ΠΕΔΟΜΕΤRON

The Newsletter of the Pedometrics Commission of the IUSS

Issue 28, December 2009

Chair: Murray Lark *Vice Chair & Editor: Budiman Minasny*

From the Chair

Dear Colleagues

Welcome to the final Pedometron of 2009. This has been another productive year in pedometrics, most notably because of the Pedometrics 2009 conference so ably managed by the China Agriculture University in Beijing. The first Pedometrics meeting in Asia is an important milestone, and I hope that it will prove in due course to have provoked an expansion of pedometrical activity in China and beyond. Thanks again to Yuanfang Huang and Baoguo Li, and their colleagues for all the hard work which lay behind this event.

There has been a recent controversy centred at the University of East Anglia, not far from where I write, concerning hacked emails which appear to show less-than-ideal behaviour by climate scientists when dealing with dissenting views on climate change. I still maintain (as in an earlier Pedometron) that honest debate is essential for science, but it doesn't take too much imagination to understand why someone might try to suppress it on this issue. If, like most environmental scientists, you are convinced by the case that there is current anthropogenic climate change that threatens to undermine the long-term sustainability of many of the planet's key life-support systems then it follows that you want to see appropriate global-scale commitment by government to do everything possible to reduce climate disruption. As a scientist you know that pretty much all scientific knowledge has an attendant uncertainty, but your scientific formation means that you are habituated to this, and are capable of making rational choices and assessments given appropriate information about uncertainty (be it statistical or conceptual). The great problem comes in communicating uncertain scientific knowledge to the decision makers in democratic states so that the uncertainty is dealt with rationally, and not allowed to paralyse the decision-making process.

In Shakespeare's *Othello* the Duke of Venice is presented with conflicting intelligence about the actions of the Turkish navy and their intentions towards Cyprus. He responds:

> *Nay, it is possible enough to judgment:*
> *I do not so secure me in the error,*
> *But the main article I do approve*
> *In fearful sense.* (Act 1 Scene 3).

In short, he admits the uncertainty, but does not allow it to distract him from the serious implications of the information he has received. The Duke could make such a judgement, and expect those who matter to follow him. Modern governors face a tougher challenge, particularly when there are powerful vested interests, with substantial financial backing, which would like nothing better than to see governments 'so secure[d].. in the error' of climate models that the 'main article' is lost, and we have business as usual.

But the solution is not to suppress uncertainty, rather we need a kind of socio-psychological ergonomics of

Inside This Issue

uncertainty. By this I mean a way to present the uncertainty in scientific information which is honest, but which also helps the general public and their representatives to set it properly in context. As pedometricians uncertainty is our stock in trade. Any prediction or estimate should have some kind of uncertainty assessment attached to it. But perhaps we forget how difficult these assessments may be for those coming to them cold. Consider a power analysis, for example. If we tell a customer (maybe a government official) that our proposed sampling scheme will detect a 10-tonne per hectare change in soil carbon at 95% significance with 80% power we are asking them to deal with a probability assessment (the power) that a probabilistic event (our estimate of change falls within a given interval) conditional on the true value of that even being of a certain size. This probability of a conditional probability requires careful handling by statisticians, still less those who have to make decisions based on them. I recently attended a meeting where a senior UK civil servant with responsibilities related to climate change kicked off proceedings by telling us that his degree was in medieval literature. Some of my colleagues groaned, but I am all in favour of generalists taking on such roles. It is probably better if scientists are compelled to find appropriate ways to express uncertainty to non-specialist officials so that these are passed on in comprehensible and relatively un-mangled form to their final consumer — voters under democratic government.

I offer no solutions, but a statement of a problem which has to be tackled if pedometricians are to be effective citizens in the wider, and rather troubling world in which we all live.

May I draw your attention to the call for nominations for the Richard Webster medal (Page 3). I hope to see many of you at the Pedometrical symposia at World Congress of Soil Science next August. I hope to see even more at Pedometrics 2011 which, as was announced in Beijing, is to be organized by Luboš Borůvka and held at Třešt′ in the Czech Republic.

Best Paper 2008

We have 40 votes cast at the end of July 2009, and the winning paper was:

Grinand, C., Arrouays, D., Laroche, B. and Martin, M.P., 2008. Extrapolating regional soil landscapes from an existing soil map: Sampling intensity, validation procedures, and integration of spatial context. Geoderma, 143(1-2): 180-190.

Congratulations to Clovis and colleagues for an excellent paper. The award was presented to Dominique Arrouays at Pedometrics 2009 in Beijing.

Call for Nominations for Best Paper 2009

We 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 2009 (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 2009 should be sent to Murray Lark (murray.lark@bbsrc.ac.uk) by end of March 2010.

THE RICHARD WEBSTER MEDAL:

AN AWARD BY THE PEDOMETRICS COMMISSION OF THE INTERNATIONAL UNION OF SOIL SCIENCES

Photo by Mark Mallott

The Richard Webster Medal: an award by the Pedometrics Commission of the International Union of Soil Sciences

The Richard Webster medal was established before the last World Congress of the International Union of Soil Sciences (IUSS). The award is for the best body of work that has advanced pedometrics (the subject) in the period between the IUSS World Congress of 2006 and the next one in 2010. However, achievements before that period will also form part of the evaluation (see more detail below). The award will be made at the next meeting of the IUSS World Congress. The first award was made to Professor Alex McBratney (University of Sydney) at the World Congress in Philadelphia (USA).

Guidelines for the award of the Richard Webster Medal

The official rules are also at http://www.iuss.org/popup/Webster_medal.htm

Requirements and eligibility for the award of the Richard Webster Medal

1. Soil scientists eligible for the award will have shown:

   a) a distinction in the application of mathematics or statistics in soil science through their published works,

   b) innovative research in the field of pedometrics,

   c) leadership qualities in pedometrics research, for example, by leading a strong research team,

   d) contributions to various aspects of education in pedometrics (e.g. supervision of doctoral students, teaching of pedometrics courses in higher education, the development of courses for broader professional needs),

   e) and service to pedometrics (e.g. by serving on a committee of the Pedometrics Commission or promoting pedometrics to the IUSS).

2) A nominee should be a member of the IUSS at the time of the nomination and have been involved in activities associated with pedometrics, in particular.

3) The nominee must be living at the time of the selection; retired pedometricians still active in pedometrics research will be eligible for the award. The nominee should be willing to receive the medal at the time and place designated by the IUSS World Congress, and be a keynote speaker at the next conference of the Pedometrics Commission (held biannually) following the presentation of the medal.

4) The Pedometrics Commission will pay for the recipient's travel expenses to attend the Pedometrics meeting where the keynote address will be given.

5) Members of the Awards and Prizes Committee shall be ineligible to receive the medal while serving on the Committee.

6) The award of the Richard Webster Medal shall not be presented to any one individual more than once.

Nominations procedure

1) Nominations for the Richard Webster Medal should be made by a colleague or colleagues who know the person's work well. The nomination should include a résumé and a short statement (a maximum of 750 words) summarizing the relevant qualifications of the nominee with respect to the conditions outlined in the section, requirements and eligibility, above.

2) The proposer(s) should submit the following on behalf of their nominee two months before the next IUSS conference (August 2010), i.e. before the 1st of June 2010:

   a) their published work for the four-year period between consecutive IUSS meetings,

   b) a suitable curriculum vitae that gives:

      - all previous publications,
      - positions held,
      - research undertaken,
      - education of others,
      - teaching courses developed,
      - and leadership and management of research projects.

This material should be sent to the Pedometrics Awards Committee chair, Professor Margaret Oliver at m.a.oliver@reading.ac.uk.

Inclusion of any of the above must show clear relevance to pedometrics.

Pedometrics 2009 Report

A-Xing Zhu

The biennial Pedometrics conference was successfully held in Beijing at the JinMa Hotel in China Agricultural University, August 26-28, 2009. The conference was hosted by the Soil and Water Sciences Department, College of Resources and Environmental Sciences, China Agricultural University, sponsored by China Agriculture University, the Pedometrics Commission of the IUSS, National Natural Science Foundation of China, and Beijing soil fertilizers work station.

Many distinguished guests attended the opening ceremony, including Prof. Qixin Sun, the Vice-President of China Agricultural University, Dr. Murray Lark, Chair of the Pedometrics Commission of the International Union of Soil Science (IUSS) from Rothamsted Research, UK, Prof. Yuanshi Gong, the Administrative Vice-President of the Scientific Research Institute of China Agricultural University, Prof. Alex McBratney from University of Sydney, Prof. Baoguo Li, the Vice-Dean of College of Resources and Environmental Sciences, China Agricultural University, Dr. Budiman Minasny, the Vice-Chair of Pedometrics Commission of the International Union of Soil Science (IUSS) from University of Sydney. Prof. Yuanfang Huang from China Agricultural University chaired at the opening ceremony. Prof. Qixin Sun on behalf of China Agriculture University gave a welcome speech at the beginning, and briefly introduced soil science as a traditional subject of China Agriculture University, emphasized the significance of quantitative research on soil science, and wished the conference a complete success. On behalf of the IUSS, Dr. Budiman Minasny expressed gratitude to China Agriculture University for its elaborate organization of the conference. Prof. Baoguo Li introduced the historical development and present situation of soil science in China Agriculture University at length and briefly presented the development of Pedometrics in China Agriculture University in terms of integration research of soil science and information technology.

Tom Hoogland on his award winning presentation.

Nearly 50 foreign scholars from 18 countries including Australia, United Kingdom, the United States of America, France, Netherlands, Germany, Italy, Canada and about 60 domestic researchers participated in the conference. The conferences covered a wide array of topics in pedometrics including soil proximal remote sensing, interfacing Geostatistics and GIS, soil sampling and monitoring, spatial statistics, soil-landscape reconstruction, space-time modeling, from point to globe scale issues, and application of pedometrics in agriculture and environmental sciences. Prof. Zhou Shi of Zhejiang University (China), Dr. Christian Walter of INRA-Agrocampus Ouest (France), Prof. Peter Finke of Ghent University (Belgium) and Prof. Alex McBratney of The University of Sydney (Australia) presented key papers, respectively. The conference hosted 40 oral papers and 26 posters.

During the conference, Dr. Murray Lark, Chair of the Pedometrics commission, presented Prof. Alex McBratney the research award of Pedometerics, the Webster Medal, and the award of Best Paper in pe-

Tom Hoogland, Christian Walter receiving best oral presentation & poster award from Li BaoGuo.

Brian Murphy & Christian Walter.

dometrics 2008. At the closing ceremony presided by Prof. Tusheng Ren, Prof. Baoguo Li presented Dr. Tom. Hoogland, the author of the best paper “Spatio-temporal modeling of the lowering of peat soils” and Prof. Christian Walter on behalf of Dr. Blandine Lemercier, the author of the best poster “Extrapolation at sub-regional scale of local knowledge embedded in detailed soil maps using ancillary data” the prizes: the Chinese yoyo.

Prior to the conference a two day Pre-conference Workshop on Sampling for Survey and Monitoring of Natural Resources by Dick Brus (Wageningen University) and Martin Knotters (Alterra-Wageningen University) was held. About 40 scholars from 10 countries attended the workshop.

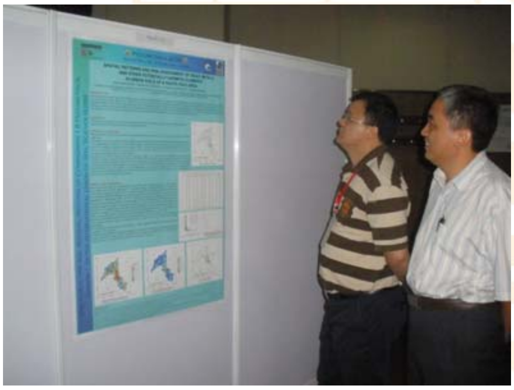
After the academic conferences, many of the participants took the conference field trip examining the typical soil pedons around Beijing, and visited the demonstration farm of National Precision Agriculture of China.

In addition to the academic program, pedometricians also visited the world renounced cultured sites in Beijing such as the Great Wall, the Forbidden City and as well as the lately booming street bars in the Bei Hai area. The pictures below are examples of these happy times. Homework for all of your pedometricians: Try to figure out what they were doing in these pictures.

On pedometrics agenda with Peter Finke.

PM 2009

Raphael & Alex receiving Best Paper 2007 award.



The Soil Texture Wizard

R functions for plotting, classifying and transforming soil texture data

Julien Moeys

Department of Soil and Environment, Swedish University of Agricultural Sciences, Uppsala, Sweden

"Soil texture", "clay, silt and sand", and "soil texture classes" are concepts familiar to every soil scientist. But do we really know them? When a French, an Australian and a Swedish soil scientists are discussing about the "sand fraction" (or "sand content") of a soil, are they talking about the same thing? Maybe not. The French soil scientist may be thinking about the 50-2000µm fraction, the Australian about the 20-2000µm fraction and the Swedish about the 200-2000µm fraction.

When the same soil scientists are discussing about a texture class named "clay" for a given soil horizon, are they talking about the same thing? Maybe not! Maybe the French soil scientist is thinking about the "Argile" texture class of the so called French "Aisne" texture triangle, but maybe he / she is thinking about the "Argile" class of the French GEPPA texture triangle (which are not identical). The Australian is thinking about the "Clay" texture class of the Australian soil texture triangle, which more or less the "Argile lourde" (heavy clay) of the French Aisne and GEPPA triangles. And the Swedish? The Swedish is used to work with 4 particle size classes, so instead of clay, silt and sand, he / she often uses 'ler', 'mjäla', 'mo' and 'sand'. But when needed he / she may uses the FAO texture classification system or the USDA one.

So unfortunately for soil scientists, some of the most common concepts we use don't have a universal definition, and may vary between different countries:

[1] there are different 'norms' for the silt and sand classes particle size boundaries;

[2] there are different soil texture classification systems, that have been build using different 'norms' of particle size classes;

[3] The classes defined by different soil classification systems are not directly comparable; and

[4] the usual representation of soil texture classification as a 'texture triangle' graph is based on different triangle geometries, but hopefully that aspect only affect the representation, and not the definition of texture classes!

There is no international consensus to normalize soil particle size classes and soil texture classification systems, so soil scientists using heterogeneous soil texture datasets have to convert data from one system to another.

The Soil Texture Wizard is a set of functions for R "language and environment for statistical computing and graphics", and a future R 'package'. This toolbox provides different functions for

[a] plotting soil texture data,

[b] allocating the soil texture data to a texture class,

[c] transforming soil texture data (that is to convert particle size fraction between different systems of particle size class limits; such transformation are never exact) and

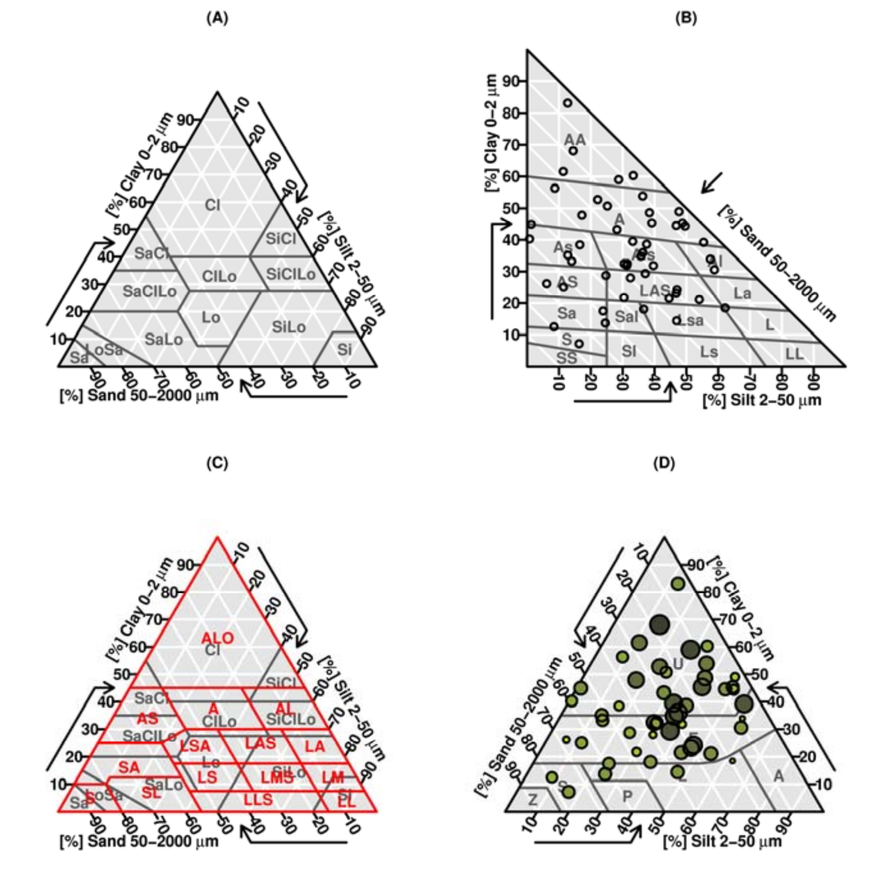


Figure 1. Four examples of texture triangle plots that can be produced with The Soil Texture Wizard. The code for these examples is given in the article.

[d] perform a few other useful operation such as the normalisation of the sum of the texture fractions or plotting different texture classification systems on top of each other.

Most of these features already exists in different softwares, but to the author knowledge not as an integrated 'bundle'. To cite a few 'competitors', R has a package called 'plotrix', by Jim Lemon *et al.*, for plotting USDA and UK soil texture triangles. Standalone programs such as 'TRIANGLE' by Aris Gerakis and Brian Baer, or 'Texture Auto-Lookup (TAL)' by Christopher Teh Boon Sung, can classify soil texture data according to one or several texture classification system(s).

The specificity of 'The Soil Texture Wizard' is that the functions are **integrated, multi-texture triangle, multi-particle size classes systems** (for clay, silt and sand) **and multi-triangle geometry**. The functions are provided with **8 pre-defined soil texture triangles** (USDA, FAO, French Aisne, French GEPPA, German, Soil Survey of England and Wales, Australian and Belgian), and with one pre-defined particle size classes transformation function (log-linear). It is possible for the user to define its own texture classification system(s), and its own particle size classes transformation function(s), and to use it as any of the pre-defined settings. It means that it is (virtually) possible to plot or classify any soil texture data in any soil texture classification system and using any triangle geometry. Icing on the cake, it is possible to chose between 7 languages for the default triangle plot axis labels and title.

Short example of R code using The Soil Texture Wizard:

Preparing the work:

Before using The Soil Texture Wizard, it is necessary to install R on your computer <www.R-project.org>. You need to download The Soil Texture Wizard source code <http://julienmoeys.free.fr/?Soil-textures-triangle> (Functions source code, .R), and save it on your computer. Finally, you need to open R (or start an R cession from the command line), set the working directory to the folder where you saved the source code (in our example C:\PEDOMETRON\R) and load The Soil Texture Wizard. Here are the first lines of code:

```
setwd( "C:/PEDOMETRON/R" )

source( "FUNCTION_TEXTURE_WIZARD.R" )
```

If you want to use your own soil texture dataset, and have it as a CSV file with CLAY SILT and SAND as headers, you can load it using the code:

```
soils <- read.csv( "filename.csv" )
```

(replace `filename.csv` by your file name)

But if you have no dataset to test, you can create a random dummy soil texture dataset with the code:

```
soils <- TT.dataset( n = 50 )
```

This generates a dummy dataset of 50 clay silt and sand values, with an additional 'Z' dummy variable.

Working with The Soil Texture Wizard:

A few 'typical use' examples are presented here. First we plot a simple USDA texture triangle plot, without data (fig. A):

```
TT.plot( class.sys = "USDA.TT" )
```

Then we plot the French GEPPA triangle, with the 'soils' dummy dataset plotted on top of it (fig. B):

```
TT.plot( class.sys = "FR.GEPPA.TT",
         tri.data = soils )
```

We can compare the USDA and the French 'Aisne' soil texture triangles by over-plotting them (fig. C):

```
geo <- TT.plot( class.sys = "USDA.TT" )

TT.classes( geo = geo,
            class.sys = "FR.AISNE.TT",
            class.line.col = "red",
            class.lab.col = "red" )
```

And finally we can visualise the value of a 4th variable (here called 'Z') with a bubble plot on top of the Belgian texture triangle (fig. D):

```
TT.plot( class.sys = "BE.TT",
         tri.data = soils,
         z.name = "Z" )
```

If you want to allocate your texture data according to the FAO texture classification, type the following code:

```
TT.points.in.classes(
    tri.data = soils,
    class.sys = "FAO50.TT" )
```

and R outputs a table with as many columns as texture classes in the FAO triangle, each row corresponding to a soil texture point. The value is 0 when the point is out of the class, 1 when it is inside, and 2 or 3 when it is on an edge or a vertex of one or several class(es), respectively. Here is an example of output (only the 5 first points). As we have used a random dummy dataset, the result varies every time.

```
      VF F M MF C
[1,]   0 1 0  0 0
[2,]   0 1 0  0 0
[3,]   0 0 1  0 0
[4,]   0 1 0  0 0
[5,]   0 0 0  1 0
```

With the same function, it is also possible to output logical values, or a vector of text strings containing classes abbreviations (in the latter case, if the data lies between several texture classes, the text item will contain 2 or more concatenated classes abbreviation, instead of one). The classification algorithm underlying this function is in fact the function 'point.in.polygon()' of R 'sp' package.

The Soil Texture Wizard sets of functions don't yet have an 'integrated' help (reason why the functions are not yet provided as a package), but the toolbox homepage <http://julienmoeys.free.fr/?Soil-textures-triangle> provides **an extensive tutorial** providing code examples and explanations on most of the functions.

How does The Soil Texture Wizard toolbox works?

Figure 2 presents an overview of the way The Soil Texture Wizard handle texture data and triangles. The toolbox distinguish the "soil texture triangle characteristics", the "soil texture data characteristics" and the "plotted triangle geometrical characteristics".

Both the characteristics of the "soil texture classification systems" and those of the "soil texture data" are defined in a (3D) clay, silt and coordinate system, as well as a particle size class system (limits). Because texture data are 'fractions' that sums to 100%, there 3D coordinates lies on a plane, and it is possible to simplify the system into a 2D coordinate system and only 2 of the 3 texture classes are needed to plot soil texture data. "soil texture classification systems" are defined in 2 parts: [i] a list of all the vertices of the classification system, with there clay, silt and sand coordinates, and [ii] a list of texture classes, with the vertices that delimits them. So texture classes are in fact defined and represented as 'polygons'.

The 'geometrical characteristics' of the triangle plot then determines which trigonometric rules can be used to convert the clay – silt – sand coordinates into x – y coordinates. These characteristics are: the position of clay silt and sand on the bottom, left and right axis; the top, left and right angles of the triangle; and the 'direction' of the 3 axis (clockwise, counter-clockwise or centripetal). By default, the 'geometrical characteristics' of the triangle plot are the 'usual' geometry of the first texture classification system plotted on the graph, but the user can customise any of the above mentioned parameters if he / she wants.

All the base elements that constitute the triangle plot – axes, axes ticks, axes labels, axes arrows, and grid lines are also defined in a clay – silt – sand coordinate system that is transformed into x – y coordinates when the triangle is plotted. The plot 'scene' thus inherently supports all the possible triangle geometry.

Additionally, The Soil Texture Wizard distinguishes the particle size classes systems of the classification system, of the texture data and of the texture plot. By default, The Soil Texture Wizard assumes that the particle size classes systems of the "plotted triangle" and of the "soil texture data" is identical to the particle size classes systems of the texture classification. But the user can 'force' any of the default values, and ask the functions to transform 'on the fly' the soil texture data or the texture classification system when they are plotted into the texture triangle plot.

Credit, license and (no) guarantees:

Although the functions and the triangle definitions have been built with great care, guaranteeing the conformity of these triangles to the "official" triangle definition is out of the scope of this (personal) project. So these functions are provided without any guarantees from the author or his employer. The toolbox comes with functions that helps the users to check the conformity of the triangle to their own 'references' (see the tutorial).

Although now almost entirely different, The Soil Texture Wizard is initially derived from the soil.texture() function of Jim Lemon's plotrix R package.

The Soil Texture Wizard is free and open source. It is licensed under a GNU-GPL version 3. Anyone is allowed to check the source code, reuse it or modify it.

References and useful readings:

Gerakis A. and Baer B., 1999. A computer program for soil textural classication. *Soil Science Society of America Journal*, 63:807-808.

Minasny B. and McBratney A.B., 2001. The australian soil texture boomerang: a comparison of the australian and usda/fao soil particle-size classication systems. *Australian Journal of Soil Research*, 39:1443-1451.

Nemes A., Wösten J.H.M., Lilly A., and Oude Voshaar J.H., 1999. Evaluation of different procedures to interpolate particle-size distributions to achieve compatibility within soil databases. *Geoderma*, 90:187-202.

Richer de Forges A., Feller C., Jamagne M., and Arrouays D., 2008. Lost in the triangular diagrams of soil texture. Poster.

Teh C.B.S. and Rashid M.A., 2003. Object-oriented code to lookup soil texture classes for any soil classication scheme. Communications in Soil Science and Plant Analysis, 34(1):1-11.

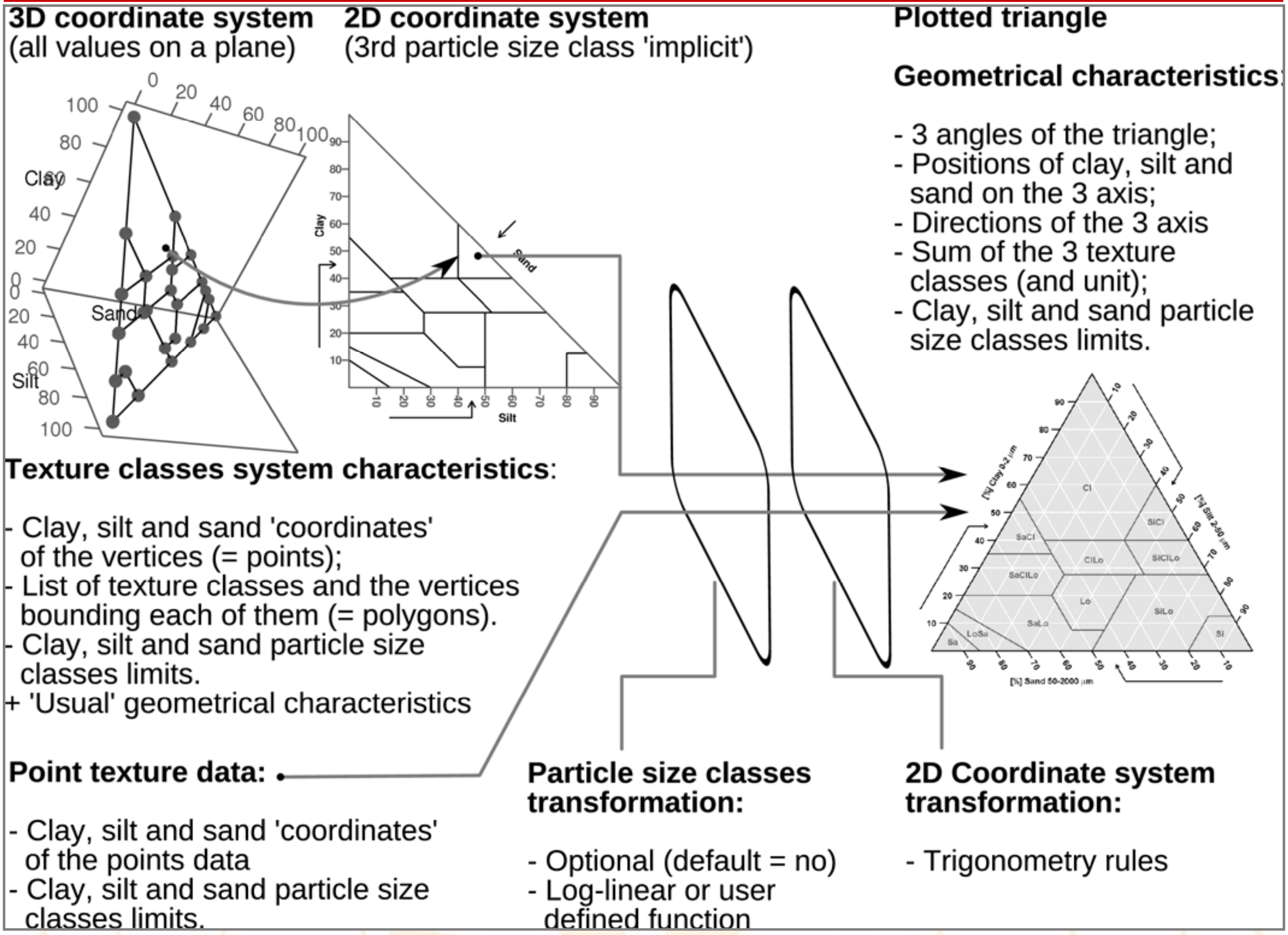


Figure 2. Scheme of the different components (soil texture classification system, soil texture data and texture triangle plot), components' characteristics and operations realised by The Soil Texture Wizard functions to produce a texture triangle plot.

Appendix: algorithm for transforming clay – silt – sand coordinates to x – y plot coordinates.

F_B, F_L, F_R the particle size fraction displayed on the bottom, left and right axis, respectively.

A_T, A_L, A_R the top, bottom left and bottom right angles of the plotted triangle, respectively, expressed / converted in Radian.

S_F the sum of the 3 particle size fraction.

D_B, D_L, D_R the direction of the bottom, left and right axis, respectively. Possible values are 'clockwise', 'counter-clockwise' and 'centripetal'. Only 4 combinations are possible for D_B, D_L *and* D_R (respectively): full 'clockwise' triangle (like the USDA triangle); full 'counter-clockwise' triangle (like the Belgian triangle); 'counter-clockwise' – 'clockwise' – 'centripetal' (like the French GEPPA triangle); and 'clockwise' – 'centripetal' – 'counter-clockwise'.

$$if\ (D_B = 'clockwise')\ then\ \{F'_B = S_F - F_B\}$$
$$else\ \{F'_B = F_B\}$$

$$if\ (D_L = 'counterclockwise')\ then\ \{F'_L = S_F - F_L\}$$
$$else\ \{F'_L = F_L\}$$

$$if\ (D_R = 'clockwise')\ then\ \{F'_R = S_F - F_R\}$$
$$else\ \{F'_R = F_R\}$$

$$if\ (D_L = 'clockwise')\ then\ \{y = F'_L . \sin(A_L)\}$$
$$else\ \{y = F'_R . \sin(A_R)\}$$

$$if\ (D_B = 'clockwise')\ then\ \{x = F'_B - y / \tan(A_R)\}$$
$$else\ \{x = F'_B + y / \tan(A_L)\}$$

Pedometrics 2011

Pedometrics 2011 conference will be held in the Castle Hotel Třešť situated in a nice countryside of the Czech-Moravian Highlands in the Czech Republic in summer 2011. In addition to the conference programme, field trip will include visit to historical sites like UNESCO World Cultural Heritage sites of Telč and Zelená Hora.

Exact dates of the conference will be specified in January 2010. Please check the conference website http://2011.pedometrics.org or

contact Luboš Borůvka (boruvka@af.czu.cz).

Incorporating soil aging in digital mapping of soils

Jay Stratton Noller

Dept. of Crop & Soil Science, Oregon State University,
jay.noller@oregonstate.edu

Introduction

Time is a widely recognized factor in conventional soil survey, typically documented as a descriptor of parent material (Soil Survey Staff, 1993), and is considered as the *t* factor of soil formation (Jenny, 1941). In digital soil mapping (DSM), which is the next mode of global soil resource inventory and analysis (Lagacherie et al., 2007), this factor is expressed as *a* (for age) in the formulation of $S_c = f(s, c, o, r, p, a, n)$ by McBratney et al. (2003).

Age (*a*) is used by digital soil mappers in *implicit* or *explicit* form. The norm to date in DSM has been to carry age in the implicit form, typically as age of parent material (*p*) and class of landform (*r*). In nearly all global cases, parent material information is carried into DSM analysis as class data derived from geological and geomorphological (thematic) maps (e.g., Bui and Moran, 2001). These data are represented in GIS as polygon vector or raster representation, with age information conveyed as geological time-series names. From Precambrian to Holocene, the spectrum of time-series names generally connotes the age range and type of earth materials from ancient, hard and difficult-to-weather crystalline rocks to recent, soft and weatherable unlithified sediments. A ubiquitous geological map unit representing the latter is Quaternary alluvial (fluvial) deposits, present along the world's waterways.

Explicit use of age as a co-variant in DSM is rare (e.g., Scull et al., 2005). It is this form of age information that, if operationalized, would present the most advantages to the digital soil mapper. Explicit forms of age information could be represented in GIS in the form of point observations, area classes, or continuous rasters collected by a remote-sensing method. Point data would be derived from field-sampled soil profiles. Many of the known geochronological methods, particularly those yielding numerical age results (the most sought-after, high confidence type). Table 1 could be used in this regard. Lithostratigraphic sections, represented as map point data, are typically used to associate age with area classes of surficial geological map units. Soil surveyors or geomorphologists/geologists could step in and date surficial geological units of interest. Spatially continuous observations of soil age, including proxies, from remote sensors would provide explicit measures of properties related to age. Currently, age is not directly determined from such sensor data without the use of a calibration data set, such as a chronosequence (e.g., Kahle et al., 1998). And so the best results for the continuous data set are a calibrated age result (Table 1) (Noller et al., 2000).

Geochronology

Establishing the age of soil is not an easy nor straightforward enterprise. Soils develop over a considerable range of time and these biological, chemical and physical changes are not considered to be as "instantaneous" as perhaps the formation of its constituent minerals and/or organismal tissue. Most strategies to date soils focus on estimating the age of origin for its parent materials. The difference in time between this origin age and today is considered the *duration* of pedogenesis. The use of duration of pedogenesis, or maturity of soil profile, is a widely held concept in soil survey (Soil Survey Staff, 1993). For most study areas, we generally assume the age of the parent material, provided it formed during the Quaternary, is the duration of soil aging.

Age information for soil studies is not just a measure of time reported in years. It is important to consider the source of material that was dated, context (or stratigraphy) of the sample, method applied, type of result, and community confidence. These topics are considered in depth by Noller et al. (2000). A brief overview of key points follows.

The 35+ known and applied methods of Quaternary geochronology classify into six types based on 1) principles and laws, 2) biological, chemical and physical processes, 3) accumulated or complexes of results of processes, and 4) logical arguments (Noller et al., 2000) (Table 1). These six types of methods yield age estimates that are classified into four types of results (Table 1). At least eight of the dating methods would be amenable to remote detection (Table 1) and hence of considerable promise and use in digital soil mapping. Before considering the use of age in a DSM project some key questions should be asked. What type of temporal information is available and can it be used? Is there enough geochronological data to support use of an age layer? Is it appropriate, within the study context, to use age as one of the environmental covariates?

Digital soil mapping as a field of study developed and has been practiced with considerable success without specifically using soil age. If we look at the successes of (predictive) DSM around the globe, one might be inclined to say it is not important to muddle our models with another covariate. On the other hand, one could argue that soil age has been used all along, albeit in implicit form, and thus contributed to successes without operator knowledge. I generally believe this to be true and so urge the community to cognizantly consider soil age the next time we approach a DSM project.

Experimental Setups

My students and I have been examining the potential roles of age carried in implicit and explicit forms to the predicitive digital soil mapping models. Estimates of age used in these studies were collected from the literature, derived using established methods and handled and reported using geochronologic community protocols (Noller et al. 2000). Methods for the establishing the ages of soil parent materials in Oregon have focused on ^{39}Ar-^{40}Ar geochronology of basalt rocks and radiocarbon (^{14}C) dating on organic matter in sediments. Argon and radiocarbon geochronology data were collected from the literature and their geographic locations were inducted as a point-data layer in ArcGIS.

In several experimental setups, surficial geology is mapped on the basis of geomorphic expression of landforms and other criteria standard in geological and geomorphological field studies. Surficial geology polygon vector data is attributed with deposit/flow/landform age based on point geochronology and surface relative–dating techniques using field techniques, aerial photo interpretation, and digital terrain model (e.g., hillshade, topographic contours). Cross-cutting relations, vertical separation, and surface-

Table 1. Classification of Quaternary geochronologic methods and their application to digital soil mapping

Type Of Results

═══════Numerical-Age═══════[1]———————————————————————————————————

——————————═════ Calibrated-Age ═══════[1]——————————

——————————═════ Relative-Age ═══════[1]——————————

——————————————————————═══Correlated-Age═══[1]

Type Of Method

Sidereal	Isotopic	Radiogenic	Chemical and Biological	Geomorphic	Correlation
Dendrochronology	Radiocarbon[3]	Fission track	Amino-acid racemization	***Soil-profile development***[3]	Stratigraphy[3]
Sclero-chronology and annual growth in other organisms (e.g. mollusks)	Cosmogenic isotopes[3] ^{36}Cl, ^{10}Be, ^{26}Al, ^{14}C, ^{3}He, and others[2]	Thermoluminescence	***Obsidian hydration and tephra hydration***[3]	***Rock and mineral weathering***[3]	Paleomagnetism
Varve chronology	K-Ar and ^{39}Ar-^{40}Ar[3]	Optically stimulated luminescence	Rock-varnish cation ratio	***Scarp morphology and other progressive landform modification***[3]	Tephrochronology
Historical records	Uranium-series	Infrared stimulated luminescence	***Lichenometry***[3]	***Rock-varnish development***[3]	Paleontology
	U-Pb, Th-Pb	Electron-spin resonance	***Soil chemistry***[3]	Rate of deposition	Archaeology
			^{10}Be accumulation in soils[3]	Rate of deformation	Astronomical correlation
				Geomorphic position	Tectites and microtectites
				Stone coatings ($CaCO_3$)	

[1]Triple-dashed line indicated the type of result most commonly produced by the methods below it; single-dashed line indicated the type of result less commonly produced by the methods below it.

[2]***Bold-italicized*** methds are particularly useful in Digital Soil Mapping

[3]Methods used in this paper.

Table after Noller et al. (2000).

weathering features are typical criteria used to assess relative surface (and by extension soil) age. Where vegetation cover is sparse, remote sensing works well to discriminate a number of time-dependent characteristics, including changes in surface (e.g., volcanic glass) reflectance (Kahle et al., 1988), collapse and smoothing of microtopography (Farr, 1991), lichen growth (Stretch and Viles, 2002), and pedogenesis (Vaughan, 2008). Age classes of surficial geologic map units used in these studies are presented in Table 2.

Overlapping age classes, for example, Latest Pleistocene and Quaternary (Table 2), present regional problems in fully utilizing the power of age in (predictive) DSM. The age classes indicate uncertainties in the age assigned to the geological map unit, which at local level was probably sorted out by the geoscientist during their mapping work. For our local geoscientist, they would indicate the relatively narrow age range of Latest Pleistocene for a fluvial terrace. Whereas, at regional level where two or more geological works are involved, the correlation of map units across study boundaries will likely err to a conservative, higher-order age class. In this example it would be Quaternary, and it would be in accordance with established norms of the International Stratigraphic Guide (Salvador, 1994). Implied here is that soils within the subject surficial geological map unit would be interpreted by one's DSM algorithm to have had up to 2.3 million years to form. This is several orders of magnitude off and the possible morphologies of predicted soils would likely be quite diverse and variable if one uses the less certain and potentially longer period of soil aging in one's predictive models. Compare that with the local situation where the map unit is constrained to an age of less than 15,000 years. The uncertainty in soil age is of considerably lesser magnitude, and resulting predictions are likely more successful. To circumvent this problem we use something akin to fuzzy sets of geological age (Table 2). Rather than use Quaternary in the above example we could use Late Pleistocene, which probably carries the field geologist's uncertainty but not the 2+ million-year range of time.

Table 2. Ages assigned to surficial geologic map

Years B.P. (Thousands)	Geological Period/Epoch	Symbol
0	Active	A
0-4	Late Holocene	Hl
4-7.5	Middle Holocene	Hm
7.5-10	Early Holocene	He
10-15	Latest Pleistocene	Pll
10-125	Late Pleistocene	Pl
0-125	Late Quaternary	Ql
125-700	Middle Pleistocene	Pm
700-2300	Early Pleistocene	Pe
0-2300	Quaternary	Q
2300-5000	Pliocene	P
700-5000	Plio-Pleistocene	Pp
2300+	Tertiary	T
1500-5000+	Tertiary-Quaternary	TQ

Our experiments on age in DSM were run using decision-tree analysis (See5 -- www.rulequest.com; Imagine 9.1 -- gi.leica-geosystems.com) following the reference area approach (Lagacherie et al., 1995; Lagacherie and Voltz, 2000; Scull et al., 2005). Experiments were designed considering both implicit and explicit modes of age information exist for use in creating predictive DSM map (e.g., Hash, 2008; Noller, in press). Between these forms, a number of forms were selected as experimental constructs: age implicit in remote sensing data; age implicit in lithological (thematic) map; age explicit in geochronological (thematic) map; and age explicit in geological (thematic) map. Control experiments were run without dependent data layers that would obviously carry geological age information.

Table 3. Environmental variables used to develop predictive models.

Independent variable	Symbol	Cell Size	Source
TOPOGRAPHY			
Slope		30m	Derived from 10m DEM
Aspect		30m	Derived from 10m DEM
CLIMATE			
Mean annual temperature	MAT	800m	PRISM climate data[1]
Mean annual precipitation	MAP	800m	PRISM climate data
Mean Jan min. temp.	Tmin	800m	PRISM climate data
Mean Jul max. temp.	Tmax	800m	PRISM climate data
VEGETATION			
Tasseled cap transformation			
Wetness index	wi	30m	Landsat TM, acquired 7/5/1989
Normal. Vegetation Index	NDVI	30m	Landsat TM, acquired 7/5/1989
PARENT MATERIAL, TIME			
Lithology	L	30m	Original work this project
Geologic age	a	30m	Original work this project
Geology (L + a)	La	30m	Original work this project
Fe-oxides (Landsat b3b1)	L31	30m	Landsat TM, acquired 7/5/1989

[1]PRISM data source is http://prism.oregonstate.edu.

Results

If one wants to test the influence of implicit and explicit forms of soil age on (predictive) digital soil mapping, five experimental setups should be adequate. I use some recent results from the Jordan Volcanic Field in eastern Oregon, USA (Fig. 1) to illustrate this (Noller, in press). The base setup, experiment 1, is run with input layers (first eight listed in Table 3), without lithology and age information. Environmental data that might strongly carry this sort of information, e.g., Landsat imagery bands, were excluded from all runs. The experiments are numbered in order of the addition of increasingly more specific (explicit) age information. In experiments 2 through 5, only one layer of information is added to the base for a total of nine.

Absence of age information is taken here to mean geological map units classed according to lithology but not age. So, the most meaningful soil age information is of the implicit form, and only in terms of weatherability of the parent material. Weatherability would relate to rate of soil aging, but is not an indication of soil age. In this matrix of environmental covariates, the landform map carries units that are not differentiated on the basis of activity, preservation of original topography, other age implicit or explicit forms. Assessment of accuracy of prediction without geology or soil age information in experiment 1 yielded the lowest values of the experiments (Overall Accuracy OA=78%, Khat=0.84) (Table 4).

Traditional addition of a parent material or geology (lithology) layer in experiment 2 yielded improved results (OA=82%, Khat=0.88). In experiment 3, the addition of a Fe-oxide reflectance (b1b3) layer yielded values (OA=79%, Khat=0.84) that are essentially unchanged from the experiment without geology and age. The lack of significant change between experiments 1 and 3 may be that while b1b3 is directly tracking the surface weathering of exposed rock (in this case, basalt flows), it is also tracking the inverse of the vegetation cover (greenness), which is greatest in areas of oldest (Tertiary) bedrock and youngest (Holocene) fluvial landscapes. In the final results, the b1b3 layer was ranked lower in producing the prediction than the vegetation layers.

The presence of explicit age information in experiment 4 significantly increased prediction (OA=83%, Khat=0.89) (Table 4). In the final go of experiment 5, the addition of geology and surface (soil) age yielded the most accurate of the experimental setups (OA=86%, Khat=0.94). Overall, incremental additions of age information yielded a corresponding incremental increase in the successfulness of prediction results.

Table 4. Results of age experiments in predictive soil map of Three Mile Hill Quadrangle, Oregon[a]

Experimental Setup	Accuracy Assessment			
	Overall	Producer's	User's	Khat
1. No lithologic/age information	78.2	73.4	65.9	0.84
2. Lithology only	81.8	75.1	68.7	0.88
3. Implicit age information only[b]	78.7	73.4	66.7	0.84
4. Explicit age information only[c]	82.8	79.4	70.2	0.89
5. Lithology and (soil) age[d]	85.9	78.1	71.5	0.94

[a] Experimental results of a study on a Miocene to Holocene age volcanic field in Oregon (Noller, in press).
[b] Geomorphologic relative age (Fe-oxide reflectance) is linearly related to surface character; no other geology-related inputs.
[c] Geochronologically established age classes of geological (thematic) map units; no other geology-related inputs.
[d] Combined lithologic and age (geological thematic) map units.

Discussion

Because age is in the mind of the soil surveyor as "*t*" of Jenny's (1941) soil-forming factors and in the digital soil mapping model as "*a*" of McBratney et al. (2003), it follows that inclusion of this factor will improve study results. However, due to a general lack of numerical methods for dating soils, age is commonly expressed as duration of exposure of the surface of the parent material to the atmosphere and pedogenic process (Birkeland, 1999). As discussed by pedologists, e.g., Jenny ibid.), McBratney et al. (ibid.), and Birkeland (ibid.), time is involved in all environmental factors of soil formation and so it may remain difficult, if not impossible, to fully control for the age/time factor in DSM.

Although the studies are few in which age is used for digital soil mapping, as suggested by McBratney et al. (2003) and a review of the literature for this paper, nearly all DSM studies incorporate data layers that carry implicit age information. It is reasonable to presume that some of these layers carry spatial information on (1) old vs. young soils, surfaces and parent materials, (2) actively changing vs. equilibrium landscapes, and (3) fast vs. slow rates of surficial processes. In disaggregating their geoinformation for digital soil mapping study, Bui and Moran (2001) came very close to explicitly demonstrating this very point.

The addition of geological event(s) could improve prediction. Apart from the static data layers repre-

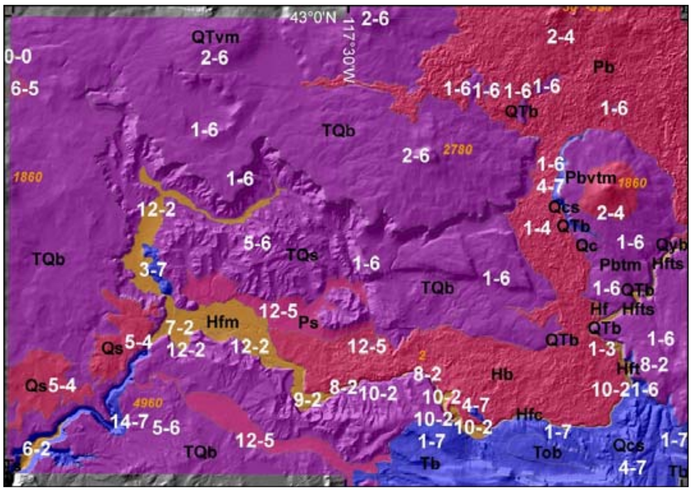


Figure 1. Age information useful in digital soil mapping is presented in this surficial geologic map, overlain on hillshaded IFSAR digital terrain model of the Jordan Volcanic Field in eastern Oregon. Colors of thematic geologic age classes (labels are black) range from Holocene (young)(yellow [warm] colour) to Miocene (old)(blue [cold] colour). Lithology of the map units are (o,y)b = (old,young) basalt, fc=fluvial channel; fm=fluvial meander belt, ft(s)=fluvial terrace (sand), (c)s=(coarse)sand, vm=mafic vent, vtm=Three Mile vent Dated basalt lava flows and vents are labelled in gold italics at point of geochronological sampling. Map units are labelled in white for lithology and age category (Table 2) (e.g., 5-6).

senting the typical DSM environmental covariates (Lagacherie et al., 2007), a sense of the dynamism in surficial processes can be expressed. Some soils developed under conditions quite different than today. By providing indications of the nature of soil parent materials and/or pedogenesis-altering conditions, paleoenvironmental maps interject information that is atypically carried in maps of geology, modern climate, and hydrology. Examples include glacial geology, paleofloods, and tephra distribution.

Summary and Conclusions

The addition of age information as inferred geological period/epoch classes and as numeric geochronological data yield significant improvements in the accuracy of prediction for digital soil maps. It may well be impossible to conduct a digital soil mapping study without any sort of age information or dependence on time. However, it is important to recognize that many of us already (perhaps unknowingly) use implicit form(s) of soil age. It is time to move towards applying explicit forms of the age factor in digital soil mapping studies.

References

Birkeland, P.W. 1999. Soils and geomorphology. New York, Oxford University Press.

Bui, E.N., and Moran, C.J. 2001. Disaggregation of polygons of surficial geology and soil maps using spatial modeling and legacy data. Geoderma 103:79-94.

Farr, T.G. 1992. Microtopographic evolution of lava flows at Cima Volcanic Field, Mojave Desert, California. J. Geophys. Res. 97 (B11):15,171–15,179.

Hash, S.J. 2008. Use of decision tree analysis for predictive soils mapping and implementation on the Malheur County, Oregon initial soil survey. Masters Thesis, Oregon State Univ., Corvallis.

Jackson, T.A., and Keller, W.D. 1970. A comparative study of the role of lichens and "inorganic" processes in the chemical weathering of recent Hawaiian lava flows. American Journal of Science 269:446–466.

Jenny, H. 1941. Factors of soil formation. McGraw-Hill, New York.

Kahle, A.B., Gillespie, A.R., Abbott, E.A., Abrams, M.J., Walker, R.E., Hoover, G., and Lockwood, J.P. 1988. Relative dating of Hawaiian lava flows using multispectral thermal infrared images: a new tool for geologic mapping of young volcanic terranes. J. Geophys. Res. 93(B12):15,239–15,251.

Lagacherie, P., Legros, J.P., and Burrough, P. 1995. A soil survey procedure using the knowledge of soil pattern established on a previously mapped reference area. Geoderma 65:283-301.

Lagacherie, P., and Voltz, M. 2000. Predicting soil properties over a region using sample information from a mapped reference area and digital elevation data: a conditional probability approach. Geoderma 97:187-208.

Lagacherie, P., McBratney, A.B., and Voltz, M., Eds. 2007. Digital Soil Mapping: An Introductory Perspective. Elsevier Science & Technology, Amsterdam.

McBratney, A.B., Mendonça Santos, M.L., and Minasny, B. 2003. On digital soil mapping. Geoderma 117:3-52

Noller, J.S. n.d., Applying geochronology in predictive digital mapping of soils, in J. Boettinger, D. Howell, A. Hartemink, A. Moore and S. Kienast-Brown, eds., Digital Soil Mapping: Bridging Research, Environmental Application, and Operation. New York, Springer.

Noller, J.S., Sowers, J.M., and Lettis, W.R., Eds. 2000. Quaternary Geochronology: Methods and Applications. American Geophysical Union Reference Shelf Series. American Geophysical Union, Washington, D.C.

Salvador, A. 1994. International Stratigraphic Guide: A Guide to Stratigraphic Classfication, Terminology, and Procedure. Boulder, CO, Geological Society of America.

Scull, P., Franklin, J., and Chadwick, O.A. 2005. The application of decision tree analysis to soil type prediction in a desert landscape. Ecological Modeling 181:1–15.

Soil Survey Staff, 1993. Soil Survey Manual. U.S.D.A. Agriculture Handbook No. 18. U.S. Government Printing Office, Washington, D.C.

Stretch, R.C. and Viles, H.A. 2002. The nature and rate of weathering by lichens on lava flows on Lanzarote. Geomorphology 47:87-94.

Vaughan, K.L. 2008. Pedogenesis at Craters of the Moon National Monument and Preserve, Idaho. Ph. D. dissertation, Univ. of Idaho, Moscow. 156 p.

Scale-specific relationships between soil properties: Hilbert-Huang Transform

Asim Biswas, Lindsay Tallon, and Bing Cheng Si

Department of Soil Science, University of Saskatchewan, Canada

Soil scientists have recognized that variability of soil properties is the rule and not the exception since the early 1900's. In spite of this reality, the systematic study of soil variability only began in earnest during the latter half of the previous century. Furthermore, it has been over three and a half decades since the classic study of Nielsen et al. (1973) on the spatial variability of soil properties. Extensive research has been performed in an effort to understand characteristics of soil spatial variability, such as spatial similarity or dependence, periodicity, scale dependency, and nonstationarity. Myriad analysis techniques have also been developed and applied to address soil spatial variability, including geostatistics, spectral and coherency analyses, fractals/multifractals, wavelets and wavelet coherency. While these techniques have proven to be extremely useful in helping researchers uncover the underlying variability in a system, they typically suffer from the same deficiency; spatial analyses assume the spatial series to be linear. In practice, the total effect from multiple processes is not additive, thus the principle of superposition does not apply. Put simply, the response of the system to multiple processes at different scales cannot be determined simply by observing one process at a scale and subsequently adding the individual results together. In this situation the system cannot be explained by a linear equation and thus it is called a nonlinear system.

Hilbert-Huang transform (HHT) is a new method that has been developed to simultaneously deal with both nonlinear and nonstationary data series (Huang et al., 1998). An advantage of the HHT method is that it does not impose any mathematical rule in the analysis, but instead explains the hidden physical mechanisms directly from the data (Huang and Wu, 2008). Unlike other data analysis methods, there is no a priori basis in HHT; rather it is adaptive and derived from the data (Huang et al., 1998).

The HHT is a two step method. The first step is empirical mode decomposition (EMD), which works directly in the spatial domain with the basis derived from data. EMD separates variations completely based on the frequencies present in a spatial dataset and decomposes a signal into a finite set of oscillatory modes (see Box 1). Each mode is represented by an intrinsic mode function (IMF). IMFs separate the total variation into different scales, and, unlike wavelet analysis, do not depend on mathematical functions or restrictive assumptions. The decomposition to IMF is based on a simple assumption that at a given time or space, there may be different simple oscillatory modes of significantly different frequencies superimposing one other (Huang and Wu, 2008). Each frequency is representative of one scale process. Defining modes or IMFs should satisfy the following conditions: 1) the mode may or may not be linear, but the number of extrema and zero crossings must either be equal or differ at most by one; 2) The oscillation will be symmetric with respect to the local mean; that is, at any data point, the average value of the envelopes defined by local maxima and local minima is zero. According to these definitions, IMFs can be obtained after decomposing any function through a sifting process. Each IMF is comprised of a set of processes of similar scales.

Once the IMFs are separated, Hilbert transforms are easily applied to each IMF as the second step of HHT. Hilbert transform leads to an apparent space-frequency-energy description of a spatial series after separating into different scales, or IMFs, and thus will have intrinsic physical meaning at every point. The energy of each IMF can be calculated from the instantaneous amplitude, which is a function of space. The phase calculated from Hilbert transform is also converted to instantaneous frequency as a function of space. Unlike the fixed frequency band used in Fourier transform, Hilbert transform calculates frequency at every location. By examining the local frequency properties of a spatial series, the details of the nonlinear processes can be achieved. As the energy and the frequency are calculated separately as a function of space, we can express the energy as a combined function of space and frequency in Hilbert spectrum. The variable frequency resolution used in calculating the spectrum provides better spatial resolution than any other method based on uniform frequency. The Hilbert spectrum provides instantaneous frequency information, which is representative of spa-

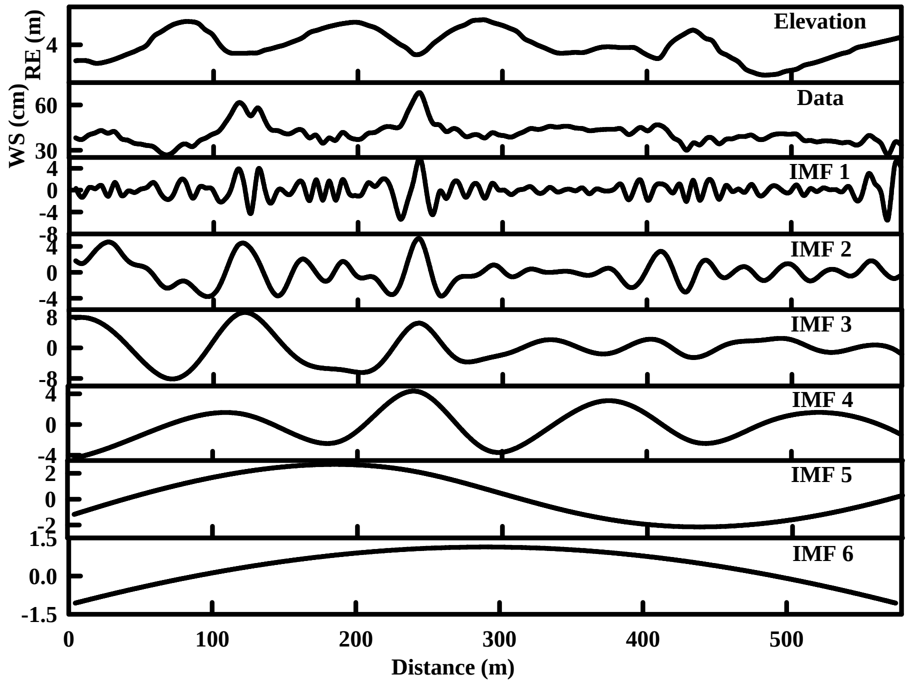


Fig. 1: IMFs of 2nd May 2008 with elevation and the soil water storage data series

tial scales of processes. The scale and location specific information of any soil process can be identified from Hilbert Spectral analysis. The total contribution of each frequency or each scale can be calculated by accumulating the energy over the entire data span from the construction of the marginal spectrum. A marginal spectrum can be calculated for each IMF, which is an alternative spectrum expression of the data to the traditional Fourier spectrum.

HHT separates different scale processes and identifies location specific scale of variation of nonstationary and nonlinear soil spatial variability. Of the methods available for dealing with soil spatial variability, HHT represents a very promising area, as is described in the following case study.

A Case study

Hilbert-Huang transform was used to delineate the scale specific controls on soil water storage in a rolling landscape (Biswas and Si, 2009). Soil water storage measurements were taken at 128 points along a linear transect on different dates in 2007, 2008, and 2009. The measurements taken on 2nd May, 2009 are shown in Figure 1. A number of factors made this study especially suited to demonstrating the utility of the HHT method. The rolling landscape had irregular variations in controlling factors, resulting in uneven spatial means. Therefore, the soil water storage series was nonstationary. Another unique aspect was that the various controlling factors operated simultaneously at different scales and the overlapping of scales made accounting for each individual effect complicated. It was clear that the natural system did not follow the principle of superposition. The nonlinearity in the spatial series made it difficult to separate the effects from different controlling factors, thus making it a good candidate for HHT.

The linear correlation coefficients with respect to measurement scale indicated a good correlation between soil water storage and organic carbon (OC) and soil texture (Fig. 2). Elevation did not correlate well at the measurement scale. Similar correlations were observed for 15 other soil water measurements with very little change in the magnitude over seasons.

Empirical Mode Decomposition (EMD) was used to separate the overall variations in the soil water storage series at different scales directly from the data. Variation at each scale was represented by an Intrinsic Mode Function (IMF). Among the 6 IMFs for each soil water series, IMF 2 and IMF 3 represented majority of the total variations. Figure 1 showed the IMFs extracted from 2nd May 2008 soil water series measured at St. Denis national Wildlife Area, Saskatchewan, Canada (Biswas and Si, 2009). From visual comparison it appeared as if IMF 3 represented the topography.

The linear correlation coefficient between soil water storage and its controlling factors at different scales, as indicated by IMFs, is also presented in Fig. 2. IMF 1, encompassing the highest frequency data, was not correlated to any factors. IMF 2 and IMF 3 were highly

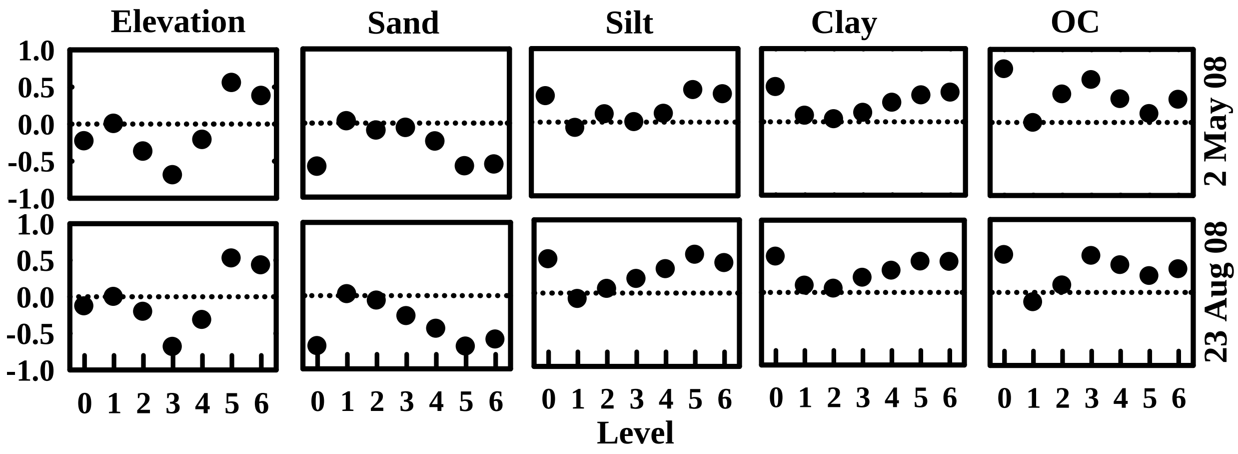


Fig. 2. Scale specific correlation between soil water storage and its controlling factors for two different dates of measurement (Level 0 indicates the untransformed data)

correlated to elevation and OC. Almost all the factors were correlated to IMF 4 indicating a combined control from factors at that scale. IMF 5 indicated from the effects of sand and elevation, while IMF 6 represented a mixed control. The change in the sign (- or +) of correlation over scales between soil water storage and elevation neutralized the overall controlling effect, which indicated a weak correlation at the measurement scale. The correlation between OC and soil water storage was high in IMF 2 and IMF 3, which constituted the maximum contribution towards total variation. The high correlation at important IMFs resulted in a high correlation at measurement scale.

A scale specific regression relationship was built up between soil water storage and the controlling factors at each scale (Table 1). Elevation and OC were the major predictor at IMF 2 and IMF 3, which explained the majority of the variation in the spatial series. Because OC was inversely correlated to elevation, this confirmed that elevation was the major control at those scales.

Hilbert spectrum of 2nd May 2008 (Fig. 3) indicated a strong variation at large scale (80 m or more), which was representative of elevation. Some medium scale variations were also observed.

The HHT method was used successfully in this application, providing a better understanding of the variability found at the site. EMD separated the scale specific process and dealt with nonstationarity. Hilbert Spectral Analysis (HSA) calculated the energy and frequency at each location and each scales, and addressed nonlinearity. There was also the additional benefit that the instantaneous frequencies that vary locally were calculated directly from the data without the help of any mathematical functions. The HHT method, combining EMD and HSA, could be very useful in elucidating the underlying spatial variability in a landscape due to its inherent advantage in dealing with nonstationary and nonlinear spatial series. Despite the apparent advantages, further research on how to identify if a spatial series is linear or nonlinear is needed. HHT currently can only be used for one dimensional spatial or temporal data series, and extension to two or three dimensions will be necessary. Finally, statistical tests need to be developed to make quantitative interpretations of HHT spectra and IMFs.

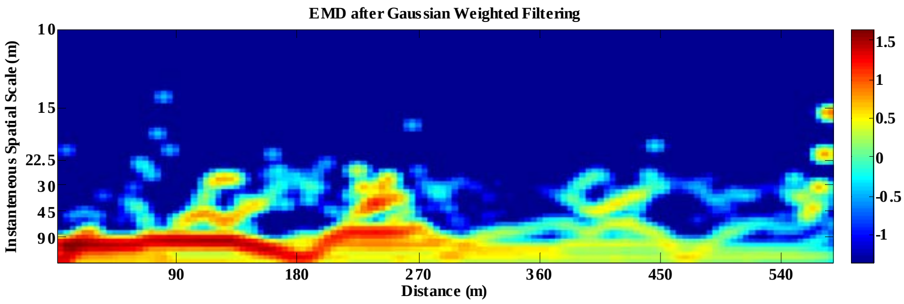


Fig. 3: Hilbert Spectrum of 2nd May 2008 soil water storage

References

Biswas, A., and B.C. Si. 2009. Analysing landscape variability in soil water storage using Hilbert-Huang transform. To be submitted.

Huang, N.E., Z. Shen, S.R. Long, M.C. Wu, H.H. Shih, Q. Zheng, N.C. Yen, C.C. Tung, and H.H. Liu. 1998. The empirical mode decomposition and the Hilbert spectrum

Table 1: Predictive relationship of soil water storage at different scale.

Level	Model
0*	OC + Sand + Ele + OC × Ele + Sand× Ele
1	Intercept only
2	Ele + OC + Ele × OC
3	Ele + OC + Ele × OC
4	OC
5	Sand + Ele + OC + Clay + Silt + Ele × OC + Ele × Clay
6	Sand + OC + Ele + Silt + Sand × Ele + Ele × OC

*- Untransformed data, OC- Organic Carbon, Ele- Elevation

for nonlinear and non-stationary time series analysis. Proceedings of Mathematical, Physical and engineering Sciences. 454(1971):903-995.

Huang, N.E. and Z. Wu. 2008. A review of Hilbert-Huang transform: method and its applications to geophysical studies. Review of Geophysics. 46:RG2006, doi:10.1029/2007RG000228.

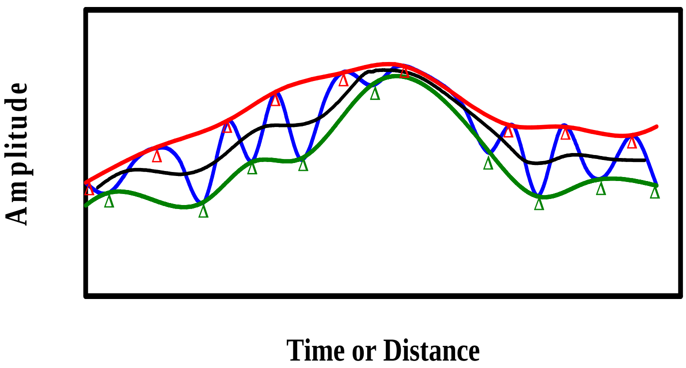
Neilsen, D.R., J.W. Biggar, and K.T. Erh. 1973. Spatial variability of field measured soil water properties, Hilgardia, 42, 7, 214-259.

Bing Si on Hilbert-Huang transform in Pedometrics 2009, Beijing.

Box 1. Empirical Mode Decomposition (EMD)

Here is an example of the conceptual basis for the EMD process.
The steps to obtain an IMF from a spatial data series are summarized as follows.

[a] Identify local maxima (red triangle) and minima (green triangle) of the spatial series $y(x)$ (blue line).
[b] Perform cubic spline interpolation between the maxima and the minima to obtain upper and lower envelopes, $U(x)$ (red line) and $L(x)$ (green line), respectively.
[c] Calculate mean of the two envelopes (black line) $m(x) = (U(x) + L(x))/2$



[d] Obtain $c(x) = y(x) - m(x)$
[e] $c(x)$ is an IMF if the number of local extremes of $c(x)$ is equal to or differs from the number of zero crossings by one. Then, the average of $c(x)$ is zero. If $c(x)$ is not an IMF, then repeat steps a)–d) with $y(x)$ replaced by $c(x)$, until the new $c(x)$ obtained satisfies the conditions of being an IMF.

[f] Calculate the residual $r(x) = y(x) - c(x)$.

Once an IMF is sifted out of the dataset, the residual data remains as the process continues in the same fashion for the remaining IMFs. In following this sifting process, very high frequency (fine scale) processes are initially extracted. Subsequent iterations result in extracting IMFs of continually decreasing frequency (increasing scale). The IMF extraction proceeds until the residue becomes a monotonic function or a function with only one extreme from which no IMF can be extracted. The final residue represents the adaptive trend in the data or is a constant. The original spatial series can be reconstructed by adding all IMFs and the final residue.

Digital Soil Mapping in the Irish Soil Information System

Ron Corstanje, Thomas Mayr, Reamonn Fealy, Joanna Zawadzka,
Giuseppe Lopapa, Rachel Creamer, and Roger Schulte.

There is a growing need in Europe to support policy-makers with a harmonised soil information system and the current status of available information on soils in Europe is inconsistent at best. A soil map and information system at the scale of 1:250,000 has been identified by the EU as an economically feasible intermediate scale and the proposed approach must take into consideration existing methodologies, e.g. the SOTER project (Dobos et al., 2005), and Directives, e.g. the INSPIRE directive, [Directive 2007/2/EC, Infrastructure for Spatial Information in the European Community]. Harmonised soil data across Europe with a 1:250 000 geo-referenced soil database will allow for exchange of data across member states and the provide the information needed by the European Commission and European Environment Agency for reporting on issues relating to soil quality under a future Soil Framework Directive. Within this context, the Environmental Protection Agency of the Republic of Ireland commissioned a project run by Teagasc to produce a 1:250 000 soil map of the Republic of Ireland. Delivery of this map and associated database is a collaborative effort between Teagasc, the National Soil Resources Institute at Cranfield in the UK and University College Dublin.

Figure 1 Map of the Rep. of Ireland organized by counties

In Ireland, a complete set of soil information at the target scale identified at European level (1:250 000) does not exist. Results from a study by Daly and Fealy (2007) indicate that the soil data coverage of Ireland is incomplete in both detail and extent. This has created difficulties for users of Irish soil information and has often led to inappropriate use of soil data. The General Soil Map of Ireland only provides a highly generalised and often inappropriate level of information for the many national applications for which it is used. The overall objective the Irish Soil Information System (ISIS) is to conduct a programme of structured research into the distribution of soil types over the whole of Ireland and construct a soil map, at 1:250 000 scale, which will identify and describe the soils according to a harmonised national legend. The project contains a unique combination of soil taxonomic efforts and digital soil mapping which will inform subsequent field work by pedologists to generate the 1:250 000 soil map of the republic.

The project will begin with an intensive study and analysis of the areas surveyed in detail that cover 44% of Ireland (Terra Cognita). Soil class criteria will be reviewed and redefined where appropriate to augment the current Irish soil classification. The General Soil Map (GSM) of Ireland currently comprises 367 soil series and surveys of an additional four counties, surveyed since the production of the GSM, have identified a further 112 soil series. This classification will have to rationalized to produce a consistent, robust soil classification on which to base the 1:250,000 soil map. This classification will be correlated to wider international soil classification standards, principally the World Reference Base classification (WRB) (FAO, 1998; 2006).

As the soil classification is rationalized, parallel efforts will be placed in generating predictive model of the soil series of Ireland with a view to informing the subsequent field programme (Digital Soil Mapping). These efforts are divided into a number of key phases of activity, each designed to determine the most robust preliminary landform classification that will form a key resource for the following field programme. In first phase, the methodologies proposed will address and assess two different approaches available for predictive soil mapping as suited to the Irish context; i) Physiographic Soilscapes mapping and ii) Digital Soil Mapping (such as Stratification, Soil association level analysis, Soil series level analysis and Features space analysis).

Physiographic Mapping

As there is no physiographic map for Ireland, a first impression of the landscape units will be obtained using a combination of a modified Hammond approach (Figure 2), Iwahashi and Soter. This can be implemented relatively quickly. More sophisticated approaches will be applied, such as a modified LENZ approach. This is a hierarchical system for mapping land environments based on data on terrain, climate and soil. This method would require a new set of parameters developed specifically for this project and a modified MacMillan hill-shed analysis using LandMapR, in which three different approaches are considered to determine landform units at the national, regional and local scale. The modified MacMillan is a more experimental approach but relies on DTM data which are currently available for Ireland.

Digital Soil Mapping

The digital soil mapping efforts will be based on those environmental covariates that form part of the Scorpan factors of soil formation (McBratney et al., 2003) and can be used in the soil-landscape analysis. These include terrain, soil (small scale), geology (parent material), climate and land use information. A number of inference techniques will be considered as part of this project including See5, Netica and Random Forests. See5 uses decision trees whereas Netica uses belief networks.

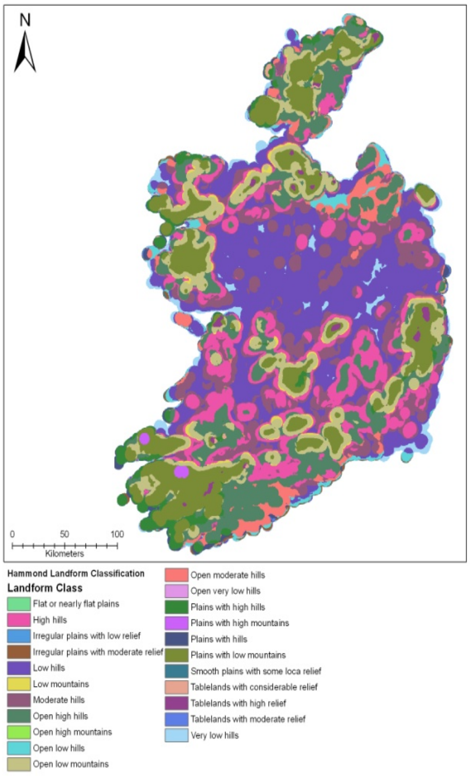


Figure 2 Hammond classification of the Rep. of Ireland

In the second phase of this part of the project, the base delineations for all of the previously unsurveyed area of the country, referred to as terra incognita will be developed. This involves extending the results from the spatial inference engines into areas where the soil information is sparse, about 66% of Ireland. Extrapolation of the modelling efforts will need to consider physiographic units in this landscape, as well as ensuring that there is correspondence in the feature space used for model development. Map unit definitions as well as scale issues will need to be fully considered in this activity. There are a number of existing approaches that may also be considered such as region growing. These approaches will be inventoried and assessed in order to choose the most appropriate and pragmatic method. Validation methods that will be assessed are common methods such as bootstrap, or Jack-knife methods; alternatively we may also consider rates of misclassification to obtain measures of performance for the different approaches.

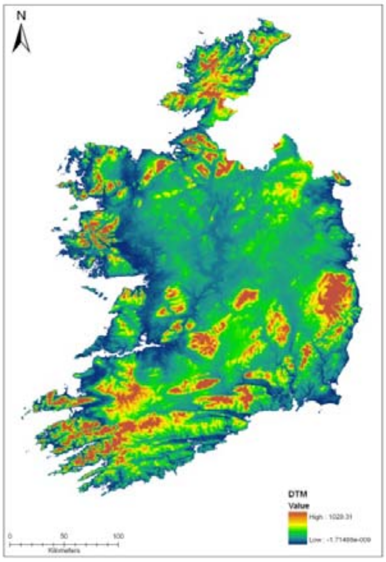


Figure 3 DTM of the Repub. of Ireland

The SCORPAN covariates

The terrain of the republic is characterized by a hilly interior lowland surrounded by a broken boarder of rugged hills and low mountains to the west and south (Figure 3). The geology of the central plain is predominantly carboniferous limestones (Figure 4). To the northeast of this there are the Lower Palaeozoic shales, grits and greywackes, with limestones and younger shales further east. In the northwest is characterized by a complex mixture of shists, quartzites and granite whereas in the south west and east the geology is generally dominated by Old Red Sandstone. Ireland experienced at least two major glacial episodes, an earlier Munster General Glaciation (200,000 to 130,000 years ago) and the Midlandian General Glaciation (75,000 to 10,000 years ago). There are significant glacial deposits (boulder clay) in the southern quarter of the country associated to the earlier period. The later glacial episode generated extensive drumlin deposits in Ulster and Leinster, eskers in the Kildare, Offaly, Galway and Roscommon and thick sand deposits in Curragh.

There are, in broad terms, nine major soil types iden-

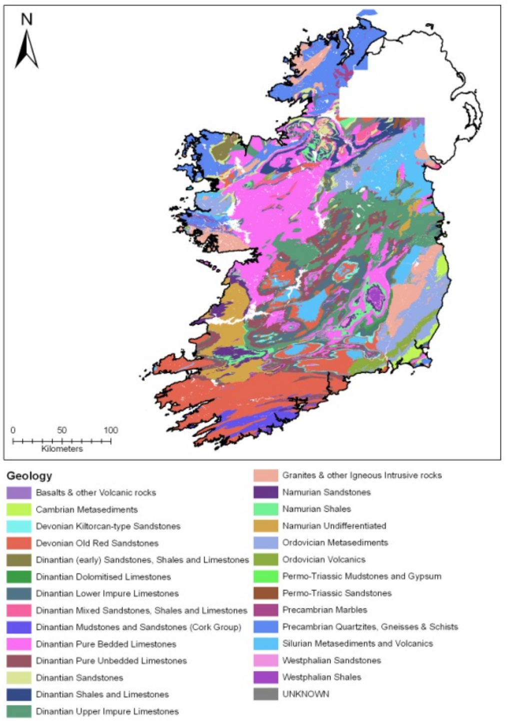


Figure 4 Map of the main geological formations in the Rep. of Ireland

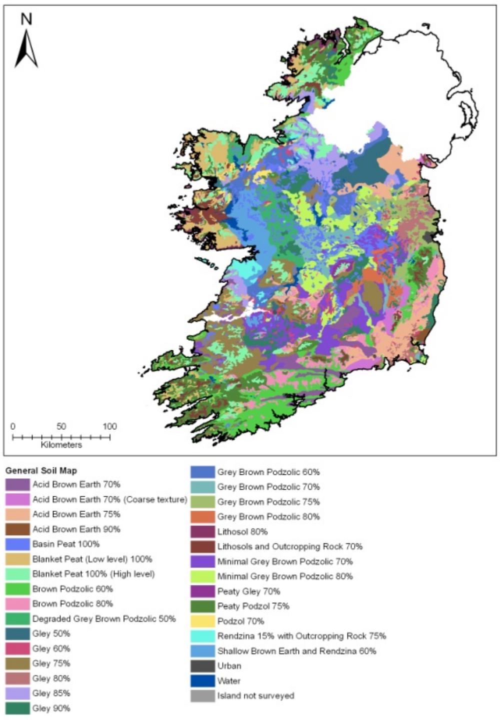


Figure 5 Main Soil groups in the Rep. of Ireland.

tified in the General Soil of Ireland: Podzols, Brown Podzolics, Grey Brown Podzolics, Acid Brown Earths, Brown Earths, Gleys, Rendzinas, Lithosols and Peat. Podozols are typically leached, poorly drained soils that predominate in the mountainous and hill areas; the less depleted Brown Podzolics are mostly observed in the southern and south eastern areas associated with sandstones and shales. Rendzinas and Grey Brown Podzolics are usually formed from calcareous parent material and are therefore primarily found in the central lowlands, underlain by limestone geology. The poorly drained Gleys are often found in association with Basin peat (groundwater dominated) and are found in the central lowlands of Ireland. Blanket peat (rainfall dominated) is found predominately in the western areas of the country and in the upper parts of the mountain ranges. The predominant climate in Ireland is wet and mild, with the majority of precipitation associated to Atlantic systems arriving from the west and moving east. Predominant landuses are pasture, arable, forest or peat (Fay et al., 2007).

The Digital Soil Mapping efforts will inform a field investigation programme on the areas that have not already been surveyed in detail (Terra Incognita). The fieldwork, organized by Teagasc, will focus on boundary checking and map unit composition by examining and describing modal profile pits and auger bores. Representative soil profile pits will need be excavated, sampled and described for soil series for which there is no existing characterization. The field survey should produce the equivalent of approximately 350 map sheets (of 10km x 10km). The result from the various stages of the project will then be integrated in a Soil Information System to deliver the key final map deliverable, the digital polygon based 1:250,000 soil map of Ireland.

About the Authors:

Ron Corstanje, Thomas Mayr and Joanna Zawadzka are part of the National Soil Resources Institute, in the Dept. of Natural Resources, Cranfield University, UK.

Guiseppe Lopapa, Rachel Creamer, Reamonn Fealy and Rogier Schulte are at Teagasc, Rep. of Ireland.

Literature cited

Daly, K. and Fealy, R. (2007). Digital Soil Information System for Ireland (2005-S-DS-22), Final Report. Environmental RTDI Programme 2000-2006, Environmental Protection Agency, Johnstown Castle, Wexford.

Dobos, E., J. Daroussin and L. Montanarella. (2005). An SRTM-based procedure to delineate SOTER

Terrain Units on 1:1 and 1:5 million scales. EUR 21571 EN, 55 pp. Office for Official Publications of

the European Communities, Luxembourg. Available at http://eusoils.jrc.it/projects/soter/index.htm

McBratney, A. B.; Mendonça Santos, M. L, and Minasny, B.,2003. On digital soil mapping. Geoderma, 117(1-2): 3-52.

FAO (1998). World Reference Base for Soil Resources. World Soil Resources Report No.84, FAO, Rome, 88 pp.

FAO (2006). Guidelines for soil description (Fourth edition). Rome, Italy, 97pp. Field handbook Page 28 of 37.

Fay, D. Kramers, G. Zhang, C. "Soil Geochemical Atlas of Ireland". Associated datasets and digitial information objects connected to this resource are available at: Secure Archive For Environmental Research Data (SAFER) managed by Environmental Protection Agency Ireland http://erc.epa.ie/safer/resource?id=4856ff8c-4b2b-102c-b381-901ddd016b14 (Last Accessed: 25/11/2009)

The 4th Global Workshop on Digital Soil Mapping

From Digital Soil Mapping to Digital Soil Assessment: identifying key gaps from fields to continents

Rome, 24-26 May 2010

CNR Conference Room, Piazzale Aldo Moro, 7 - 00185 Roma

http://2010.digitalsoilmapping.org/

Join us at the World Congress of Soil Science, Brisbane,

Pedometrics Symposia

1.5.1 Quantitative monitoring of soil change (Convened by Murray Lark and Tom Bishop).

In this session we will consider the statistical problems of collecting spatio-temporal information on the soil. We will focus on the problems of designing appropriate monitoring and sampling schemes, on the use of information from novel sensing technologies, on statistical methods for spatio-temporal prediction and on integrating multiple sources of information on the soil. A keynote talk will be given by Dick Brus from Alterra, Wageningen.

1.5.2 Modelling critical processes in changing soil (Convened by Andy Whitmore and Matthew Pringle).

In this session we will consider some generic problems raised in quantitative modelling of processes in the soil. There are exciting new developments in the field of modelling which are all pertinent to the specific problems of soil modelling. In particular we will focus on data assimilation and Bayesian approaches to the estimation of model parameters and state variables, and for handling the uncertainty in our resulting estimates. We will consider the problems of predicting soil processes at appropriate spatial scales and of error propagation in process models. The outcomes of the error propagation analyses are essential to strike the right balance between model complexity and data availability. A keynote talk will be given by Gerard Heuvelink from Wageningen University.

1.3 Digital soil assessment (Convened by Florence Carré and Neil McKenzie).

This symposium focuses on Digital Soil Assessment which is the process beyond Digital Soil Mapping. Once the soil map and the associated accuracy have been produced, these serve as inputs for modelling soil processes (threats to soil, soil functions, soil-environment relationships). The accuracy produced during the DSM process should also be used in the soil-process modelling in order to obtain two kinds of outputs: the spatial distribution of the outputs of modelled soil process, and the associated accuracy of the prediction.

1.5 Soil Sense: rapid soil measurements (Convened by Viacheslav Adamchuk and Rapahel Viscarra Rossel).
Conventional methods of soil analysis can be slow and expensive and on occasions the procedures are complex and only qualitative. Proximal soil sensing (PSS) can provide good quality, quantitative, inexpensive soil information. PSS is developing into a vibrant area of multi-disciplinary research that aims to apply state-of-the-art sensing technologies to the study of soil processes and spatio-temporal soil variability. Specific areas of investigation will include sensor development, signal processing, data fusion, pedometrics, spatial modelling, and fundamental research into soil-sensor interactions.

D1.2. Modelling the direction and rates of soil formation in time and space. Convenor – Edoardo Constantini and Budiman Minasny.

Did you miss this? ...

Murray

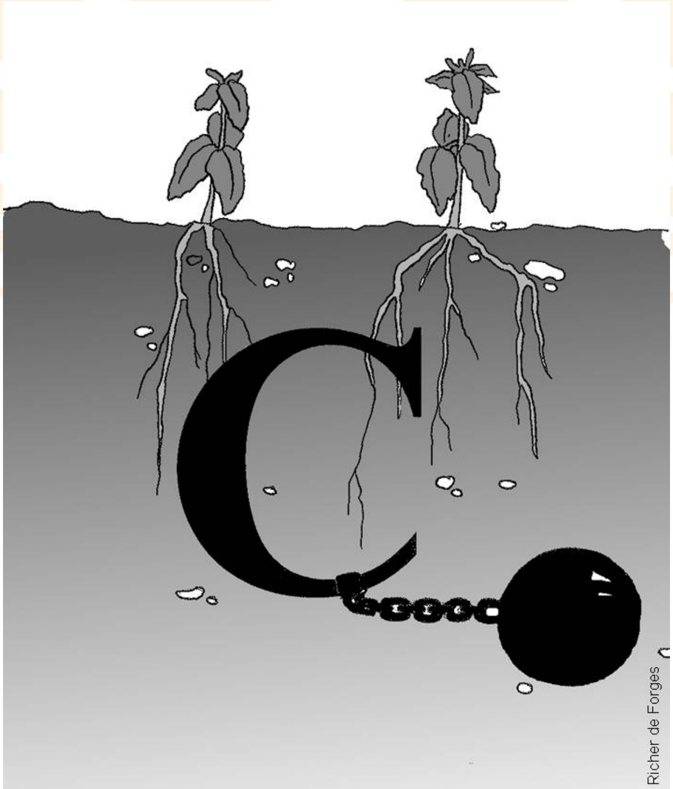
Andersen KE, Brooks SP, Hansen MB (2003). Bayesian inversion of geoelectrical resistivity data. Journal of the Royal Statistical Society Series B-Statistical Methodology 65, 619-642.

Abstract: Enormous quantities of geoelectrical data are produced daily and often used for large scale reservoir modelling. To interpret these data requires reliable and efficient inversion methods which adequately incorporate prior information and use realistically complex modelling structures. We use models based on random coloured polygonal graphs as a powerful and flexible modelling framework for the layered composition of the Earth and we contrast our approach with earlier methods based on smooth Gaussian fields. We demonstrate how the reconstruction algorithm may be efficiently implemented through the use of multigrid Metropolis-coupled Markov chain Monte Carlo methods and illustrate the method on a set of field data.

This should be of interest to anyone working with resistivity data or related information where the goal is reconstructing variables at depth. In 5 years it has had 14 citations according to ISI, but none from the soil science community as yet (apart from the present writer who mentioned it briefly in a review). It is high time that pedometricians and other soil scientists using this technology had a look at this work.

Part of the paper is physics, but the part that interested me most is the spatial statistics. Bucking the trend the authors state that a continuous conductivity field is not the most appropriate model for their data. Rather they want to discretize the field into regions associated with distinct superimposed deposits. This gives the paper its general appeal, because we might want to do something like this with any number of soil variables. Dick Webster showed how a 1-D data set might be optimally partitioned (e.g. Webster, 1978) but this paper provides a solution for higher-dimension cases. The random models is random coloured polygonal graphs in which the nodes and edges are random processes which can be modelled with an appropriate Bayesian approach. This is a great paper, and I urge pedometricians to read and learn from it.

Webster, R. 1978. Optimally partitioning soil transects. Journal of Soil Science 29, 388–402.



Soil Carbon sequestration

http://richer-de-forges.nexenservices.com/

Report from Geomorphometry 2009, Zurich

Bob MacMillan

Bob MacMillan participated in Geomorphometry 2009 in Zurich from Aug 30 – Sept 2. The main purpose of my participation was to keep informed about other efforts, similar to GlobalSoilMap.net, that have a interest in processing digital elevation data and other digital data sets globally or at least for extremely large areas.

This conference actually contained a large number of presentations of direct relevance for the GlobalSoilMap.net project. Perhaps first and foremost were the descriptions of efforts being undertaken in Australia (Gallant and Read) and Europe (Köthe and Bock) to process SRTM DEM data at 30 m (Australia) and 90 m (Europe) grid resolution to reduce artefacts and produce a filtered and cleaned DEM that is more suitable for use to produce inputs for the GlobalSoilMap.net project. Both of these presentations highlighted the significant advantages that can be realised by applying a series of filtering and conditioning routines to the original raw SRTM DEM data. It is obvious that similar procedures would prove equally useful if applied to SRTM DEM data sets for other parts of the world under the jurisdiction of other GlobalSoilMap.net nodes. Gallant has offered to help with efforts in other Nodes if asked.

Also of great interest were several projects that demonstrated that it is indeed possible to process and produce digital output for global scale digital data sets, including global scale SRTM DEM data sets. Reuter and Nelson presented a description of WorldTerrain, a contribution of the Global Geomorphometric Atlas. Peter Guth described processing of global scale SRTM data to identify and classify organized linear landforms (dunes). Peter also provided examples of multiple scale analysis and illustrated what you get to "see" from DEMs of 1 m, 100 m and 2 km grid resolution. Guth intends to publish the many different grids of DEM derivatives he produced for his project and make these processed data available for free and widespread use by others. Marcello Gorini described a physiographic classification of the ocean flood using a multi-resolution geomorphometric approach.

Several authors presented methods that may prove of interest to the GlobalSoilMap.net project. Gallant and Hutchinson described a differential equation for computing specific catchment area that could be applied to produce an improved terrain covariate for use in the GlobalSoilMap.net project. Similarly, Peckham, gave a new algorithm for creating DEMs with smooth elevation profiles that could be used to condition rough SRTM or GDEM data sets to smooth out noise and produce more hydrologically plausible surfaces. This algorithm was of particular interest to the GlobalSoilMap.net project because it appeared to be able to introduce hydrologically and geomorphologically relevant detail into 90 m SRTM DEMs of relatively low spatial detail. Romstad and Etzelmuller described a new approach for segmenting hillslopes into landform elements by applying a watershed algorithm to a surface defined by the total curvature at a point instead of the raw elevation value. The resulting watersheds were bounded by lines of maximum curvature, effectively structuring each hillslope into components partitioned by lines of maximum local curvature. This is harder to explain than to understand when illustrated but it is remarkably simple to implement and may provide a new way of automatically segmenting hillslopes in a simple and efficient fashion.

Metz and others presented an algorithm for fast and efficient processing of massive DEMs to extract drainage networks and flow paths. This is of considerable interest and relevance to the GlobalSoilMap.net project because of the project's need to process SRTM data globally to compute hydrological flow networks and various indices that are computed based on flow networks (e.g. elevation above channel, distance from divide). This algorithm can process data sets of hundreds of millions of cells (11,424 rows by 13,691 cols) in a few minutes instead of a few days (or not at all for some algorithms that fail on data sets this large).

Overall, this was an excellent conference, dominated by leading edge research in the area of geomorphic processing of digital elevation data that is of direct relevance and interest to the GlobalSoilMap.net project. We have much to learn from these researchers and much to benefit from maintaining contacts and working relationships with them.

Soil Bibliometrics

NIR and soil science: a teen-age love story

Véronique Bellon Maurel

Montpellier SupAgro-Cemagref

Near infrared (NIR) spectroscopy is a well-know technique and has been used for more than 40 years for measuring the quality and composition of agricultural and food products. Although first attempted in the 60s (Hart, 1962, Massie and Norris 1965), especially with cereals and fruits, this method has spread rapidly into the food and agricultural field (excluding remote sensing and soil issues). What is the research pathway and trend of this method in soil science is the question of this bibliometric study. More particularly, I am focussing on the use of NIR spectroscopy for agricultural/ environmental soil analysis, excluding urban soils and contaminated soil issues.

First of all, let us have a look at the development of NIR as a characterisation method in food-agricultural products applications.

Figure 1 shows the evolution of NIR publications dedicated to food and agricultural products characterisation, excluding remote sensing, using Scopus database. The lowest curve represents the Scopus output (from journals and conferences) whereas the highest curve represents the overall articles found on this subject (including non-refereed articles, referred to as "more" references in Scopus, which contributes a lot to the database after the 90s).

In 1987, the peak paper production is mainly related to the several chapters contained in "Near-Infrared Technology in the Agricultural and Food Industries", a 330-pages multi-authored book edited by Phil Williams and Karl Norris, who with Osborne and Fearn, published the first NIR book (Near Infrared Spectroscopy in Food Analysis) in 1986. Williams and Norris can be considered as 'Near Infrared discoverers' (although NIR radiation was discovered by William Herschfel, in the early 1800s) because they really brought this technique into application. Again in 1989, a multi-authored handbook edited by USDA/ARS about the analysis of forage quality (USDA/ARS Handbook No.643) gives a boost to the number of publications.

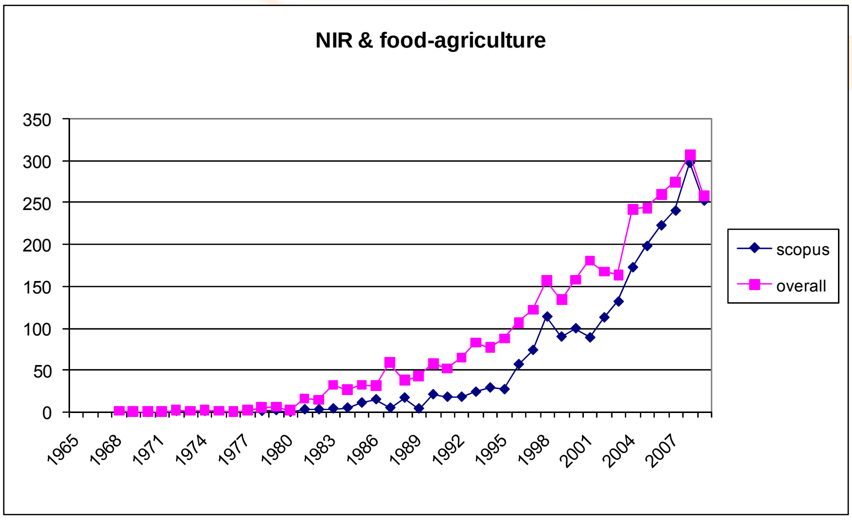


Figure 1: Publications in NIR for food and agricultural products (remote sensing is excluded); in blue: Scopus referred publications and in purple overall publications (provided by Scopus), i.e. including non-reviewed articles.

Actually the first leap of NIR spectroscopy publication for food and agriculture occurred in 1981, with a number of publications leaping from 2 (in 1980) to 16. This first step is followed by book publications in the late 80s as mentioned above. Then in the first half of the 90s, the number of Scopus referred publication is maintained around 20-30 papers per year, whereas the number of "other" publications, mainly conference proceedings keep on growing.

In 1996 and up to 2002, another jump is observed, with the growing of Scopus referred articles (mainly journal articles but also some conferences, in particular

SPIE), this can be seen as the worldwide spreading of the NIR community. A next level is seen in the last 6 years (2004-2009), where the number of publications keeps on growing, supported by the biennial international conference on near infrared spectroscopy, by the publication of several multi-authored books, by the introduction of new themes as the hyper-spectral imaging for food products and of course by the natural spreading of this technology which still poses enough methodological issues for still being studied by research teams. The interesting feature of this community is that it has grown enough to approach the issues of NIR spectroscopy, no longer throughout the application objectives, but with specific questions related to this analytical field, such as instrumentation, light-matter interaction, chemometrics…

So does this give the same trend with studies carried out on soil?

One must be more cautious when conducting a bibliometric study on NIR characterization of soils. Indeed, the outputs are rapidly biased by publications dealing with remote sensing (such as canopy/soil discrimination) or with geological analysis, on earth or on various planets. My aim was definitely not just to give information about the output (number of papers) but I want to focus more on the emergence of a new community of research targeting NIR as a rapid tool, either in the laboratory or in the field, for soil analysis. This is why, after doing a search ("nir" or "near infrared") and soil, I manually removed the papers corresponding to remote sensing or to geological issues (this takes a long, long time).

As shown in Figure 2, the first papers in this field appear at the turn of the 80s and in the early 90s. After the pioneering works of Schumann & Meyer (1989), Ben-Dor & Banin (1990), Sudduth & Hummel (1991), a latency time – with less than 10 publications/ conferences on the subject- was experienced up to the mid 1995 when the community started to be structured and to publish around 15 papers a year. A new level was reached eight years later, after 2004 with more than 25 papers published annually. The trend is still growing and I can foresee that this trend will continue for years because there's a lot of questions are still unsolved in this area. Interestingly, new investigations have appeared in the last 5 years, dealing for instance with new parameters to analyze (in relation with biological properties of the soil) or with the most appropriate data processing techniques to apply to soil data.

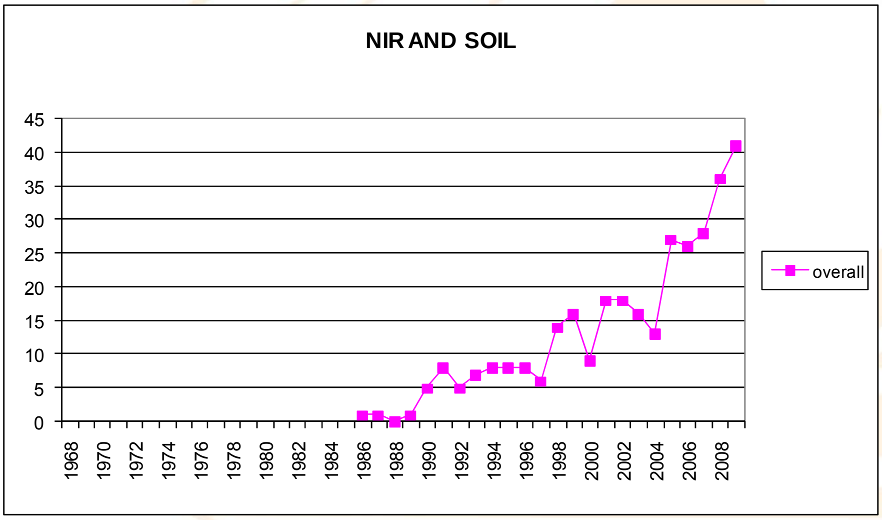


Figure 2: Overall publications in NIR for soil science (remote sensing excluded).

Of course, NIR in soil science - having started later and being limited to a smaller field than food and agricultural products – does still represent a small part or the overall publications dealing with NIR applied to natural resources, as shown in Figure 3. It is also interesting to note that, up to now, only a limited number of researchers who have started their career in NIR applied to food/ agricultural products have swapped to soil. James Reeves III, for instance, is one of them. This means that basically the soil and NIR community is mainly made up of soil scientists who investigate this powerful technique. More crossing-over between the still-

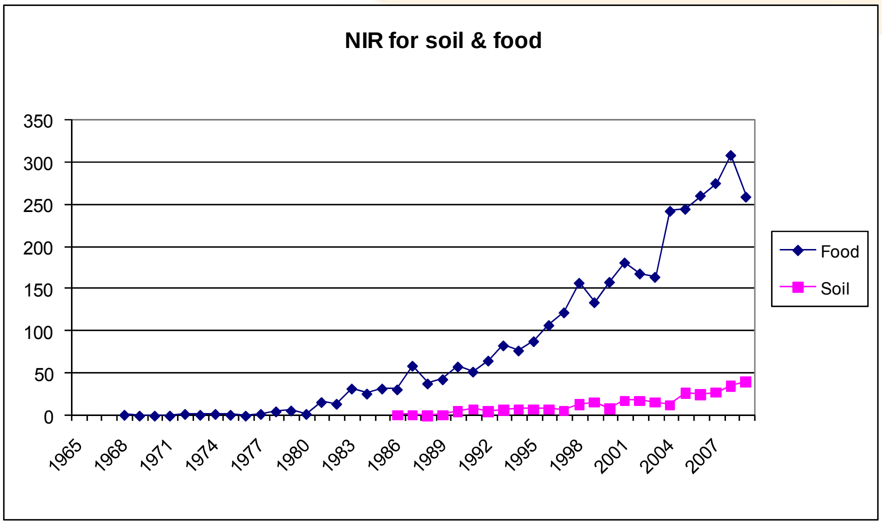


Figure 3: Publications on NIR applied to soil science and to food science.

somewhat-separated communities is un-escapable: for long-time NIR scientists, it opens an exciting field with new paradigms (the heterogeneity of soil samples, the issue of sampling, the light-matter interaction in highly diffusing media), for the NIR-soil scientist community, this is the opportunity to take advantage of the knowledge and advances already experienced and to go further and grow much faster. New spaces that allow better exchanges between these two communities are necessary. As the President of the 2013 International Conference on Near Infrared Spectroscopy ICNIRS, I commit to organizing a session specifically dedicated to NIR spectroscopy applied to soil.

Acknowledgment

This study has been carried out as part of a travelling scholarship of Pr V Bellon Maurel at ACPA, The University of Sydney. This scholarship is supported by the European Commission (IRSES program, IRSES project nr 235108) and the Languedoc Roussillon Council (Regional Plat-form GEPETOS – ECOTECH-LR).

References

Ben-Dor, E., Banin, A. 1990 Near-infrared reflectance analysis of carbonate concentration in soils, Applied Spectroscopy 44 (6), pp. 1064-1069

Hart, J.R., Norris, K.H., Golumbic, C. 1962 Determination of the moisture content of seeds by near-infrared spectrophotometry of their methanol extracts , Cereal Chem. 39 (2), pp. 94-99

Massie, D.R., Norris, K.H. 1965 Spectral reflectance and transmittance properties of grain in the visible and near infrared , Transactions of the ASAE 8 (4), pp. 598-600 18

Schumann, A.W., Meyer, J.H. 1989 Proceedings of the South African Sugar Technologists' Association 73, pp. 72

Sudduth, K.A., Hummel, J.W. 1991 Evaluation of reflectance methods for soil organic matter sensing, Transactions of the American Society of Agricultural Engineers 34 (4), pp. 1900- 1909.

About the Author

Prof. Véronique Bellon Maurel is professor of Agricultural Engineering at Montpellier Supagro, France, and head of the joint research laboratory UMR ITAP (Information and Technologies for AgroProcesses).

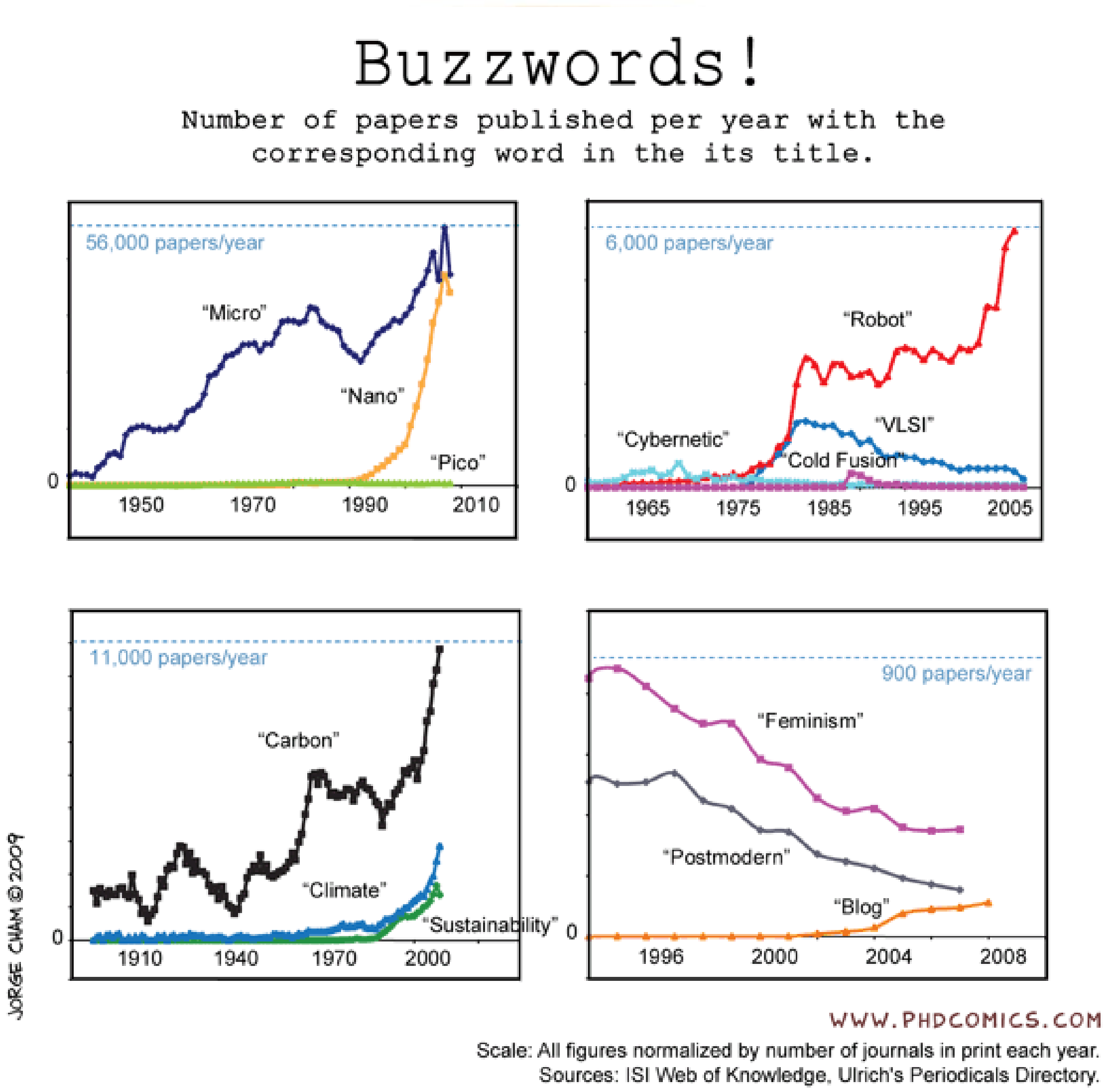
She has been involved in NIR spectroscopy for more than 20 years, mainly for fruit quality assessment. In 2008, she got the Tomas Hirschfeld Award, for her contribution to the NIR community. In 2009, she spent 6 months as a visiting scientist at ACPA, University of Sydney. She will be the convenor of ICNIRS2103, the International Conference on NIR Spectroscopy, to be held in France.

DSM-SSSA

The Digital Soil Mapping Workgroup* under the umbrella of the Soil Science Society of America (SSSA) Div.S5 Pedology was recently formed and invites members to join the Workgroup.

The Workgroup shares and discusses topics of interest by Digital Soil Mappers and maintains a Google Group DSM-workgroup-sssa@googlegroups.com and meets once a year at the annual ASA-CSSA-SSSA Meeting. If you are interested in to join the Workgroup please contact the current Chair (Sabine Grunwald, sabgru@ufl.edu). It is not required that you are a member of SSSA to join the DSM-SSSA Workgroup – everybody interested in DSM is welcome to join.

*** Not to be confused with IUSS's Working Group on Digital Soil Mapping, established at the 2006 World Congress of Soil Science, and advertising its fourth Global Workshop on page 24.**



Pedometrician profile

Sabine Grunwald

University of Florida, USA.

How did you first become interested in soil science?

Way back in graduate school I had a wonderful teacher (Dr. Harrach), who sparked my interest in soil science. He was engaged to teach students about soil mapping and explained which soil forming factors may lead to the development of specific soils. I remember jumping into soil pits and trying to grasp the characteristics and differences of soils. Reading a soil-landscape, learning of pedogenesis, and explaining different soil patterns was fun and inspired me to deepen my study in this subject matter. As an Environmental Scientist by training, I developed an interest in soil science early on in my career, due to the central role of the critical soil zone for transportation and transformation processes that impact biogeochemical cycles (carbon, nitrogen, phosphorus and other cycles) and environmental quality.

How were you introduced to pedometrics?

My Ph.D. on water quality simulation modeling using a mechanistic model to simulate water flux, sediment yield, and nitrogen and phosphorus loads allowed me to gain insight into process-based modeling. However, it also became clear that such models are constrained by soil and other landscape properties that are often not available at a scale and resolution that match reality. I was enthusiastic to learn about the spatial distribution and variability of soil properties as they relate to environmental factors and stressors. The Post-Doc position under supervision of Kevin McSweeney, Birl Lowery and Phil Barak (University of Wisconsin-Madison, Soil Science Department) allowed me to develop holistic 3D soil-landscape models and enhance my skills in geostatistics and digital soil mapping (DSM) techniques. The Pedometrics meeting in 1997 held in Madison, Wisconsin provided the kick-off to connect with the Pedometrics community, which I keep enjoying to engage.

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

The paper by Malone, McBratney, Minasny and Laslett (Geoderma) inspired to look at continuous mapping of soil carbon storage and available water capacity. I am a big fan of the research work of Alex's team that aims to integrate various pedometical methods and datasets and provides novel approaches advancing pedometrics.

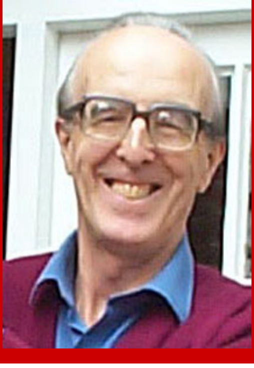
What problem in pedometrics are you thinking about at the moment?

Scaling of soil prediction models considering the extent of the study area, grain (pixel size), aggregation/disaggregation of SCORPAN factors, and spatial dependence structures of soil and environmental properties. We can learn much from other disciplines such as landscape ecology (landscape indices), remote sensing, and hydrology, which have focused research on up and down-scaling of properties and processes. Large amounts of soil data are needed to conduct research on spatial and temporal scaling of soil models. Our research team at the University of Florida has started to focus on scaling of soil carbon pools considering variations in space and time.

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

Synthesis of soil and environmental data sets, including legacy and reconnaissance soil data, as well as integration of soil and remote sensing techniques into multi-tier soil prediction models will facilitate to address topics of profound importance such as global climate change, food security, and degradation of land and water resources.

Pedometrics not applied for the sake of developing another mathematical routine, but in context of environmental problems, soil risk assessment, and ecosystem service valuation will elevate its value and contribution to science.



Non-Pedometrician profile

Anthony Young

How did you first become interested in soil science?

When I completed a PhD in Geomorphology in 1958 I realised there were too many geomorphologists around and not enough soil scientists. One of my lecturers at Cambridge had been approached by the Colonial Office to know if there was anyone suitable for the post of Soil Surveyor, Nyasaland (Malawi). There was a history behind this, involving C. F. Charter's favourable experience in recruiting Cambridge Geographer Hugh Brammer for the Gold Coast (Ghana). At interview I was asked if I would like training in soil survey at the Imperial College of Tropical Agriculture, Trinidad; but, being destined for the African Rift Valley, they sent me instead to train under Alan Crompton in the Vale of York.

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

What is the role nowadays for the study of soil in the field? Nowadays, not many organizations are doing soil surveys, in the sense of making soil maps. I would like to see more soil monitoring, especially in developing countries. But even more, it would be very desirable to develop a role for the 'Soil Management Adviser', as a backup specialist comparable to Entomologist, Plant Pathologist, etc., to agricultural advisory services.

Even more generally, who is looking at soils, their management, and their problems in the field? Where does one find soil scientists helping farmers with such problems?

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

Since I retired, remote sensing linked with statistical methods is being increasingly used. Others can give more information.

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

Quite probably, but refer to other informants.

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

To apply pedometrics to soil monitoring, specifically soil degradation; with a view to putting estimates of soil change on a reliable statistical basis.

Answers to Pedomathemagica Pedometrics 2009 Special

1. When the game was interrupted there was only one possible future situation in which Bert would win: if the next three pits did not contain iron pans; otherwise Alf would win. The probability that Bert would have won is $(½)^3$ =⅛, and since the only other possible outcome is that Alf won, p_a is ⅞. Alf should therefore have 70 coins and Bert 10.

2. This second problem is a little more challenging. If you start listing the possible outcomes it becomes apparent that the probability that Alf wins, p_a, is

$$p_a = \frac{1}{3}\left\{\left(\frac{2}{3}\right)^0 + \left(\frac{2}{3}\right)^3 + \left(\frac{2}{3}\right)^6 + \left(\frac{2}{3}\right)^9 + \cdots\right\},$$

similarly

$$p_b = \frac{1}{3}\left\{\left(\frac{2}{3}\right)^1 + \left(\frac{2}{3}\right)^4 + \left(\frac{2}{3}\right)^7 + \left(\frac{2}{3}\right)^{10} + \cdots\right\}$$

$$\text{and } p_c = \frac{1}{3}\left\{\left(\frac{2}{3}\right)^2 + \left(\frac{2}{3}\right)^5 + \left(\frac{2}{3}\right)^8 + \left(\frac{2}{3}\right)^{11} + \cdots\right\}.$$

If we write

$$\varphi = \left\{\left(\tfrac{2}{3}\right)^0 + \left(\tfrac{2}{3}\right)^3 + \left(\tfrac{2}{3}\right)^6 + \left(\tfrac{2}{3}\right)^9 + \ldots\right\},$$

then $p_a = \frac{1}{3}\varphi, p_b = \frac{2}{9}\varphi$, and $p_c = \frac{4}{27}\varphi$.

Since $p_a + p_b + p_c = 1,\ \varphi = {}^{27}/_{19}$.

Substituting this into the three expressions above gives

$$p_a = \frac{9}{19}, p_b = \frac{6}{19}, p_c = \frac{4}{19}.$$

A quicker solution (thanks to Kathy Haskard) is to notice that at any stage in the game, while it is still running, the probability that the next person to play will turn out (perhaps at a later stage) to be the winner is a constant (and equal to p_a before the game begins). It therefore follows that, before the game begins,

$$p_b = \frac{2}{3}p_a, \text{ and } p_c = \frac{4}{9}p_a.$$

3. By 'outcome' I mean the binomial variable: *Pedometrician 128 sits in his proper chair*, or *Pedometrician 128 has to sit in another chair*. If Alf is the *j*th in the sequence then he can select at random from chairs *j* to 128, all pedometricians 1 to *j*-1 will have sat in their proper chairs. Let p_j = 1/(128-*j*+1). If *j*<128 then p_j <1 (we are told that *j* ≤ 64); and p_j is both the probability that Alf will sit in his own chair (in which case it is fixed that all remaining pedometricians, including 128, will sit in their proper chairs) *and* the probability that Alf will sit in chair 128. With probability 1-2p_j he will sit in some other chair, in which case the outcome remains undetermined for the present — let us say he sits in chair *m*, *j*<*m*<128. All pedometricians from *j*+1 to *m*-1 can now sit in their proper chairs. We now have a new sequence in which pedometrician *m* acts as if he were a new drunk pedometrician. He will sit with probability p_m in seat *j* (in which case all remaining pedometricians, including 128, will sit in their proper chairs), with the same probability he will sit in 128 which also determines the outcome, or, with probability 1-2p_m the outcome remains undetermined for the present and he sits in the chair of some pedometrician labelled *m*+1 or *m*+2 or…or 127. And so on.

We can have a very short sequence of 'pedometricians who select an arbitrary seat', with just one member (Alf) if he happens to sit in his own chair, or in chair 128. We assume only (see above) that Alf is not 128, (since if he were then he does sit in an arbitrary seat). Alternatively we might have a longer sequence. If pedometrician 127 is in the sequence then his arbitrary chair will either be Alf's chair, or 128 i.e. there is no possibility that his action won't determine the final outcome, but the two possible outcomes have equal probability. Thus, when the exercise begins any pedometrician from *j* (Alf) to 127 might be the one whose action determines the outcome, we don't need to work out their respective probabilities because in any case the probability that seat 128 will be left free or will be occupied is the same. And therefore the probability that Pedometrician 128 will sit in his own chair is 0.5. Note that the actual number of pedometricians is immaterial, and that the statement that Alf is one of the first 64 pedometricians is something of a red herring since *any* condition such that he is not Pedometrician 128 would do.

Pedomathemagica Answers

Answers to Pedomathemagica No. 26

by Gerard Heuvelink

Problem 1 (MEDIUM-HARD)

See Pedometron 21

Problem 2 (MEDIUM-HARD)

This problem was perhaps more difficult than I had announced (instead of 'medium-hard' I would now classify it as 'hard').

Let the three people at the crossing be named A, N and S, where A is the person that always tells the truth, N the one that never tells the truth and S the one that sometimes tells the truth. Of course we do not know who is A, N and S. Let us also name the same three people 1, 2 and 3, in arbitrary order.

We use the first question to identify a person other than S. We ask nr 1 the following question: "Does nr 2 speak the truth more often than nr 3?"

If the answer is yes, then there are three possibilities: 1) if 1=A then 2=S and 3=N; 2) if 1=N then 2=S and 3=A; 3) if 1=S then either 2=A and 3=N or 2=N and 3=A. Note that in all three cases we have 3≠S.

If the answer is no, then there are also three possibilities: 1) if 1=A then 2=N and 3=S; 2) if 1=N, then 2=A and 3=S; 3) if 1=S then either 2=N and 3=A or 2=A and 3=N. Now, in all three cases we have 2≠S.

If the answer to the first question was yes then we pose the second question to nr 3, if it was no we pose it to nr 2. This ensures that this person is either A or N, but not S.

The second question is "Would the person, who always says 'yes' when you say 'no' and vice versa, say that this (point in one of the two directions) is the road that leads to town A?". If the answer is no, then take the road that you pointed at. If the answer is yes, take the other road. It is not difficult to verify that this road leads to town A both when the person you posed the second question to is A and when it is N.

Pedomathemagica

Pedometric Bookworm by Dick Webster

The Encyclopedia of Soils in the Environment, compiled by Daniel Hillel and published by Elsevier in 2005, is packed with information. It contains all you need to know about soil and a lot more that you might never need. It comes in four hefty volumes, and you would not want all of them in your rucksack when you venture out into the environment on the off-chance that you might wish to consult them on a matter or two.

In the library, however, a bookworm decided he would like a taste of all that in-formation. So he climbed his way up to the bookshelf where the four volumes were arranged in their usual order, 1 to 4 from left to right. He wriggled his way to the first page of Volume 1. Then he steadily munched his way straight from there to the last page of Volume 4. Volumes 1 and 2 contain 548 and 542 pages respectively, each spanning 28.5 mm. Volume 3 contains 570 pages spanning 30 mm, and Volume 4 is the slimmest with only 459 pages spanning 25 mm. The encyclopedia is handsomely bound in hard covers 3.5 mm thick. How far did the bookworm travel on his grand pedological tour?

Pedomathemagica by Gerard Heuvelink

Problem 1 (EASY)

Christmas time is upon us, which means that many of us will be celebrating with lots of good food and drinks. In case you are a beer drinker, you may like this problem which is not difficult to solve. You should be able to solve it even after ten glasses of beer (try it!).

The mass of one beer glass and one beer bottle equals the mass of one beer mug. The mass of one beer bottle equals that of one beer glass and one beermat, and the mass of two beer mugs equals that of three beermats. How many beer glasses equal the weight of one beer bottle?

Problem 2 (HARD)

Once you are sober again and have a fresh mind, you may try this one, which is pretty hard.

In front of you lie five rocks that are similar in size but have quite different weights. The weights are 1, 2, 3, 4 and 5 kg.

Can you, without any other aid than five times using the scales (see picture) identify the weight of each of the five rocks? Note that the scales only tell you which side is heavier (or that both sides have equal weight) but not how much heavier one side is compared to the other. You are allowed to put more than one rock at each side of the scales.