Issue 38, December 2015

Chair: Budiman Minasny

Vice Chair: Lin Yang

Editor: Jing Liu

ΠΕΔΟΜΕΤRON

Inside this issue

From the Chair

Dear Colleagues

It has been a great year for Soil Science, we were celebrating the International Year of the Soils. We have a great celebration in Cordoba, Spain. And I would like to thank the organiser, especially Tom Vanwalleghem for their hospitality and hard work.

An important event happened few weeks ago, COP21. This is the first agreement requiring all nations to restrict global warming to well below 2 degree C above pre-industrial levels. And I was wondering what contributions can pedometricians make to help achieve this target. Fortunately the French has started the 4 per 1000 initiative.

The French minister of agriculture launched the "4 pour mille" initiative to boost the carbon stock in the world's soils by 0.4% each year. We can help to put this into practice. Our digital soil mapping skills can readily translate the digital maps into required sequestration rates to achieve the 0.4%. We can also identify areas of potential sequestration, carbon saturation and much more. An as Nicolas Saby explained (on page 26) we can verify the carbon stocks change due to changes in land use or management practice. So there is a lot of potential work for us to make the 4 per 1000 initiative work.

Please enjoy this issue of Pedometron, it has information on our activities and initiatives, including our new award for early-career researcher.

It is the time of the year again, students and colleagues have gone for holidays and offices are quiet. So looking forward to 2016, International Year of Pulses.

Wish you a productive and successful new year.

Budiman Minasny, Sydney, December 2015.

Brendan and I sitting alone in a classroom anticipating 2016.

Pedometrics 2015, Cordoba

The 2015 Pedometrics meeting in Córdoba Spain has come and gone. To call the meeting a great success would be an understatement. Tom Vanwallegham and his colleagues at the University of Córdoba arranged a stellar meeting. They compiled a program that not only helped to facilitate a week of fantastic pedometrics discussion but provided meeting attendees a taste of the Spanish and Andalucían way of life.

The meeting began with an informal icebreaker Sunday night. Foreshadowing the many stellar meals to come, the evening consisted of a seemingly endless stream of pinchos and tapas. Lubricated by generous servings of Jamón Ibérico, conference attendees settled easily into the Spanish way of life and readied themselves for the start of the conference.

The scientific program started in earnest Monday morning with a preconference workshop on soil landscape modeling led by Peter Finke and the IUSS working group on soil-landscape modeling. In the morning session, researchers presented their latest work in this exciting arena and in the afternoon walked attendees through first-hand demonstrations of their models. As a complete novice to the soil-landscape modeling realm, I must say I was quite impressed by the sophistication and depth of the current efforts in the field.

The workshop on Monday proved to be the perfect start for the conference. By providing a glimpse of what the future of pedology and pedometrics may hold, the presenters and organizers left me wondering about what is to come for our science. What are current and future limits of pedometrics as a tools set? What can pedometricians contribute to other scientific disciplines? What will pedometrics look like in 10 years' time? In addition to many fascinating pedometrics presentations, several speakers gave more introspective talks; addressing these questions directly.

Several speakers asked us to think about the role of pedometrics in the larger scientific community. Jed Kaplan asked pedometricians to provide the kind of data needed for land-surface modelers. Cristine Morgan asked us to think about what questions pedometrics could answer for those studying soil security. Philippe Baveye asked us think critically about the utility of our work and keep the larger scientific questions in mind when approaching our research. Such thought provoking presentations provided an ideal backdrop for three days and nights of pedometrics discussions.

A major highlight of the conference was the conference dinner on Wednesday night. After a remarkable meal and some lively conversation, Budiman Minasny took center stage as he presented the 2014 Richard Webster medal to Gerard Heuvelink. Gerard's acceptance speech was not the only

Pedometrics 2015

entertainment for the evening as we were all treated to a breathtaking performance of the Flamenco, a truly mesmerizing performance.

Pedometrics 2015 was one of the most rewarding conferences I have ever attended. The small number of participants lent the conference an intimate feeling. Pedometrics is a great community full of passionate and skilled scientists. I am glad to call myself a pedometrician and I look forward to see you all at Pedometrics 2017 in Wageningen.

Jason Ackerson,

Texas A&M University, USA.

Pedometrics 2015 was an opportunity for soil scientists from all over the world to meet and discuss their science. I was lucky enough to be amongst them thanks to the Pedometrics Student award. This was the first time I had attended the Pedometrics coference and I learned a great deal from the experience, not only from the wide range of presentations and posters but also through discussions with others at the meeting.

I had the opportunity to present my work on “The use of an unbalanced nested sampling scheme to reveal scale-dependent variation in soil properties” and found the discussion that followed highly stimulating. I received lots of positive feedback and a few suggestions on the directions my work could take in the future.

The program was varied and covered a wide range of topics under the banner of pedometrics. However, across the sessions a common theme emerged. Many of the presentations drew on the idea that whilst pedometrics is an exciting and developing field, if we are to make an impact we need to work closely with scientists from other disciplines.

Helen Metcalfe,

Rothamsted Research, UK

Photos courtesy of John Triantafilis

Pedometrics 2015

You can view some of the highlights on twitter #pedometrics15, twitter.com/hashtag/pedometrics15?f=tweets . At the conference, we presented the following awards, congratulations:

Best Student Oral Presentation
Jason Ackerson- Continuous depth profiles of soil clay content from penetrometer-based in-situ visible near infrared spectroscopy
Mario Fajardo- Detection of soil morphological horizons using Vis-NIR spectroscopy

Best Oral Presentation
Jacqueline Hannam - Can soilscape stratification optimise the feature space in digital soil mapping?

Best Student Poster
Mouna Feki - Comparative assessment of different methods for determining soil hydraulic properties: measurements and estimations

Best Poster
Sebastien Drufin - Towards a methodology for harmonization of soil maps assisted by digital mapping

Margaret Oliver Award for Early-career Pedometricians

The Pedometrics Commission of the IUSS is pleased to introduce a new award, which is intended to recognize up-and-coming talent in pedometrics. The award is named for Margaret Oliver, in recognition of her outstanding commitment to the promotion and encouragement of pedometricians in the early stages of their careers as well as her overall service to pedometrics.

Margaret Oliver is well-known in the pedometrics community for her many papers, book chapters, standard textbooks, and most recently Basic Steps in Geostatistics: The Variogram and Kriging (M. A. Oliver and R. Webster, 2015, Springer). She has also contributed to the use of numerical methods in many other applications such as childhood cancer, pollen and radon. She made a major contribution to our subject by popularizing it with clear descriptions of its method and applications; for example the early paper in Soil Use & Management (7:206) "How geostatistics can help you", aimed at non-pedometricians working in applied soil science. She has been a mentor to a number of doctoral students, including some of the first female pedometricians. She taught geostatistics and other numerical methods to undergraduates and postgraduate students at the Universities of Birmingham and Reading.

The award will be given at each biennial international meeting of the Pedometrics Commission; the first award will be given in Wageningen (NL) in 2017. The recipient will also receive a voucher for £50 from Wiley, which can be redeemed for digital books.

Requirements and eligibility for the award of the Margaret Oliver award:

Nominees must have:

(1) received a PhD degree or equivalent no longer than 5 years before the nomination deadline;

(2) made high-quality contributions to pedometrics, as evidenced by published work, conference presentations, workshops, field guides, etc.

(3) at the time of the award be active in pedometrics and with a prospect of so continuing

"Pedometrics" is broadly defined as the application of mathematics or statistics in soil science.

Nomination will be announced in mid-2016.

New science for an old art

Alfred Hartemink (hartemink@wisc.edu)
University of Wisconsin-Madison
FD Hole Soils lab

“The soil is a complex mixture of many minerals, decaying organic matter, bacteria, fungi and other living organisms. All of these constituents play an important part in the fertility of a soil” (Truog, 1930). A definition that closely resembles that of Carl Sprengel of 1844: ““Soil is a changed mass of material derived from minerals and containing the decomposition products of plants and animals”(Sprengel, 1844). Emil Truog was a soil fertility professor at The University of Wisconsin, and he was particularly interested in the mineral matter of the soil (“the foundation material”). He realized that part of the mineral matter can be identified by petrographic microscope, but that the colloidal fraction is too small and that the x-ray may be used to investigate what mineral is present (Truog, 1930). The x-ray of a soil sample creates a diffraction pattern from the constructive interference of scattered x-rays passing through a crystal. In the late 1920s, Truog used the x-ray diffractometer of the Department of Geology and started investigating what phosphates are obtained when phosphate fertilizers are applied to soil. This was needed as there was a poor response to phosphate fertilizers on some soils. In those soils, the phosphate combined with hydrated iron oxide and formed a basic phosphate which was not well available to crops. They also used x-rays for assessing the nature of inorganic substances that caused soil acidity (Fig. 1).

The optimism of their research was large, and so they wrote: ”The X-ray now used to unlock the secrets of the soil”. It was in the year 1930.

The x-ray machine was invented by the German WR Röntgen in 1895, and he titled his discovery x-radiation, where x stood for the unknown. He won the Nobel Prize in 1901. There were rapid developments when connection between x-rays radiating from a sample and the atomic weight of the sample was discovered (Shackley, 2011). Half of the Nobel prizes in physics were awarded to developments in x-rays from 1914-1924. The first portable x-ray machines were from the 1910s and used by Marie Curie in World War I ambulances. In soil science, x-rays were mostly being used to investigate clay minerals including aluminosilicates, ferric and ferrous silicates and other minerals (Kubiëna, 1938) .

One important development was that fluorescent emission of secondary x-rays in the sample could be induced by using an x-ray source with a metal target. Simply said, it is the determination of elemental composition by measurement of the intensity of secondary or fluorescent x-rays emitted from a sample bombarded with high energy x-rays from an x-ray tube (Vandenheuvel, 1965). It had some disadvantages, but overall the x-ray-fluorescence spectrometer (xrf) is non-destructive, requires minimal sample preparation, and is relatively fast and easy to use (Shackley, 2011). The xrf can be used for chemical analysis of major, minor and trace elements in rocks, soils and sediments. In the 1950s, the xrf

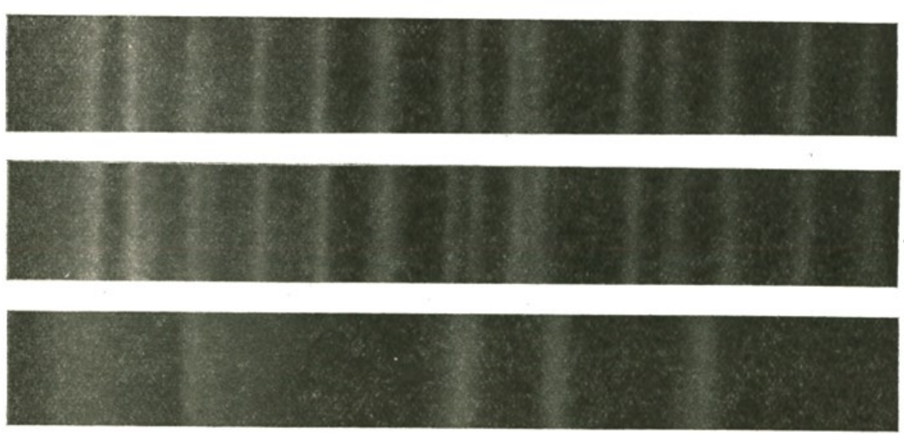


Fig.1. Emil Truog (1884-1969) probably best known for the Hellige-Truog pH Test, used the x-ray to assess the form of phosphate after fertilizers have been applied. The upper x-ray diffraction pattern shows a basic iron phosphate; the middle picture is x-ray pattern of material formed when hydrated iron oxide is treated with soluble phosphate. The two patterns match. The lower picture is the x-ray pattern of alumino-silicic acid causing soil acidity, having other lines that are differently spaced.

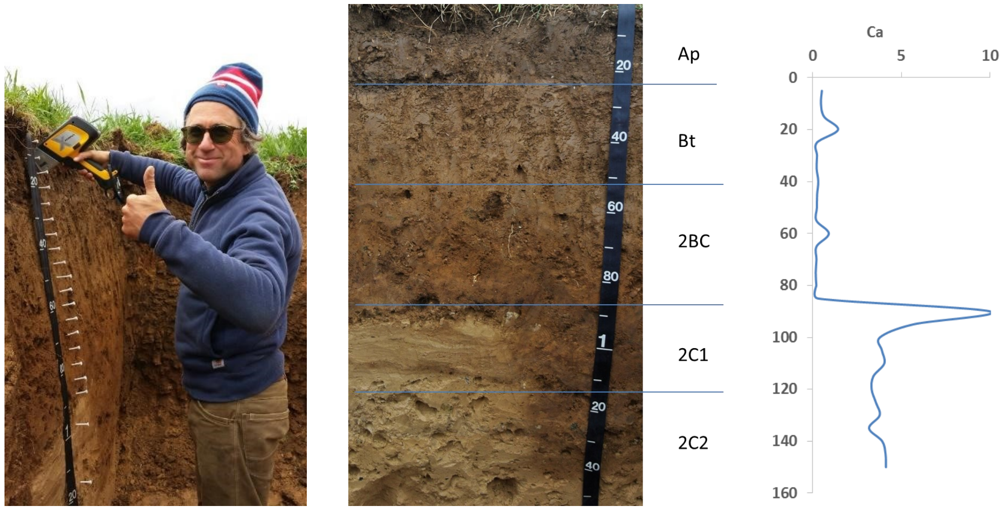


Figure 2 In situ measurements of elements using pXRF and rapid assessment of parent material boundaries – loess over glacial outwash in an Alfisol in Wisconsin. Indispensable: thumb up, and wear sunglasses.

became commercially available. There must have been papers on xrf and soils before the 1950s, but ploughing through the database it seems that the first papers started to appear in the 1950s and 1960s (e.g. Handy and Rosauer, 1959; Van Compernolle et al., 1965). At the International Congress of Soil Science in Bucharest in 1964, x-ray fluorescence spectroscopy was presented as a new method for investigation soils (Fripiat, 1967).

The use of x-rays for the studies of soils has found wide applications, not just for elemental or mineralogical analysis, but also in x-ray computed tomography and x-ray microscopy. Handheld or portable x-rays fluorescence (pXRF) as we know them now became available in the 1990s, and currently the technology is being used for rapid assessment of elements in the field (Fig. 2).

Progress in soil science has come from sound theory guiding research, and from measurement techniques and data leading to the development of sound theory. Readers from this newsletter will probably would like to add to that: quantitative thinking and pedometry. Some of that thinking is currently being steered to update and modernize soil profile descriptions through (i) in situ measurements using a range of machines (e.g. pXRF, vis-NIR), (ii) development of soil depth functions for a range properties reflecting horizons as well as key soil processes, and (iii) to map the soil profile through fine raster sampling as well as digital image analysis. These techniques, collectively termed *digital soil morphometrics*, work at the pedon scale and have the potential to enhance our understanding of soils – how they form, how horizons could be delineated, and for soil classification purposes.

This little piece was prompted when reading the short note of Truog in a report from 1930. Naturally, I had to think of the "The sun has shone here antecedently" (McBratney and Minasny, 2010). It should be noted that the 1930 report also contains articles entitled "Making soil microbes work for man" and "Insects – both friend and foe to the farmer". Sounds familiar perhaps.

References

Fripiat, J. J. (1967). New methods of investigation in soil chemistry. In "8th International Congress of Soil Science", Vol. 1, pp. 171-192. ISSS, Bucharest, Romania.

Handy, R. L., and Rosauer, E. A. (1959). X-ray fluorescence analysis of total iron and manganese in soils Proceedings Iowa Academy of Sciences 66, 237-247.

McBratney, A., and Minasny, B. (2010). The sun has shone here antecedently. In "Progress in Soil Science" (R. Viscarra Rossel, A. B. McBatney and B. Minasny, eds.). Springer, Dordrecht.

Shackley, M. S. (2011). An Introduction to X-Ray Fluorescence (XRF) Analysis in Archaeology. In "X-Ray Fluorescence Spectrometry (XRF) in Geoarchaeology" (M. S. Shackley, ed.), pp. 7-44. Dordrecht, Springer.

Truog, E. (1930). "The X-ray is now used to unlock the secrets of the soils," University of Wisconsin, Madison.

Vandenheuvel, R. C., ed. (1965). "Elemental analysis by X-ray emission spectrography," pp. 1-771-821. ASA, Madison.

The challenge of sampling remote tropical mountain areas

Mareike Ließ (mareike.liess@uni-bayreuth.de)
University of Bayreuth,
Dept. of Geosciences/ Soil Physics Division

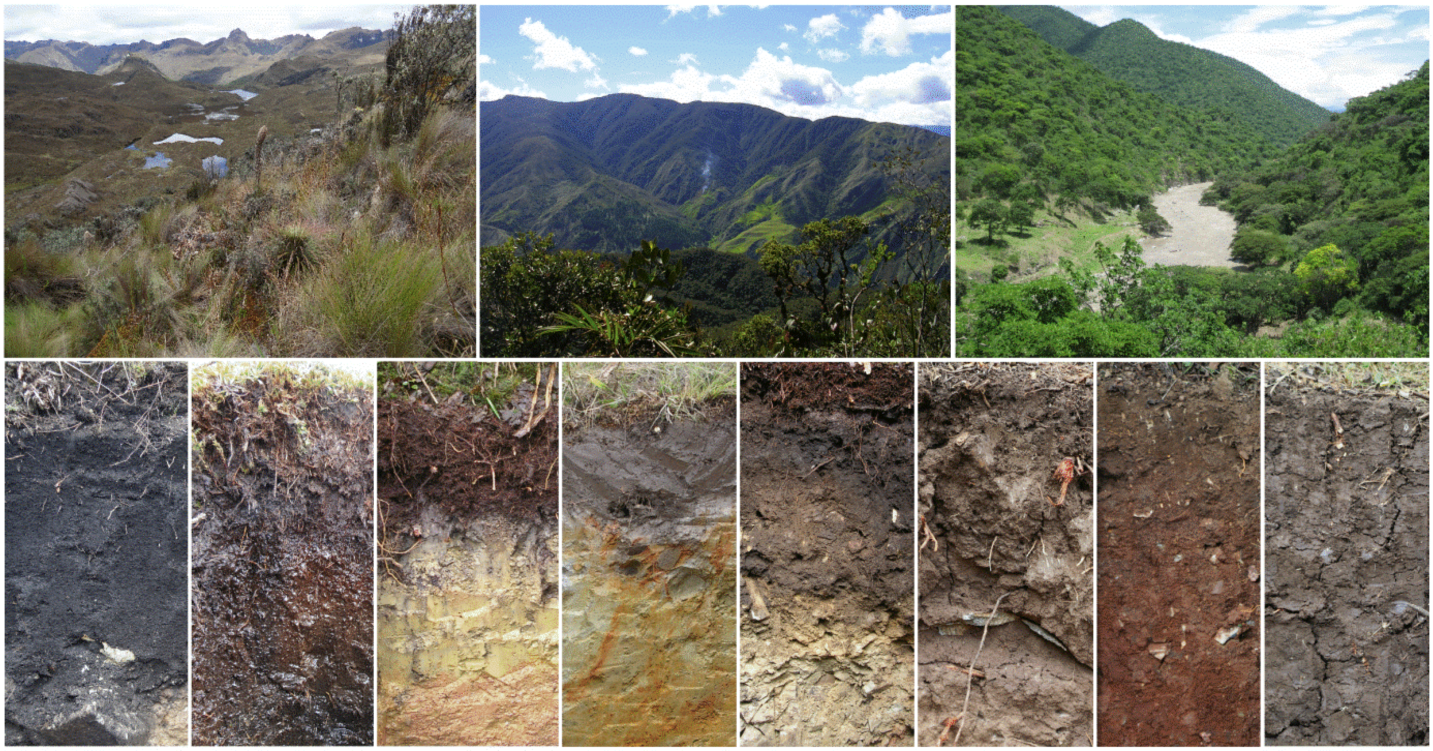


Figure 1 Investigated mountain landscapes and their soils (Photos V. Brito, M. Ließ).

1 Introduction

Due to the strong link between soil science and agriculture, most scientists involved in digital soil mapping focus on agriculturally used areas. But there is also another reason for the high attention these agricultural areas receive. They are easy to access, while the highly diverse remote soil-landscapes of tropical mountain areas provide many challenges for soil sampling.

The soil-landscapes of Ecuador have captured my fascination. Here, areas of very different climate, vegetation and geological origin are found in close neighbourhood. The Andean mountain range subdivides the country into three main zones: Hot humid Amazon lowlands are contrasted by hot dry regions along the western coast. The mountain area includes ecosystems of different rainfall and temperature regimes leading to very different soil conditions. The three areas currently investigated include mountainous areas of different size, altitude, climate, vegetation and last but not least soils. Let's refer to them as humid páramo (Figure 1, upper left), montane cloud forest (Figure 1, upper middle) and dry forest (Figure 1, upper right) region according to their climate induced vegetation. Figure 1 shows some photographs of the three landscapes and their corresponding typical soils.

2 Strategies for sampling remote mountain areas

In sampling for spatial prediction, the above all aim is to ensure that the spatial variability of the soil is well-captured without introducing any bias, while the design remains feasible according to operational constraints such as accessibility, man power and cost. The here presented approach from Ließ (2015) solves the problem of many statistically-based sampling designs which select unreachable points.

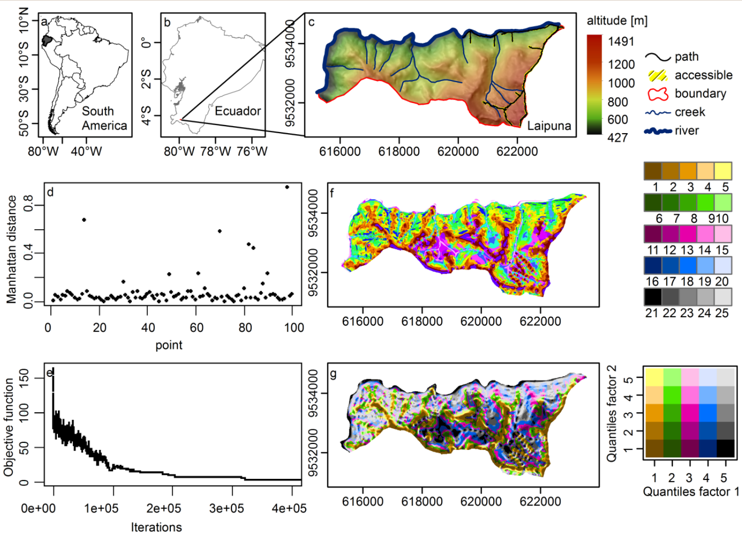


Figure 2: Sampling designs for areas with limited accessibility. a) South America, b) Laipuna reserve within Ecuador, c) Laipuna reserve, d) Manhattan distance, e) Development of the objective function in clhs, f) landscape units from cluster analysis g) landscape units from predictor quantile combinations. Adapted from Ließ (2015).

According to Jenny (1941) it is the interacting effect of the factors of soil formation which forms the soil in a particular landscape position – the basis for any regression-based digital soil mapping approach. As a consequence, we do not necessarily need a good geographical coverage of an area, but of the predictor space, if we want to capture an area's pedodiversity. Any sampling scheme for regression modelling can easily be adapted to a situation where most or all of the area is inaccessible. To demonstrate this, a selection of sampling schemes was adapted to guarantee suitability for any investigation area, while limited accessibility is taken into account. The tropical mountain landscape of the Laipuna reserve with its low accessibility along a very limited footpath network serves as an example (Figure 1, upper right & Figure 2c). The area's variability is captured by a number of parameters representing vegetation and topography.

To select a representative dataset in this area of low accessibility requires some adaptations to the applied sampling design. Here are some useful strategies documented in Ließ (2015):

1) You can replace randomly selected points by accessible points of similar landscape positions, i.e. points which are close in the predictor space but not necessarily in geographic space. The Manhattan distance may be used as a distance measure (Figure 2d) between the randomly selected points and their replacements, or

2) You can adapt the conditioned latin hypercube algorithm (cLhs, Minasny and McBratney, 2006) in such a way that the quantiles for the Latin hypercube are based on the whole research area while selecting points to optimise the hypercube is only permitted from the accessible subarea. Figure 2e shows the development of the objective function while optimising the Latin hypercube.

3) You can create landscape units by different procedures, e.g. by clustering the environmental predictors (Figure 2f) or by making use of the predictors quantiles (Figure 2g) and sample these units according to their spatial coverage within the whole area, while again selecting points is only permitted from the accessible subarea.

4) For some of the designs it is useful or even necessary to reduce the number of predictors. e.g. cLhs will not give satisfying results with a high number of predictors. As a better alternative you can condense the predictor space by applying an a-priori factor analysis. This will leave you with two or three factors instead of hundreds of possibly correlated predictors without having to choose.

Finally, instead of sampling a previously unsampled area with limited accessibility, the here presented strategies may also be used to adapt designs to subsample an existing dataset. So you may use your existing data without starting to sample all over again. But please be aware that even if the here described strategies will assist you to adapt any sampling design for regression-based prediction, some designs might be more suitable than others. For it is the interacting effect of the factors of soil formation which forms the soil in a particular landscape position (Jenny, 1941). As a consequence, this interaction needs to be considered also within the sampling design. Last but not least you can simply test for the sample's representativeness by comparing the predictor space of the selected sample to the predictor space of the whole area.

Figure 3: The adventure of sampling montane tropical cloud forest soils (Photos M. Hitziger, M. Lara, M. Ließ)

3 Experience from sampling a montane cloud forest area

After closing the gap between statistical desires and operational applicability, sampling these remote mountain areas still remains a challenge. Here is some personal experience from sampling the montane cloud forest area (Figure 1, upper middle).

The never drying vegetation gets you soaked within minutes, while hiking the steep and narrow paths. To gain a bigger database with limited resources in terms of man power and time, you will soon consider auger sampling besides profile sampling. However, 2 m augers (necessary due to the thick organic layers) are not necessarily saving time and man power in these soils which are influenced by frequent landslides. Nevertheless, tilting with stones or penetration of some belowground horizontal tree trunk improves your skills in problem solving. And as you may have guessed, accessibility is a challenge. The upper cloud forest area seems impervious and the footpath network only runs along the mountain ridges, but doesn't enter the deep side-valley structures. So you have to climb the dense root network extending way above ground without falling too often into one of the root network's loops - a real challenge while you try to look ahead through the dense vegetation in order to recognise the suddenly occurring landslides before it is too late. The solution to all problems came by opening hillslope transects to cover upper, middle and foot slope positions of different exposition, slope and curvature at different altitudes. This permits walking instead of climbing the length of the side valley slopes, but still needs some awareness in avoiding being too close to the person working with the machete, as even machete handles get slippery when wet. Hence, please be aware: Tropical montane cloud forests are the toughest to sample, but there is a high adventure factor involved as you can see from the photographs displayed in Figure 3.

Acknowledgements

The here presented research is embedded in an interdisciplinary ecosystem research group funded by the German Research Foundation (DFG, http://www.dfg.de/). In 2013 the group changed its organisational structure to become a Platform for Biodiversity and Ecosystem Monitoring and Research in South Ecuador (http://www.tropicalmountainforest.org/). The main interest was deepening previous research while at the same time diving into a stronger cooperation with a multitude of local public stakeholders to include the knowledge transfer aspect and guarantee long-term monitoring. Logistic support was provided by the local stakeholders, Nature and Culture International (http://natureandculture.org/) and the municipal public agency ETAPA (http://www.etapa.net.ec/).

References

Jenny, H., 1941. Factors of Soil Formation.: A System of Quantitative Pedology., New York.

Ließ, M., 2015. Sampling for regression-based digital soil mapping: Closing the gap between statistical desires and operational applicability. Spatial Statistics 13, 106–122.

Minasny, B., McBratney, A.B., 2006. A conditioned Latin hypercube method for sampling in the presence of ancillary information. Computers & Geosciences 32 (9), 1378–1388.

Mareike Ließ is a researcher at the University of Bayreuth, Germany. She will soon join the Helmholtz Centre for Environmental Research (UFZ) in Halle (Saale). Her research interests are soil science, soil-landscape modelling, machine learning & data mining, multivariate statistics, geostatistics, remote sensing and terrain analysis

Turning a smartphone into a tricorder for soil monitoring

Matt Aitkenhead, David Donnelly, Malcolm Coull
The James Hutton Institute,
Aberdeen AB15 8QH, Scotland, UK

The use of spectral information for estimating soil attributes is a large and growing research area. Some of these attributes, such as organic carbon content, pH and moisture content are not only well-estimated but are very important indicators of general soil health and fertility. These attributes and others provide broad information about the capacity of the soil to provide nutrients, water and physical support to crops. They also provide information about erosion and compaction risk.

Smartphones are potentially useful for field-based monitoring of soils as they have (A) cameras and other sensors, (B) processing power, (C) data transmission/reception capacity and (D) the ability to incorporate highly customised software and user-specific software tools (apps) into their operation. However, while their cameras may have multiple mega-pixel imaging capabilities, they cannot provide high spectral resolution.

At the James Hutton Institute in the UK, we have been working to find out whether it is possible to use a smartphone camera to provide estimates of soil attributes, or whether we need more detailed spectral information. It turns out that while some information is lost in using a camera that captures in broad visible wavelength ranges only (RGB), there is another trick that can be applied – photographs taken with a smartphone can be tagged with their capture location, and this location can be used to include information about the soil's environmental characteristics.

To develop the underlying models, we have used a database of Scottish soil samples that allows us to link soil attributes to colour and site descriptions. This database has been developed at the James Hutton Institute over several years, and includes data from the National Soil Inventory of Scotland, a systematic survey of Scottish soils using a grid-based system (10km grid).

A mobile phone app (Android/Apple) has been produced that provides an estimate of soil organic matter rapidly and for free, while a later version is being developed that will provide information on several more soil indicators. This app currently only works for Scottish soils, but we are working on expanding this! We are also working on models that can provide estimates of several other soil attributes, with the aim of providing land managers, consultants and other stakeholders with a tool that will allow them to rapidly assess soil condition.

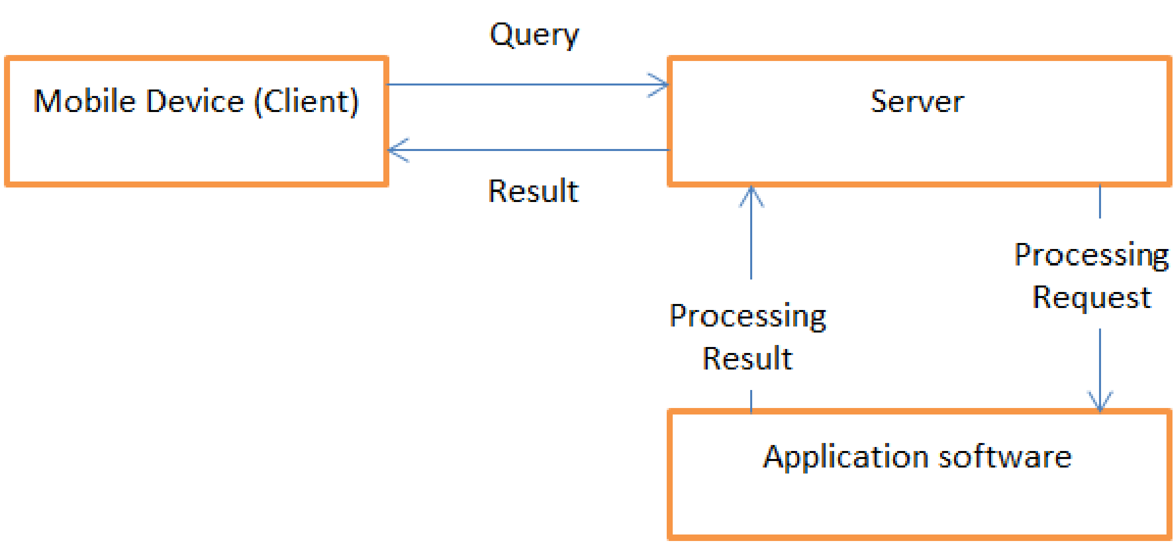
So how does this app work? From the user's perspective it is quite straightforward and simple, but there are a number of things going on behind the scenes. First, you take a photograph of the soil, with the colour correction card in place

This image is geotagged and sent to our server, where it is split into image data and positional information. The image data is corrected using the colour correction card's pixels, and the location information is used to extract site characteristic descriptors from a number of spatial datasets. These descriptors include topography (e.g. elevation, slope, aspect), climate (rainfall, temperature), land cover, soil map unit and other information.

Why correct the image colour? A number of factors can influence the colours in a digital photograph, including lighting conditions and camera design. Colour correction allows us to reduce these external effects and get a more accurate representation of the soil colour. This is important because colour is such a strong indicator of several soil attributes (particularly organic matter content). So we take the colours visible in the correction card, and use this to adjust the image so that the correction card matches what we know it should look like.

What about location? Knowing where the photograph was taken can provide information about the site, as described above. Soil character is influenced by a number of different factors (as you well know!) and if we know these specific factors for a site we can model the soil's attributes. Integrating information about colour and site descriptors provides a more accurate model than just using colour or site descriptors alone, which is why both are included in the underlying model.

There are a number of reasons why the imagery is passed to a server for processing, rather than being handled on the phone itself. Firstly, having all of the processing on the phone would require having all of the spatial datasets there also, and this is impossible for a number of reasons (storage capacity, intellectual property rights, data security). Secondly, the amount of data processing required is fairly high and beyond the capacity of nearly all smartphone pro-



cessors. This has made the data processing framework more complex but in some ways also more robust, as it means the app is smaller (and therefore less likely to crash) and that the data generated can be stored more securely.

The requirements for taking a photograph of soil in the field are not as strict as you might think – all you have to do is make sure that the scene is evenly lit and that the colour correction card is visible in the shot. The photograph can be taken from up to a metre away, and requires very little preparation of the soil (a 30-cm pit will do for the soil organic matter app, which is only providing information on topsoil OM). So it can take less than five minutes from arriving at a site to having your first estimate values – and it's totally free. The app itself takes between 10 and 30 seconds to provide a response, depending on how strong the mobile phone signal is .

This describes what we currently have. The next step is to add more sophisticated image analysis and improved models, to allow us to estimate more than just soil organic matter. Current work is focussing on soil structure and texture, using smartphone image analysis to extract image 'texture'. Currently, the minimum image pixel resolution that can be achieved is around 10 microns with this limit being up to 100 microns for some older models. Try holding your phone's camera as close as possible to a ruler, and you will find out how much detail can be resolved.

And after that? Well, there are a lot of soil attributes that can be estimated using this approach, and a lot of soils around the world that need monitoring. We are hoping to work with other soil scientists internationally to produce a number of smartphone tools for rapid, low-cost soil monitoring. If you are interested, get in touch! Links to the existing apps can be found at http://www.hutton.ac.uk/research/groups/information-and-computational-sciences/esmart, and we are happy to receive feedback and comments.

Matt Aitkenhead is a soil scientist with a background in a wide range of soil and environmental topics, and in the modelling of complex environmental system. He is currently working on Digital Soil Mapping and the monitoring of soil with smartphone apps.

David Donnelly works on the analysis, creation and management of geographical data at the James Hutton Institute, and the leasing of data products. He works on the development of smartphone apps and linking them to underlying data products.

Malcolm Coull is the assistant Soils Data Manager at the James Hutton Institute, and works on a number of projects that require integrated soils data management and GIS work. He is also currently working on GPS tracking of cattle for monitoring health and behaviour.

Is DSM trying to tell us something?

Jack Hannam
Cranfield University

Best presentation at #Pedometrics15., wow what can I say? Thank you so much for the vote and also a very stimulating conference (well done Tom et al.) that was particularly enhanced by the nightly tapas intake. The accolade should of course be extended to my co-authors Thomas Mayr, Joanna Zawadzka and Ron Corstanje at Cranfield and many of the others in the Irish Soil Information project team at Teagasc in Ireland, including all the surveyors that observed, touched and described the soils. The project was tasked to develop a national scale (1:250,000) soil association map for the Republic of Ireland using legacy data available for about 40 % of the country and digital mapping to fill in the gaps (the white areas in Figure 1). When the project started back in 2008 the GlobalSoilMap specifications were not developed. We decided to stratify the landscape into broad soilscapes and develop soil-landscape specific models to predict the soil associations in the gaps. Validation was achieved by independent data gathered through a parallel 3 year field programme (Simo et al., 2015). The project produced a web-based national soil information system for Ireland (http://gis.teagasc.ie/isis/). So looks good, nice map, nice validation, job done.

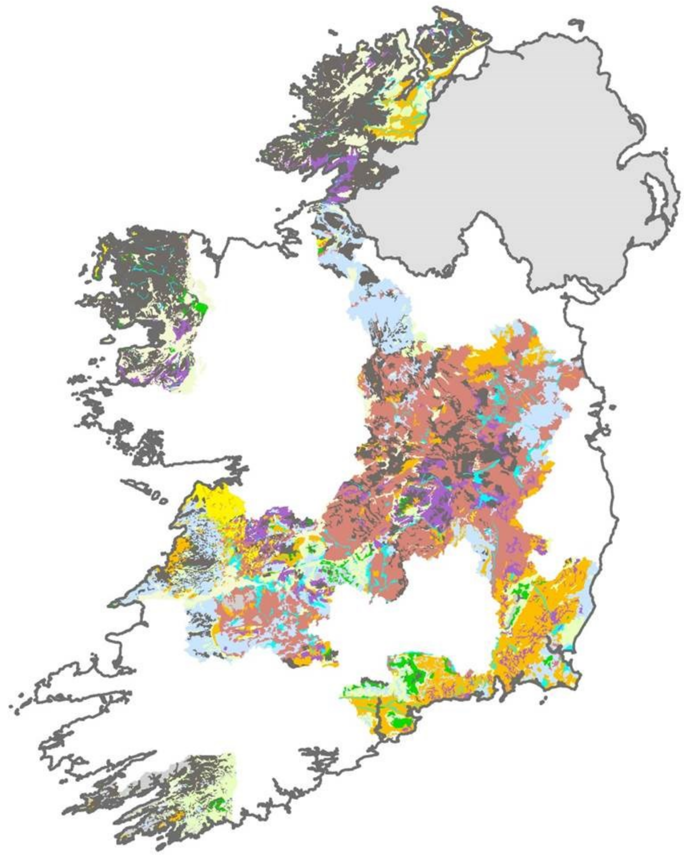


Figure 1. The task of the Irish Soil Information System. Coloured areas: Legacy soil association data. White areas: Soil associations to be predicted using DSM.

But was stratification the most suitable approach? Once we started to look into this other questions arose - is the success or failure of the predictions also telling us something else? Could it give us insight into soil processes at landscape scales? Back to the stratification question. We used a Bayesian network to extrapolate the soilscapes into the 'gaps' using 31 covariates, with the most likely soilscape assigned for the area. We also used a Bayesian network to subsequently predict the soil associations. Certainly, in some areas the soilscape was not well represented in the predicted area, due to suitability of the training data. The legacy data was collected for much of the productive agricultural land in Ireland so the semi-natural landscapes in our gaps were not always well represented. However, if we consider areas that are almost surrounded by legacy data we hope that this effect should be minimised assuming we have repeatable broad soil landscapes over this test area. The predicted soilscapes for this area looked reasonable (Figure 2) even to a sceptical pedologist like myself.

We wanted to test whether our decision to stratify was justified so we also applied a global model in this area in addition to the models stratified by soilscape. When validating the predicted soil associations using the global model vs the stratified model we found that in some cases the predictions were improved using the stratification (as we had originally anticipated) but in other areas a global model was better. In part this is due to the suitability of the training dataset or the covariates. But this was not always the case. In Figure 3 the large area (1cn) that does better using the stratified approach has both good and bad representations of the predicted covariate feature space compared with the training data (red and blue areas). The limitations in the covariate feature space does not explain the whole story. It may go some way to explaining the areas where a global model is favoured (16 in the figure), which have a relatively high number of covariates with values or categories that are absent or fall outside the limits of the training data.

So let's look at the soils. In the area where we have better prediction with stratification (1cn), the predicted soils are

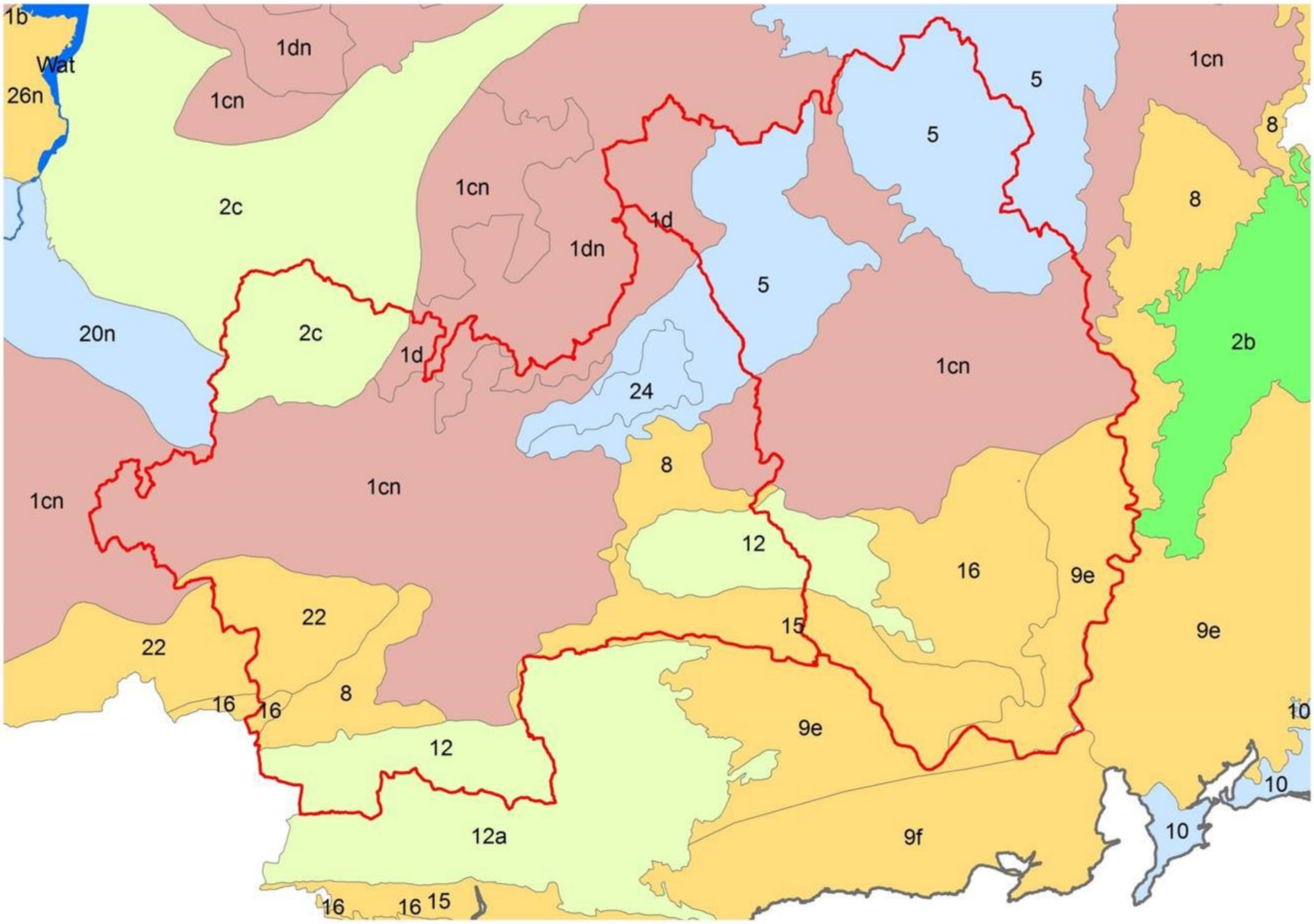


Figure 2 Predicted soilscapes (derived from a Bayesian network) for a test area delineated in red, surrounded by legacy soilscape data (derived from expert knowledge).

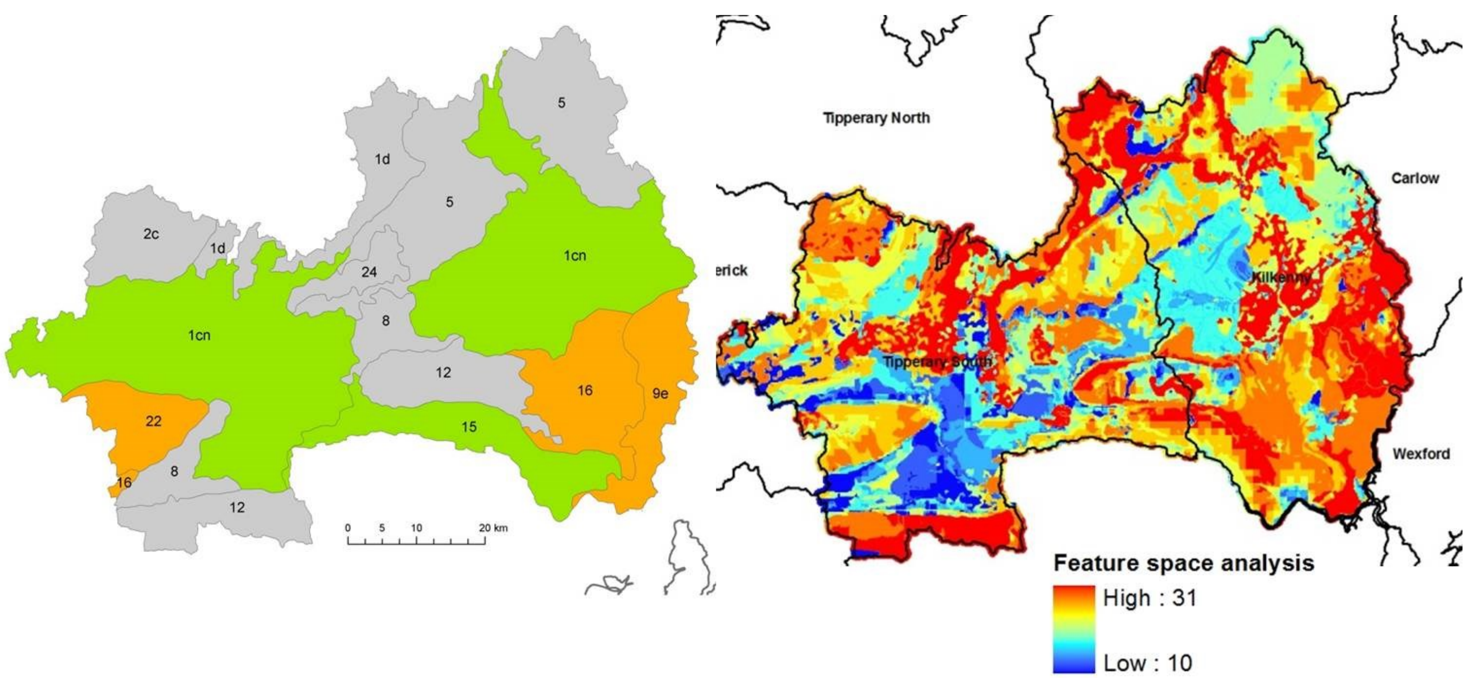


Figure 3 Left: Predicted soilscapes Green: Stratified model is better. Orange: Global model is better Grey: little difference between global and stratified model. Right: Simple feature space analysis of the sum of covariates with bins or categories absent or outside the limits of the training data.

Luvisols, Gleysols and basin peat. The stratification which constrains the feature space may be more appropriate where we have a stronger coupling between soils and landscape features. The Gleysols and basin peats occupy distinct positions the landscape such as depressions or alluvial environments, so are anticipated to have a strong coupling to landscape features or positions. Luvisols, dominated by leaching and the movement of clay within the soil profile, may also show broad-scale links to environmental processes that are captured in the covariates and training data, at least for these environments in Ireland. Indeed clay translocation has been effectively modelled at the pedon and toposequence scale using models such as SoilGen (Finke, 2012) but there are gaps in predicting these processes at the landscape scale (Opolot, et al, 2015). In our case just described, DSM can potentially offer a framework to begin to develop spatially distributed models of certain pedogenic processes.

In the area where a global model is more effective (16) the predicted soils are Podzols, 'Podzolic' soils and Cambisols. The type and scale of covariates in the models are not appropriate to predict the processes that cause divergence in these associated soils and hence from a DSM perspective a broader feature space is necessary. This also indicates that something is missing. In this area the podzolic soils are associated with lowland (< 300m) heath environments that have formed as a response to vegetation change precipitated by forest clearance in the Bronze age, subsequent grazing and more recent grassland improvement. These temporal changes in land management are not currently captured by the covariates. In some cases we need to better represent the time factor, not only in terms of the age of the landscape but also to account for temporal fluctuations in other commonly used covariates such as land use or climate.

Not all the answers are here of course and that is not the intention. Some of it is down to data, statistics and possibly sheer luck. But if we can begin to unpick our DSM outputs then perhaps they can tell us a lot more than just how successful our map predictions are. I challenge you. Move out of the comfort zone and take a look at your data though different eyes – particularly when it doesn't work very well!

References

Finke P.A. 2012. Modeling the genesis of luvisols as a function of topographic position in loess parent material. Quaternary International, 265, 3-17.Opolot E., Yu Y.Y., Finke P.A. 2015. Modeling soil genesis at pedon and landscape scales: Achievements and problems. Quaternary International, 376, 34-46

Simo I., Schulte R.P.O., Corstanje R., Hannam J.A., Creamer R.E. 2015. Validating digital soil maps using soil taxonomic distance: A case study of Ireland. Geoderma Regional, 5, 188-197

Jack Hannam is a Senior Research Fellow in Pedology at Cranfield University and a council member of the British Society of Soil Science. Her research interests include DSM, soil functions, soil health and management, and trying to find common ground with statisticians. She tweets @Dirt_Science

On usability of soil maps (and on global soil data models vs stitching together of individual disparate soil maps)

Response to the Pedometron issue 37 article by Gerard Heuvelink "It's the accuracy, stupid"

By: Tomislav Hengl, ISRIC --- World Soil Information, Wageningen

Summary: In Pedometron issue 37, Gerard Heuvelink has opened a discussion about whether the maps should be judged (only) based on the accuracy with which they represent the real world, and whether it is OK to leave artifacts in soil maps e.g. due to the stitching effect of disparate soil maps produced across administrative borders (see the article: "It's the accuracy, stupid"). Tom Hengl has replied to some of these ideas and tried to justify another point of view: attribute accuracy is not all there is (and please try not to leave ANY artifacts in the maps you produce). To Tom, success of a soil map is a product of: relevance for decision making / spatial planning × level of detail (spatial accuracy) × thematic (attribute) accuracy × completeness × consistency × accessibility × price; hence a single measure alone does not determine its success. The key problem of many continental, regional and global soil mapping projects is probably not the limited accuracy of input data or limited soil knowledge, but the fact that soil mapping projects are significantly underfunded when compared to global land cover mapping, climate change projects, even the gaming industry and space research. There is also too little collaboration across borders and soil maps are still dominantly generated in isolation. Tom suggests that a better quality of especially global soil maps can be achieved by using a hybrid top-down (combined point data) / bottom-up (focused areas) framework and ensemble methods of data merging, so that a globally consistent and complete product of maximum possible accuracy can be generated.

My colleague (and friend) Gerard B.M. Heuvelink recently wrote an interesting article in Pedometron with an intriguing title[1]. Although I completely 100% agree that pretty maps are not necessarily good maps and that this is a delusion that everyone should try to avoid (i.e. do not worry too much about the color legend, worry about the accuracy of the values you produce), I felt like responding to this article for two reasons: (1) Gerard has taken my 'Frankenstein soil maps' somewhat out of context, so I would like to provide a better introduction to the origin and context of the characterization of maps as “Frankenstein” maps, and (2) Gerard's suggestion that "maps should be judged (only) based on the accuracy with which they represent the real world" is only partially correct; his suggestion that we should not worry and leave some artifacts in the maps could encourage some people maybe too much. I felt like clarifying some important points (especially to the statisticians).

On the issue #1, I am not sure if I was the first person to coin the term "Frankenstein soil maps" but yes in 2010 at the DSM conference in Rome I presented some slides that illustrated obvious artifacts (country borders) in the Harmonized World Soil DB and in the European Soil Atlas. I then referenced the GlobalSoilMap project, that had just promised a series of new 100 m resolution soil property maps of the world by 2012, by saying "Please do not make another Frankenstein". In subsequent years, my name has been closely associated with that term (in good and bad and until death do us apart...) and OK I accept this. But what did I actually mean by "Frankenstein soil map"? Well, let me first try to define it:

"A Frankenstein soil map is a map that has been produced by combining together disparate soil polygon maps produced independently (in isolation) without harmonization of soil boundaries, scales and/or legends i.e. attributes. Such a product of binding ('stitching') diverse, dissimilar maps will exhibit many obvious artifacts so that end users may lose confidence in the product, although the overall quality

of the product might be sufficient for decision making or modeling."

So why does the term “Frankenstein soil map” have such a negative connotation? I first discovered in 2003 at the Geocomputation conference in Southampton that we (soil mappers) actually do not have the best reputation outside our own discipline when one keynote presenter spoke about the input data used to build hydrological models and said: *"We used geological maps instead of soil maps because soil maps of no use anyway."* I have heard such statements repeated on multiple occasions since then. Rossiter (2006) also recognized that once soil maps are overlaid with other GIS layers the users' perception of the map’s reliability may drop. So yes, we do not have the best reputation in other environmental, hydrology and geology-related fields, and one reason is that our maps were not the best.

However, the core reason for my "Please do not make another Frankenstein" plea actually arose from my reluctance to support building of spatial predictions models separately and in isolation (and then stitching maps together after predictions). I felt that this was highly likely to be suboptimal and have tried to inspire everyone (and still hopelessly trying!) to look instead at bringing all the world soil (point) data together to fit global statistical models (to complement local models). Here, I am mainly inspired by similar work in climatology (Hijmans et al. 2005; Smith et al. 2007), forest monitoring (Hansen et al 2013) and/or species distribution modelling (Araújo and New, 2007; Tyberghein et al. 2012). Global models mean Big Data and heavy (high performance) computing, so it is not trivial, but I think that it is a path worth the effort, and in the future computers should be able to easily handle this. This is really the essence of my fear of the soil *Frankenstein* → it's basically a fear of isolated, individual and non-coordinated soil mapping models (vs a model based on collaboration and data sharing). It is more than just a fear of incongruous visible country or administrative borders in soil maps. France and England dug the Eurotunnel independently and came to the same spot plus or minus few millimeters (note however that this was on the basis of a single engineering design) → OK it is possible but why take the risk?

On the second issue ("maps should be judged based on the accuracy with which they represent the real world"), I think that the issue of data quality deserves a more comprehensive consideration. As many of you are aware of, data quality is overall measure of degree of excellence of a data product (see e.g. Veregin, 1998). Different aspects of data quality are:

- **Accuracy** of data - the degree to which data correctly reflects the real world object or an event being described. Because maps answer two basic questions "what is here" and "where is what" (Burrough and McDonnald, 1998), there are also two types of accuracies: (1) thematic or attribute accuracy or 'the accuracy of how much' and (2) locational or positional accuracy or the 'accuracy of where'. In recent years there is also an increasing interest in communicating uncertainty of the maps (and our estimate of uncertainty also has accuracy), hence there is also the third type of data accuracy of 'how accurate is the accuracy estimate' (sounds maybe a bit silly but it can be tested and will probably be increasingly reported).

- **Completeness** of data - the extent to which the expected attributes of data are provided. Data can be accurate but incomplete and vice versa.

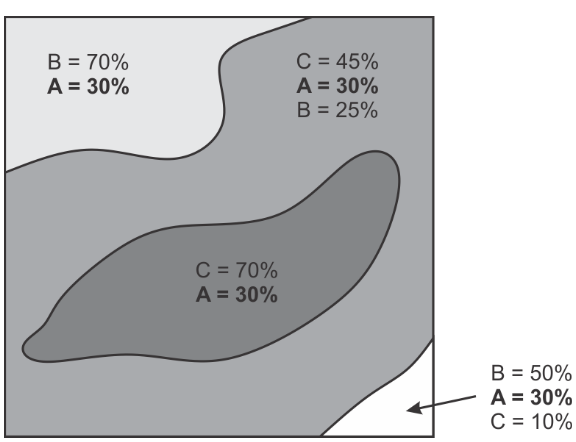
- **Consistency** of data - refers to the absence of apparent contradictions in a database. Data can be accurate and complete but inconsistent. Many models that make use of soils data do not do well if presented with spatial data that clearly represent different populations with different accuracies, resolutions or reliabilities.

- **Reliability** - the ability of a system or component to perform its required functions under stated conditions. Reliability is often the overall product of quality of all inputs, but it can also be an effect of organizational set-up. For example, a data set produced following some international standard that is described in detail and the description is accessible internationally can be more reliable than a data produced using some local standard which is undocumented.

- **Accessibility** - the ease of access and can be measured as amount of work required to access some information of interest. A data set can be of excellent quality but difficult to access even for those who are ready to pay to access the data.
- **Easy-of-use or usability** - defines how well product matches user needs and requirements. A data set can be of excellent technical quality but have very narrow user community.
- ...

Although accuracy is often considered by many as the key measure of data quality it is not the only measure. As mentioned above, data can be accurate but still of low quality and/or usability. It get's even more complex, different aspects of data quality work against each other because quality of course costs money → our budget is limited and we often need to sacrifice e.g. completeness at the cost of accuracy or vice versa. It's like designing a new mobile phone → everybody would like a larger screen and thinner device, but this comes at the cost of shorter battery life and higher price… and less time spent in real life. You can't have it all!

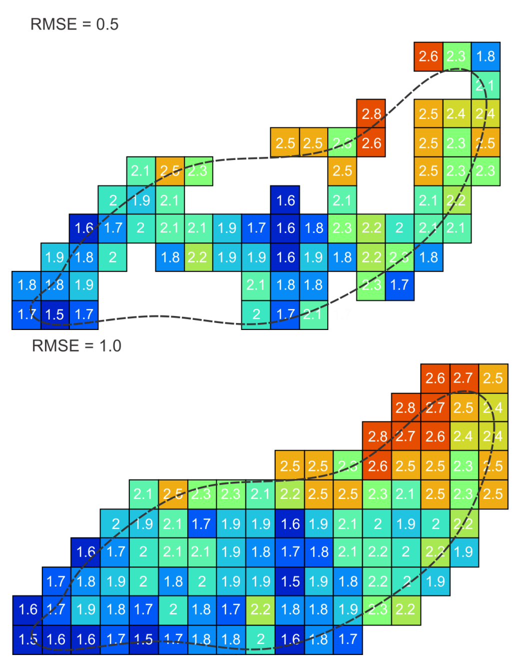
I will now illustrate that perfectly accurate map might be completely unusable with some simple examples. Consider the following example of a soil polygon map with 4 mapping units and 3 soil types:



Ignoring the accuracy of spatial boundaries i.e. let us assume that the boundaries have been delineated with high spatial accuracy, and let us assume that the (attribute) accuracy of setting the percentages is close to perfect. Imagine now that user requires an answer on "where does soil A occur in this study area?". Obviously our uncertainty about "where" i.e. location would have been the major frustration (basically, based on this map we have no idea where does soil A actually occurs, although the 30% fraction estimate might be perfectly accurate!). This hopefully demonstrates that (attribute) accuracy can be completely irrelevant and the usability of this soil map within the context of where is hence close to 0. So yes Gerard, (attribute) accuracy can also be completely irrelevant to the user, depending on the type of question they pose. This brings me to my thesis number #1:

> *Spatial data usability can be considered only within some given data processing workflow, the same way, for example, land usability can be considered only within some given land use system. To achieve maximum usability of the whole spatial Soil Information System, one should consider all aspects of data quality and then try to select an optimal combination (that is still within the project budget)."*

Consider also the following example of two soil maps showing distribution of soil organic carbon stock for a polygon delineated below:



Let us assume that the map on the top has two times higher attribute accuracy than the map at the bottom. In this case user might be interested in finding out "*What is the total soil organic stock for the whole area?*". Even though the predictions in the map on the top might be highly accurate, obviously this map would be completely unusable to the user because it is: significantly incomplete. To many, the map on the top would still be of much better quality than the map at the bottom.

I could continue listing such examples, but I hope I have proven my point: (attribute) accuracy is just one aspect of the quality of soil maps. Accuracy is not all there is. In fact, in their review of Global and Regional Soil Information, Omuto et al. (2012) have also concluded that poor accessibility and lack of harmonization are possibly bigger problems for all future global soil information initiatives than soil data availability and accuracy. This brings me to my thesis #2:

"Success of soil maps is typically a product of: relevance for decision making / spatial planning × level of detail (spatial accuracy) × thematic (attribute) accuracy × completeness × consistency × accessibility × price."

Eventually, soil maps should only be judged by the usage / impact they make. The added value of spatial soil information, especially in financial terms, is probably the best argument for investing in soil mapping.

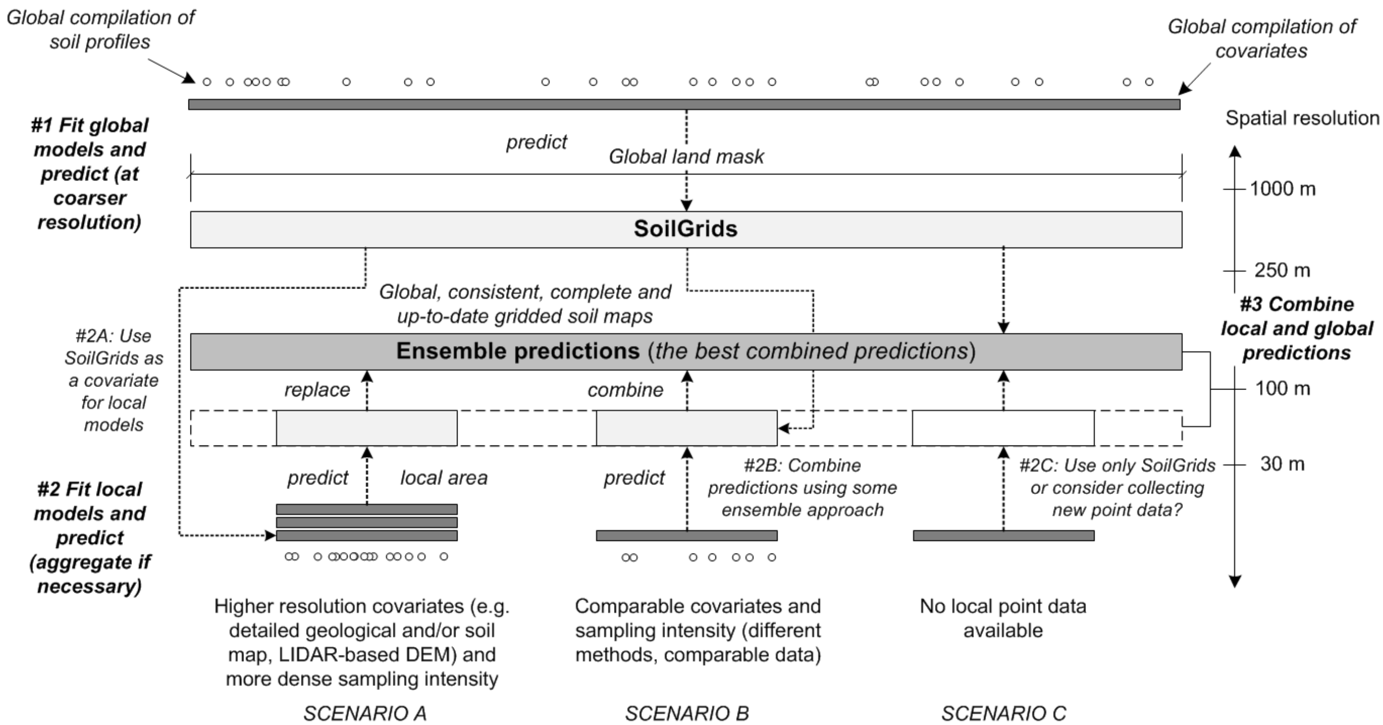
So why have different groups in the world released soil maps with such obvious artifacts (e.g. the HWSD and European Soil Atlas)? My assumption is that it was not intentional at all. These teams worked with modest budgets, and there were simply not enough money to make a more proper job with the maps, or as Pascal put it "If I had more time, I would have written a shorter letter". USDA maybe had just enough patience and resources to produce truly seamless (i.e. complete and spatially consistent and with high spatial detail) digital polygon maps of a large area (STATSGO and SSURGO; Zhong and Xu 2011), but not every soil mapping agency in the world can afford that. For those of you interested in the wider problems of soil map quality, please consider reading for example David Rossiter's "Digital soil resource inventories: status and prospects" article in Soil use and Management.

In financial terms, global and continental soil mapping projects are significantly underfunded when compared to global land cover mapping, climate change projects etc. Even the gaming industry has vastly more money to produce models of Earth relief than soil scientists. Mars has probably more accurate (geological and geomorphological) remote sensing data than Earth (Neukum and Jaumann 2004; Luo and Stepinski, 2009)! So really the key point is here not whether we should reduce all artifacts or not (of course we should try to reduce all artifacts! please do not leave obvious artifacts in soil maps) but: why is there so little investment in making better soil maps of continents / world? Neil McKenzie raised this discussion at the DSM in Rome when he asked "why do governments spend billion dollars on missions to Mars or abstract theoretical physics research, while on the other hand we know so little about the soils of the Earth? ...and soils are one of the fundamentals of the human civilisation and of our survival" (this is maybe not the exact citation, but as far as I remember this was the essence of what Neil said). Global land cover, meteorology, biodiversity, hydrology communities have pushed boundaries of global mapping by producing global data at resolutions of 30 m or better, while the best current global soil data you can find in 2015 is still limited to 1 km resolution (HWSD and SoilGrids data sets). This brings me to my thesis #3:

"We (soil mappers) need to try to make more usable soil GIS, even if there is no major new funding for the Global Soil Partnership and/or GlobalSoilMap, we need to collaborate and help each other to process Big Data and select and apply the best performing models to generate soil spatial predictions that increase user's confidence and result in growing user communities."

So to summarize this discussion, I would like to emphasize one more time three main issues:

- #1 Accuracy of soil maps is important, but consider that (a) usability is a product of multiple quality criteria, and that (b) a map that has perfect (attribute) accuracy might still be completely unusable and/or irrelevant.



Suggested use of SoilGrids for generating ensemble predictions in combination with local models (focused areas). A likely win-win system where local predictions can be combined with global predictions to produce complete and consistent maps (with some minimum artifacts). Based on Hengl et al. (2015).

- #2 An elegant way to reduce artifacts in the output predictions across borders is to run global models (instead of running local models that are geographically isolated and ignore global correlation between various climatic, hydrological, evolutionary, geological and similar processes).
- #3 It is a good practice to correct all obvious (difficult to tolerate by our users) artifacts, inconsistencies and errors as soon as possible. A simple rule of thumb system would be:

  a. First, focus on what user want → consider data quality and mapping costs within some data processing / decision making workflows.

  b. Second, do your homework and clean up the input data (keep the artifacts to a minimum). In the case of the European Soil Allas and HWSD that would mean → harmonize all soil boundaries and legends, spend more time on automating harmonization so that we can run it on other datasets too.

  c. Third, use methods that achieve the best overall combination of attribute and positional accuracies. If ensemble methods help improve accuracy but at the cost of more artifacts → use them nevertheless.

What do you think? Is it perfectly OK to create and use "Frankenstein soil maps" (ones that everyone will immediately notice have been stitched together without full harmonization) or should we attempt to make sure that there are as few artifacts and inconsistencies as possible? Should global soil property maps be produced using globally consistent models or 'stitched' from national / regional predictions… or both? Please write to pedometrics@googlegroups.com or to Gerard and Tom.

[1] It has not been clarified but the quote “It’s the economy, stupid!” comes from James Carville a campaign strategist (some say “evil genius”) of Bill Clinton's successful 1992 presidential campaign against sitting president George H. W. Bush. Carville's original phrase was meant for the internal audience of Clinton's campaign workers as one of the three messages to focus on, the other two messages being "Change vs. more of the same" and "Don't forget health care." source: Wikipedia

References:

Araújo, M. B., & New, M. (2007). Ensemble forecasting of species distributions. Trends in ecology & evolution, 22(1), 42-47.

Burrough, P. A., & McDonnell, R. A. (2011). Principles of geographical information Systems (Vol. 19988). Oxford University Press.

Hansen, M. C., Potapov, P. V., Moore, R., Hancher, M., Turubanova, S. A., Tyukavina, A., ... & Townshend, J. R. G. (2013). High-resolution global maps of 21st-century forest cover change. science, 342(6160), 850-853.

Hengl T, Heuvelink GBM, Kempen B, Leenaars JGB, Walsh MG, Shepherd KD, et al. (2015). Mapping Soil Properties of Africa at 250 m Resolution: Random Forests Significantly Improve Current Predictions. PLoS ONE 10(6): e0125814. doi:10.1371/journal.pone.0125814

Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G., & Jarvis, A. (2005). Very high resolution interpolated climate surfaces for global land areas. International journal of climatology, 25(15), 1965-1978.

Luo, W., & Stepinski, T. F. (2009). Computer□generated global map of valley networks on Mars. Journal of Geophysical Research: Planets (1991–2012), 114(E11).

Neukum, G., & Jaumann, R. (2004, August). HRSC: The high resolution stereo camera of Mars Express. In Mars Express: The Scientific Payload (Vol. 1240, pp. 17-35).

Omuto, C. and Nachtergaele, F. and Vargas Rojas, R. (2012) State of the Art Report on Global and Regional Soil Information: Where are we? Where to go? Global Soil Partnership technical report, FAO Rome.

Rossiter, D. G. (2004). Digital soil resource inventories: status and prospects. Soil use and management, 20(3), 296-301.

Smith, D. M., Cusack, S., Colman, A. W., Folland, C. K., Harris, G. R., & Murphy, J. M. (2007). Improved surface temperature prediction for the coming decade from a global climate model. science, 317(5839), 796-799.

Tyberghein, L., Verbruggen, H., Pauly, K., Troupin, C., Mineur, F., & De Clerck, O. (2012). Bio□ORACLE: a global environmental dataset for marine species distribution modelling. Global Ecology and Biogeography, 21(2), 272-281.

Veregin, H., (1998) Data Quality Measurement and Assessment, NCGIA Core Curriculum in GIScience, http://www.ncgia.ucsb.edu/giscc/units/u100/u100.html, posted March 23, 1998.

Zhong, B., & Xu, Y. J. (2011). Scale effects of geographical soil datasets on soil carbon estimation in Louisiana, USA: a comparison of STATSGO and SSURGO. Pedosphere, 21(4), 491-501.

Pedometricians' Favourite Equations

- **Ana M. Tarquis (Technical University of Madrid, Spain)**

Shannon's entropy is my favourite equation:

$$H = -\sum_{i=1}^{M} P_i \log_2 P_i$$

Always that I look at the Shannon's entropy formula I cannot stop thinking how a simple calculation can tell us so much and being so useful. This **simplicity in form gives us a sense of elegance** if I might say it.

Now that we are in the era of **complex network, big data, data science, mobile phones applications** and I don't know how many other terms the Information Theory comes to me naturally being a fascinating subject. This arose once **the notion of information got precise and quantifiable**.

Shannon's theory tackled the problem of how to transmit information most efficiently through a given channel as well as many other practical issues such as how to make communication more secure. **Shannon's entropy provides an absolute limit on the best possible average length of lossless encoding or compression of an information source.**

Shannon's article titled "A Mathematical Theory of Communication", published in 1948, as well as his book "**The Mathematical Theory of Communication**", co-written with mathematician Warren Weaver and published in 1949, **truly revolutionized science**.

From a physical point of view, information theory has nothing to do with physics. However, the **concept of Shannon entropy shares some intuition with Boltzmann's**, and some of the mathematics developed in information theory turns out to have relevance in statistical mechanics.

Statistical mechanics provides three probability measures on the phase space, the microcanonical, canonical, and grand-canonical measures. When we study the Shannon entropy of these measures it turns that coincides with the **thermodynamic entropy**.

Recently, the Shannon entropy has been widely used as a useful index of various diversities and heterogeneities such as biodiversity and pedodiversity. It uses the formulation of Shannon and is interpreted as **a measure of randomness of a probability distribution**. Also, it has been used with **complexity measures** to discriminate among water flow models.

- **Laura Poggio (The James Hutton Institute, Scotland)**

It is difficult to choose a favourite equation. However, after some thinking, I went for the Bayes Theorem. It describes the probability of an event, based on conditions that might be related to the event itself. I choose this equation because it deals with probabilities and uncertainty, something very important and discussed in Pedometrics, and because it represents an approach to uncertainty mapping different from the frequentist simulations. And here it is:

$$P(A \mid B) = \frac{P(A)P(B \mid A)}{P(B)}$$

where A and B are events observed jointly;
P(A) and P(B) are the marginal probabilities of A and B;
P(A | B), a conditional probability, is the probability of observing event A given the occurrence of B;
P(B | A), is the probability of observing event B given the occurrence of A.

- **Murray Lark** **(British Geological Survey)**

I had a chance to display some of my favourite equations in Pedometron a few years ago when I set a Pedomathemagica question for Pedometrics 2007, Tubingen. Gerhard won it! Of those my favourite is probably Krige's relation, which is the partition of variance into between-block and within-block components. It is still useful, for example I invoked it in a recent paper for Geoderma. If $\sigma^2_{\mathrm{b,B}}$ is the variance of Z on support b within a block B, and a region R can be partitioned into many such (non-overlapping blocks) then Krige's relation states that

$$\sigma^2_{\mathrm{b,R}} = \sigma^2_{\mathrm{b,B}} + \sigma^2_{\mathrm{B,R}}$$

Another good equation is for the Lorentz factor, the factor by which time and length change for a moving object in special relativity. If v is the relative velocity between two reference frames (e.g. the velocity of a moving observer relative to a stationary one) then the Lorentz factor is

$$\gamma = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}}$$

where c is the velocity of light. I like this equation because it is actually very straightforward to derive from the light-clock thought experiment. Google the experiment then derive the equation for yourself!

- **Nicolas Saby** **(INRA Orleans)**

My favourite equation is the minimum detectable change of soil property y where :

$$y = z_\alpha \sqrt{s_d^2 / n}$$

In this equation, n is the number of observations and s_d is the estimate of the variance of the differences between two sets of observations, namely two sets of observation of a soil property at two time.

I like this equation because its convenience in the near future. Everybody knows what was happening in Paris this December: COP21. The governments of more than 190 nations gathered in Paris to discuss a new global agreement on climate change, aimed at reducing global greenhouse gas emissions and thus avoiding the threat of dangerous climate change. The agricultural and forestry sectors offer solutions for climate change mitigation through preserving or increasing carbon stocks in soils and biomass. The management of organic matter, which is the principal carbon reservoir in soils, is a key determinant in the capacity of soils to produce food and material and to offer other environmental services such as the regulation of water cycles and air quality etc. Acting on soil carbon stocks also means acting on soil and environmental quality. This is the sense of the '4 per 1000' initiative proposed by France at COP21.

Along with a large amount of soil carbon data that has to be collected with appropriate statistical design, pedometricians might use this equation to verify the effective additional carbon stocks induced by the effects of changes in land use or management practice.

Digital Soil Mapping Training at The Dokuchaev Soil Science Institute, Moscow

The Dokuchaev Soil Science Institute was recently awarded the prestigious grant from the Russian Scientific Foundation on "Large-scale digital soil mapping using remote sensing data". As part of the requirements, the Dokuchaev Institute held a short course on digital soil mapping for Russian post-graduate students and young scientists.

To celebrate the Year of the Soils, some of us were honoured to be invited as instructors for the course which was held from 27th November until 2nd December 2015. The program was prepared by Pavel Krasilnikov and Igor Savin. A preliminary module was delivered on 27th Nov by Daniil Kozlov and Joulia Meshalkina from Lomonosov Moscow State University. On the 30th Nov, we were greeted by the Director of the Dokuchaev Soil Science Institute Dr. Andrey Ivanov, and Executive-Director, Dr. Igor Savin. Dr. Ivanov explained the history and role of the institute and the current need for updated soil information. Land privatization is currently happening in Russia, and there is a need for detailed soil information in order to assess its potential use and environmental impact.

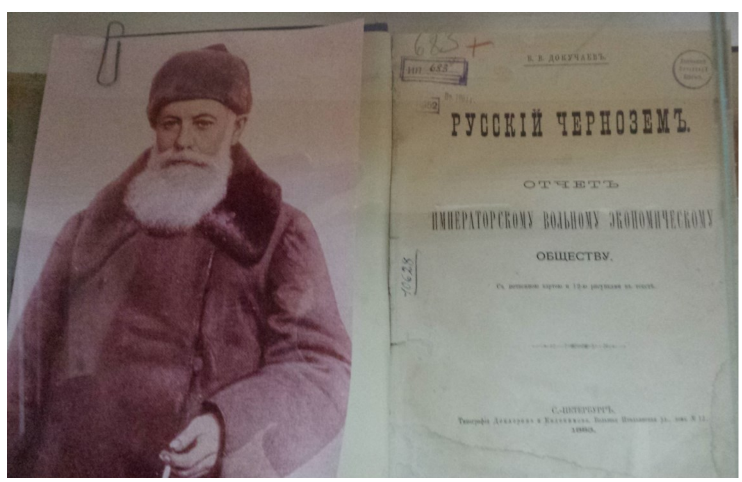
Philippe Lagacherie from INRA Montpellier, France started the program with an introduction to Digital Soil Mapping and the use of remote sensing hyperspectral infrared data. This was followed by A-Xing Zhu from University of Wisconsin, USA who introduced the use of expert system and environmental similarity approach. A-Xing then demonstrated the use of the SOLIM software for soil mapping to enthusiastic students.

The second day, Gerard Heuvelink from ISRIC and Wageningen University, The Netherlands introduced the concepts of geostatistics with lectures and practicals. This is followed by John Triantafilis from the University of New South Wales, Australia talking about the use of electromagnetic induction and gamma radiometrics data in digital soil mapping. Finally Tom Hengl from ISRIC provided demonstrations of popular digital soil mapping tools: R, SAGA, and Google Earth.

The final day, Budiman Minasny from the University of Sydney, Australia discussed and reviewed methodology of digital soil mapping. This was followed by hands-on practical sessions using R for soil depth harmonisation with Brendan Malone from the University of Sydney. The day continued with lecture and practicals on mapping soil classes, and creating web-based interactive maps.

While it was a short, we were all impressed with the students' enthusiasm on digital soil mapping. We hope that this introductory concept will prepare the students with their task ahead in producing high resolution digital soil maps in Russia.

We'd like to thank Igor Savin & Pavel Krasilnikov for their invitation. It is indeed an honour to present at the home of soil science. Maria Konyishkova helped with the arrangement of the whole administration and program. Daniil Kozlov and Joulia Meshalkina introduced us to the culture and architecturally beautiful Moscow.



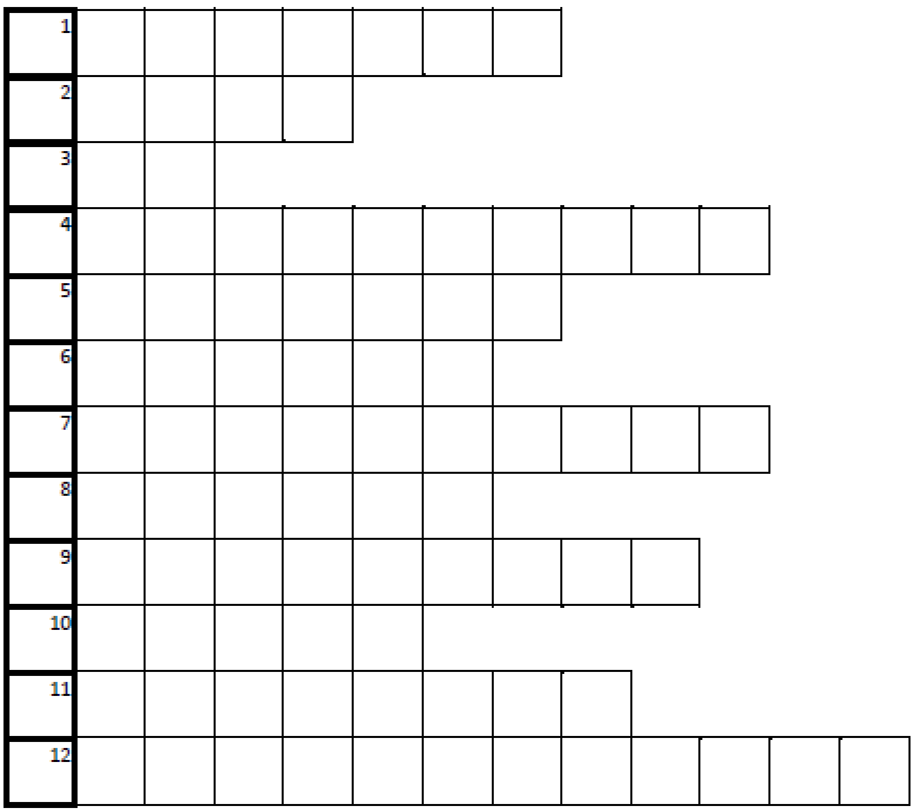
Photos courtesy of John Triantafilis

How I got involved in pedometrics

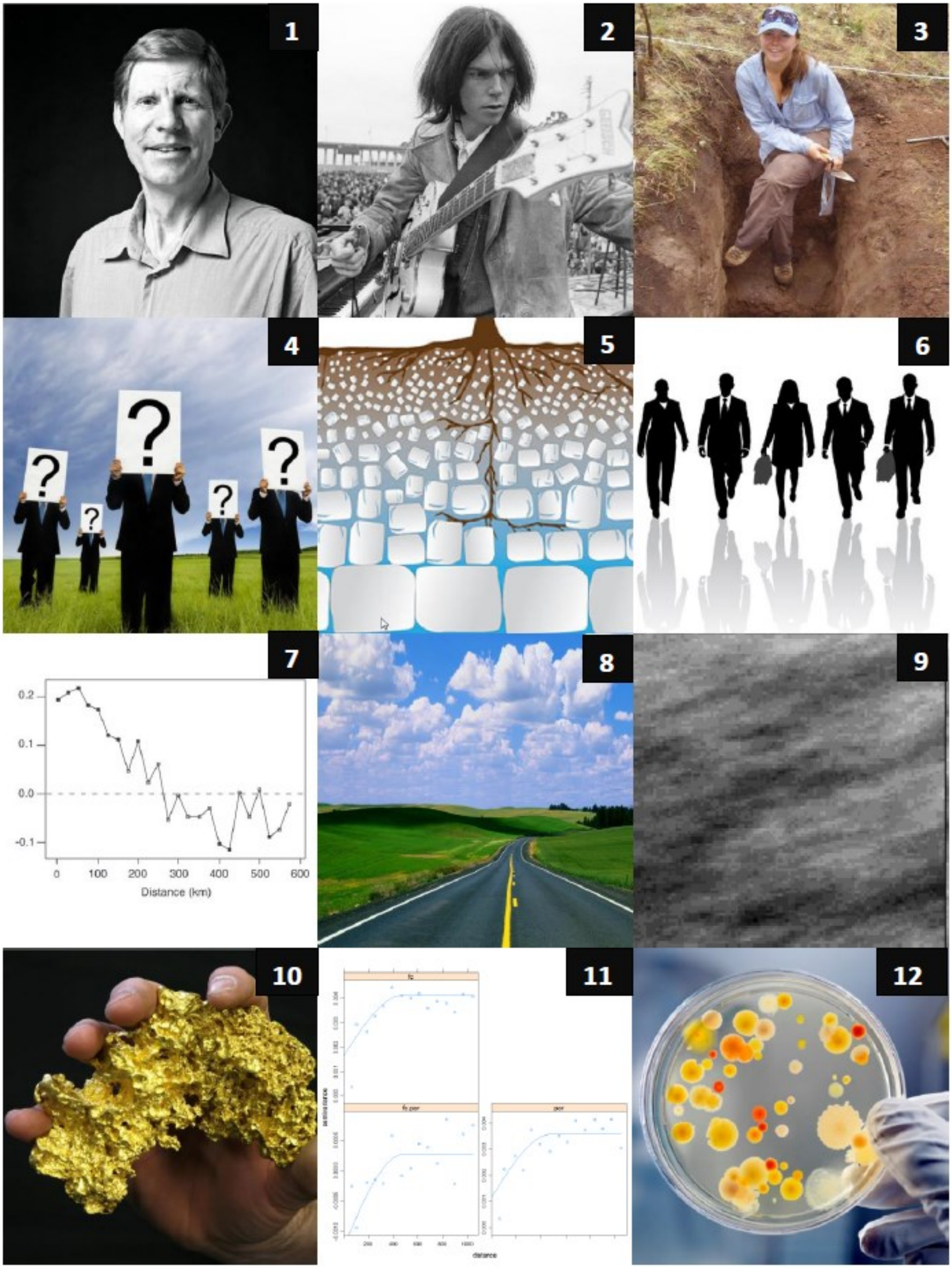
Gerard Heuvelink
ISRIC - World Soil Information and Wageningen University

Budi has started a new tradition by asking recipients of the Richard Webster medal to contribute an article on how they got involved in pedometrics. I am happy to do so, but rather than providing a historic narrative of how I ended up in pedometrics, I decided to wrap the answer in a puzzle. So If you are curious about the real reason behind me becoming a pedometrician, solve the puzzle below and be amazed!

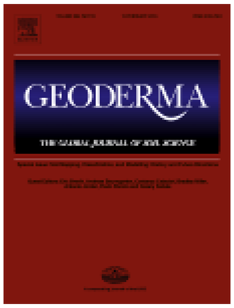
The answer is given by the first letters of twelve words. Each word is described by a sentence and a picture. Together these should provide you with just enough information to solve this puzzle!



1 = A founding father of pedometrics
2 = With fewer than 25 years, pedometrics still qualifies as such
3 = Do you dig this?
4 = The GSM specs say: quantify it!
5 = Deeper than just the soil
6 = What would we do without them?
7 = Mirror the variogram and standardise it
8 = ABC
9 = Direction matters
10 = Few soil variograms can do without it
11 = Call in the help from others to improve prediction
12 = How soil micro-organisms grow under perfect conditions


1
2
3
4
5
6
7
0.2
0.1
0.0
-0.1
0
100
200
300
400
500
600
Distance (km)
8
9
10
11
12

Geoderma Special Issue on Advances in DSM, Uncertainty and Soil Carbon Validation



This Geoderma special issue features selected papers from the biennial meeting of the Pedometrics Division 1.5 of the International Union of Soil Science in Nairobi, Kenya in August 26–31, 2013 and Pedometrics sessions at the 20th World Congress of Soil Science, Jeju, Korea in June 8–13, 2014. These meetings bring together international soil scientists and pedometricians to present the latest advances in the field of Pedometrics.

The collection of articles highlights the various spatial scales and approaches scientists' use to tackle key issues that advance the field of Pedometrics. The first paper by Vågen et al., 2015 illustrates a statistical approach for soil mapping, which combines systematic field surveys with remote sensing data to create maps of erosion risk and dynamic soil properties, for sub-Saharan Africa at a resolution of 500 m. The next paper by Lang et al., 2015 highlights the utility of existing soil databases to derive soil groups with the World Reference Base. The paper by Jiang et al., 2015 showcases an online and high-performance-enabled cyber environment for digital soil mapping.

In addition to soil mapping, pedometrics also addresses new analytical techniques to soil analysis including the use of visible/near-infrared spectroscopy. The paper by Fajardo et al., 2015 highlights a new application of spectroscopy: to identify soil horizons, which advances beyond standard prediction of soil properties.

Soil organic carbon is an important component of soil health. Thus, much attention had been paid to mapping and modelling surface and below ground carbon stocks and concentrations. The remaining three papers highlight new approaches for assessing SOC. Liu et al., 2015 present a new method for three-dimensional SOC predictions. Wiese et al., 2015 share a method of modelling the vertical distribution of carbon in a catchment in KwaZulu Natal in South Africa. The final paper by Winowiecki et al. highlights a new method for soil carbon stock accounting and mapping, which was applied in Tanzania. This method was used to assess ecosystem services and the effect of land degradation and vegetation cover highlighting the importance services provided by soil.

Read it all at: http://www.sciencedirect.com/science/journal/00167061/263

Upcoming Events

Digital Soil Mapping 2016

The 7th Global Workshop on **Digital Soil Mapping** will be held in Århus, Denmark, 27 June -1 July 2016. Feauturing all aspects and the latest development in Digital Soil Mapping

More info at http://digitalsoilmapping.org/

Geostats2016

5 to 9 of september
Valencia
GEOSTATS2016

The 10th International Geostatistical Congress will be held in Valencia, Spain, 5-9 September 2016. More info here . Abstracts submission is now open and Jaime Gomez-Hernandez (keynote speaker at Pedometrics 2015) is the chairman of the organizing committee and can be contacted at jgomez@upv.es

DSM at EGU: Digital soil mapping: from map products to a wider understanding of soil-landscape relations

(SSS 12.1)<http://meetingorganizer.copernicus.org/EGU2016/session/20969>

Convener: Jacqueline Hannam

Co-Conveners: Gerard Heuvelink, Arnaud Temme

Deadline for abstracts is 13 January 2016

ISRIC Spring School

Spring School on digital soil mapping, soil classification and assessment, ISRIC – World Soil Information, Wageningen, The Netherlands, 9-13 May 2016. For information and registration: http://springschool.isric.org

ΠΕΔΟ
METRICS
25
WAGENINGEN 2017

Pedometrics 2017

Pedometrics will be celebrating its Silver Jubilee in 2017. The first pedometrics conference was held in Wageningen in September 1992 and Pedometrics 2017 will be held on **26 June to 2 July 2017** at Hof van Wageningen . This will also be a joint conference with four of its Working Groups:

1. Digital Soil Mapping
2. Proximal Soil Sensing
3. Soil Monitoring
4. Modelling of Soil and Landscape Evolution

Mark on your calendar now for this once in a lifetime important event!

More info will be available at http://www.pedometrics2017.org/