.0	8.9	10.2	10.9	12.1
Nov.	4.1	2.7	3.0	3.6	5.5	6.8	8.0	10.0
Dec.	0.3	-0.9	-0.7	0.0	1.5	3.0	4.7	7.0

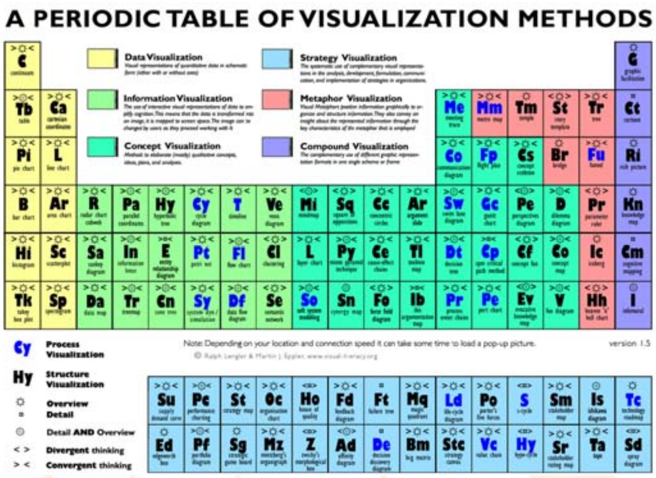
* mean air temperature at Zurich from 1881 to 1885 obtained from http://data.giss.nasa.gov/gistemp/

Worth a look on the web

There are interesting on the web related (or unrelated) to soil and pedometrics. Here's my pick, if you have others you'd like to share around send an email to vchair@pedometrics.org.

(1) Periodic Table of Visualization Methods

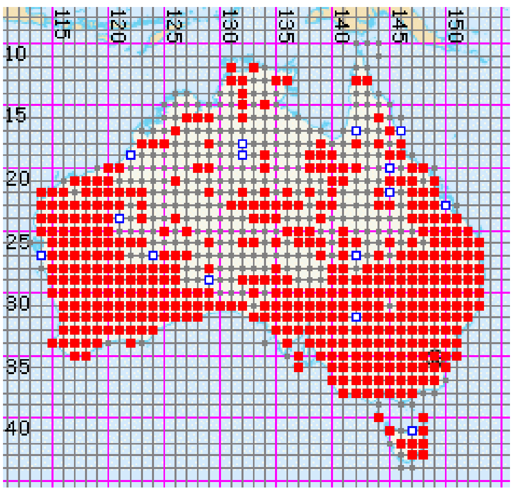
http://www.visual-literacy.org/ compiled various methods of visualization in a periodic table: histogram, data map, mindmap, Ishikawa diagram etc.



(2) Confluence.org

http://www.confluence.org

The project attempts to visit each of the latitude and longitude integer degree intersections in the world, and to take pictures at each location. Would be great if they dig a soil profile and collect the samples on each site as well. Excellent addition to the ISRIC world soil profile database. http://eusoils.jrc.ec.europa.eu/projects/soilimages/



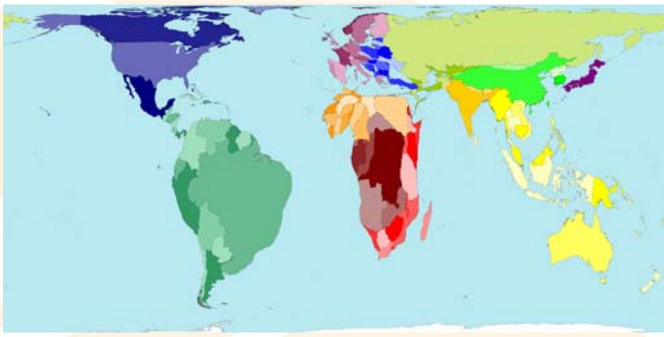
(3) Worldmapper

http://www.sasi.group.shef.ac.uk/worldmapper/

Worldmapper projects areas on the map according to the subject of interest known as equal area cartograms, or density-equalising maps. The cartogram resizes each territory according to the variable being mapped. The method is described in:

Michael T. Gastner and M. E. J. Newman (2004) Diffusion-based method for producing density equalizing maps *Proc. Natl. Acad. Sci.* USA 101, 7499-7504.

The site has 366 maps on various categories, but no soil map yet. The following example is the projection of forest area in the year 2000. Will be great for the globalsoilmap.net



(4) ElfYourself

http://www.elfyourself.com

In this festive holiday season, you can make an elf of yourself and send it around. It's fun.

Pedometrician profile

Inakwu Odeh

The University of Sydney, Australia

How did you first become interested in soil science?

The genesis of my interest in soil science started way back in Nigeria during my Higher School Certificate days, in the mid-1970s, when we undertook a number of interesting geography excursions led by our then astute Geography Teacher– a Briton– Mr David Jones. He was very passionate about nature and how the landscape is being shaped by pedo-geomorphological processes. I had even learnt about the concept of "catena" then. We also studied what are generally termed in geography as *Cuestas*, defined as "ridges formed by erosion-resistant "gently tilted hard rock layers". Since then I have been hooked to soil science.

How were you introduced to pedometrics?

I am one of the fortunate participants at and witness the birth of pedometrics in Australia in the mid-1980s when Alex McBratney first delivered a series of his seminal lectures which he then termed as "Introduction to Pedometrics". By then I was already interested in the application of geostatistics to the study of soil and its variation. I saw Alex then as an erudite, young and enthusiastic soil scientist equipped with many scientific metaphors as a means to possibilities and fresh understanding of the soil system. Evidently, Alex has not changed very much except that he is now older and much wiser. I have gained tremendously from his mentorship and scholarship over the years since I started working with him.

What recent paper in pedometrics has caught your attention and why?

I am interested in articles that further broaden our understanding of the fourth dimension- time, in the scheme of soilscape processes. There are a number of articles dealing with generic methods, namely space-time Kalman filter, wavelet transforms and time-series analysis. But what drew my attention is a recent paper by Kuzyakova et al. (2006) entitled "Time series analysis and mixed models for studying the dynamics of net N mineralization in a soil catena at Gondelsheim" Geoderma 136, 803–818. I believe their approach addresses a number of challenges in modeling soil dynamics in relation to N-cycle, which can be extended to others such as C-cycle, etc.

What problem in pedometrics are you thinking about at the moment?

As mentioned above, the immutable fourth dimension of time in the realm of soil processes is very important. As such my current thinking is to apply pedometric techniques to look into the past and future simultaneously. My focus is the study of land use and land use change in relation to soil dynamics. We should escape from the cocoon of present-day pedometricians and look at time-dependent soil processes as influenced by the anthropological factor, and then extending this to the understanding of global changes.

What big problem would you like pedometricians to tackle over the next 10 years?

The next 10 years will witness a lot of activities to tackle global climate change. I therefore foresee the pertinent role of pedometricians in monitoring the GHG emission reduction targets for countries and also provide knowledge for optimal agricultural adaptation. In the former case, and borrowing the opinion of Professor Don Nielsen in his contribution to Pedometron No. 4 & 5 in 1997, pedometricians should be bold enough to apply existing knowledge or develop new or enhanced pedometric techniques to "… link land-atmosphere energy and material fluxes to better understand the feedback loops between the landscape and atmosphere".

Non-Pedometrician profile

Daniel Hillel

Columbia University, USA

How did you first become interested in soil science?

I spent my early childhood in a pioneering Israeli settlement, on the frontier between the wilderness and the sown land, and was indelibly impressed by the physical, biological, and historical contrasts of the labile environment.

What are the most pressing questions at the moment in your area of soil science?

How to bridge over various spatial and temporal scales, heretofore treated in a disjointed fashion.

What statistical and mathematical methods are used in your area of soil science?

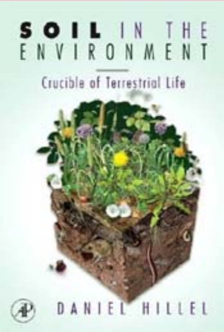
The old methods are no longer appropriate to the encompassing problems of an interconnected and interdependent world.

Are you aware of any work by pedometricians that might be relevant to your science?

I am not sufficiently aware, but am determined to rectify this lacuna.

What big problem would you like pedometricians to tackle over the next 10 years?

The issue raised above, as well as the integration of land processes with atmospheric and oceanic processes, leading to a more holistic understanding of global processes.



Prof. Hillel is a world-renowned soil scientist and hydrologist. He has authored twenty-plus books include definitive works on arid-zone ecology, low-volume irrigation, and soil and water physics which are widely adopted as standard texts in universities. He is also the Editor-in-Chief of the Encyclopedia of Soils in the Environment. His new book ***Soil in the Environment*** *attempts to unite soil science and the environment beyond what is traditionally taught .*

Pedomathemagica

With Gerard Heuvelink

Problem 1 (EASY)

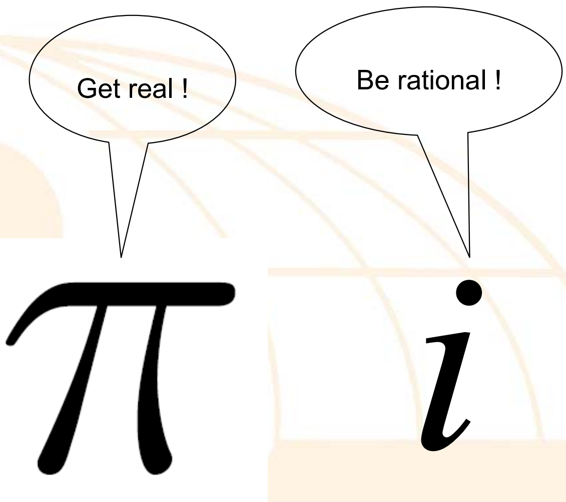
Two soil scientists Tom and David return from fieldwork. They are carrying several heavy bags containing soil samples. Tom puffs and moans, upon which David tells him: “What are you complaining about, if you would give me one of your bags then I would be carrying twice as many bags as you, while if I would give you one of mine each of us would be carrying the same number of bags”. How many bags is David carrying?

Problem 2 (HARD)

In order to carry out her fieldwork Maggie must cross an 800 km wide desert with her jeep. At the start there is a petrol station with an unlimited amount of petrol. However, the problem is that the jeep cannot take more petrol (either in the tank or in jerry cans or as a combination of the two) than takes the jeep as far as 500 km. Maggie therefore cannot cross the desert in one go but must create petrol depots in the desert, return to base and get a refill, again store part of it in the desert, and so on until she has stored enough to be able to reach the other side. Maggie is allowed to create as many depots as she likes and place them wherever she likes. What is the minimum number of kilometers that she must drive to cross the desert? What is the maximum width of a desert that she can cross in this way (if there is any)?

The Answers??

Next issue of Pedometron



☆*Pedometrics 09*☆

Will be held in Beijing, October 2009, hosted by China Agricultural University. Mark your diary and watch out for the date!

Pedomathemagica

Answers to last edition's Quiz

Congratulations to Gerard Heuvelink who won the quiz convincingly. Here are the solutions.

1. The answer is ⅔. The best starting point with such questions is to consider the sample space, the set of possible outcomes. Before we received any information from the pedologist the sample space was as follows, there are four possible outcomes.

First Pit	**Second Pit**
R	R
R	CBE
CBE	R
CBE	CBE

Where R denotes a Rendzina and CBE a Calcareous Brown Earth. From the question we know that the four outcomes have equal probablity.

When we are told "At least one of the profiles was a Rendzina" then our sample space is a subset of the initial space. Only the first three lines in the table above constitute our sample space when this piece of information has been received. Of these three equiprobable outcomes two are cases where a Calcareous Brown Earth was seen, so the answer to the question is ⅔.

If the pedologist had said, "The first profile that I examined was a Rendzina" then our sample space on receiving this information is given by the first two lines in the table, in one of them she also saw a Calcareous Brown Earth, so the probability that she saw a CBE is ½.

This problem illustrates the importance of thinking clearly about the sample space. To me it also emphasises that probabilities are statements about our rational belief, given evidence (*pace* Karl Popper), and that subtle differences in the information that we receive can have large implications for probability.

2. The abscissa (x-axis) is the size of a randomly assembled group of people. The ordinate is the probability that at least two of them share a birthday (day of the year) on the assumption that the year has 365 days (i.e. ignoring leap years), and that the probability that a randomly selected person is born on a specified day is 1/365 for all days.

It is then easy to see why the function reaches exactly 1 at 366. If we had 365 people is it possible, though unlikely, that each of them has a birthday that is unique in the group, so the probability must be less than 1. However, in a group of 366 people at least two must share a birthday as there are only 365 to go around, so the probability becomes exactly 1.

To get the function, consider first how to compute p_0, the probability that no two people in a group of m share a birthday. To visualize it, imagine that they enter the room in succession. The probability that the second person has a different birthday from the first is $1.(1 - {}^1/_{365})$, the probability that the third person has a different birthday from the first two is $1.(1 - {}^1/_{365}).(1 - {}^2/_{365})$, for the full group of m people the probability is

$$p_0 = 1.(1 - {}^1/_{365}).(1 - {}^2/_{365}) \ldots (1 - {}^{(m-1)}/_{365}).$$

The probability plotted in the quiz is then $1 - p_0$. It is clear that if $m \geq 366$ then one of the terms in the product above is 0, so $p_0 = 0$ and our probability becomes 1.

If you have time on your hands, and want to find out more (including the Taylor Series approximation) look up "Birthday Paradox" on Wikipedia. Of course it is not strictly a paradox, but for many it is counterintuitive that m only has to be 23 or more for the probability of at least 2 people sharing a birthday to exceed 0.5.

3. The first expression in the chain is Euler's relation which links the basic numbers of mathematics — e, i and π. Picture (d) is of Euler, an 18th Century Swiss mathematician. The second expression is the expectation of the Fisher score (the partial derivative of the log-likelihood with respect to a parameter). The variance of this is the Fisher information, used to obtain covariance matrices for parameter estimates. R.A. Fisher, a pioneer of applied statistics is pictured at (a). The third expression is due to Jenny (picture c) which expresses the interrelation of soil and environmental factors in the soil system. It is less familiar than the usual factors of soil formation equation, but appears earlier in the discussion in his book. The fourth expression is one of the two components of the intrinsic hypothesis of stationarity due to Georges Matheron (picture b). Matheron gave a theoretical framework to the proposed spatial estimator developed for mining exploration in South Africa by Danie Krige (picture e). The last expression is Krige's relation (slightly rearranged from its usual presentation to ensure equality with all the other expressions). And all the expressions are all equal to zero.

Murray