ΠΕΔΟΜΕΤRΟΝ

The Newsletter of the Pedometrics Commission of the IUSS

Issue 36, December 2014

Chair: Budiman Minasny Vice Chair: Lin Yang Editor: Jing Liu

From the Chair

It has been a big year for soil scientists and the IUSS. We had the 20th World Congress of Soil Science in Jeju, and the Pedometrics commission organised 2 successful sessions: Validation of Soil Carbon Sequestration, and Quantification and Application of Uncertainty in Pedometrics. In this meeting, A-Xing also formally handed the Chair and Vice-Chair positions to me and Yang Lin. Thanks to A-Xing and Dick Brus for taking care the Commission for the past 2 years. We also congratulate Gerard Heuvelink for being awarded the Richard Webster Medal.

We can proudly celebrate our successes, the Pedometrics Commission has many activities: we have the annual Best Paper Award (please don't forget to vote), the Richard Webster Award, and Pedometron – the biannual newsletter. We now also sent out a regular Pedometrics News to keep the communication "alive".

Next year is the International year of Soil, and we need to be more active, not only to the soil science community but also promoting soil science to the public. We will have our biennial Pedometrics conference in Cordoba Spain. And for the community, Ana Horta from Charles Sturt University, Australia, has proposed "The soil in my backyard" as an activity for the Pedometrics Commission. The idea is to introduce kindergarten and primary school students (ages 5 - 10) to Soil by engaging them in an outdoor activity. The soil activity with the kids will then be recorded and uploaded to our website, together with a Google Earth location to map our junior "soil scientists". Ana is currently forming a project description and every member of the Pedometrics board would be responsible to disseminate this project. I invite all of you to take this challenge!

Alongside these achievements, we also need to think about the scientific direction of Pedometrics. Geostatistics in now a common tool and already a general subject taught at an undergraduate level, and digital soil mapping has now become operational. One of our colleagues that worked for a state department in Australia said now digital soil mapping is part of their core business activity. Within the past 10 years, Digital Soil Mapping has shift from research-based activity to every-day business. It has been an exciting time to see all of the pedometrical techniques are now adopted and applied. So what's next? From my point of view, the dynamic spatiotemporal modelling would be the next challenge, and I am sure you also have other important areas. If you have any exciting challenges, we'd like to hear from you.

Finally, I wish you a wonderful festive season. Refresh, Refocus and Ready for a bigger and more productive 2015.

Budiman Minasny

Sydney, December 2014.

Inside This Issue

Pedometrics
Celebrate Year of the Soil 2015

We are very excited to invite you all to Cordoba Spain, for Pedometrics 2015, September 15th-18th, 2015.

This Conference will also incorporate meetings for the IUSS WG on Soil Landscape Modelling and Soil Monitoring. Topics include: Soil-landscape modelling: mechanistic & empirical, Soil Morphometrics (image analysis, remote sensing, 3D soil imaging), Soil sampling and monitoring, Field experimental design, Digital soil mapping and proximal soil sensing, Bayesian statistics and Hierarchical Modelling in soils, Fuzzy cognitive mapping, Soil Spatial and Temporal Scaling, and Soil Ecosystem Services. More info at: https://sites.google.com/site/pedometrics2015/home

Conference Program:

Sunday 13 Sep	• Arrival of participants + ice-breaker
Monday 14 Sep	• Pre-conference workshops WG IUSS Soil landscape modelling (hands-on modelling workshop comparing and benchmarking soil formation models) • WG IUSS Soil monitoring
Tuesday 15 Sep	• Main Pedometrics conference, integrating WG workshops sessions • Welcome reception (evening)
Wednesday 16 Sep	• Main Pedometrics conference, integrating WG workshops sessions
Thusday 17 Sep	• Main Pedometrics conference, integrating WG workshops sessions • Conference dinner
Friday 18 Sep	• Optional excursion (soil landscape modelling and validation with geophysical sensors - semi-natural area (dehesa Santa Clotilde, Cardeña) and heavily eroded olive groves with shallow soils (Montoro))

Best Paper Nominees in Pedometrics 2013

by D G Rossiter

The committee received nine nominations. These were all scored by the committee and the top five are now presented for your reading pleasure and evaluation. Following are the references, in first author alphabetic order, and abstracts. There is a nice mix: geostatistics, sampling design, a pedometrics computation toolkit, spatial scaling, and numerical methods for spectroscopy. All are quite novel in their own way, and will surely stimulate and educate the reader. We are also glad to announce that all of the nominated papers are now available free until end of the year. Thanks to the publishers. We also asked the authors to give us a summary of their paper, they really worth a read.

Now, Please vote for the 2013 Best Paper. The deadline for voting is end 2014. Please rank the papers in the "instant runoff" system (first choice, second choice... up till the last paper the voter is willing to vote for, i.e., the last paper that the voter thinks would deserve the award). Votes should then be sent to me (**dgr2@cornell.edu**) from a traceable e-mail address (to prevent over-voting). I will apply the "instant runoff" system to determine the winner. A co-author may vote for her/his own paper(s).

The Best Paper will be announced early next year and the award will be presented in Pedometrics 2015 (September) in Córdoba.

The nominees (with a summary) are:

1. Beaudette, D.E., Roudier, P., O'Geen, A.T., 2013. Algorithms for quantitative pedology: A toolkit for soil scientists. Computers & Geosciences 52, 258–268.

We developed AQP (Algorithms for Quantitative Pedology) as a software package in the R statistical environment so that we can easily perform common tasks such as visualization, aggregation, and classification of soil profile data. As soil data can be associated with location (x, y), depth (z), and property space (p); the high dimensionality and grouped nature of this type of data can complicate standard analysis, summarization, and visualization tasks. The AQP package provides pedometricians with an analysis framework that can handle the complexity of soil profile data.

AQP provides pedometricians with tools for rendering soil profiles graphically, based on horizon boundaries, horizon designation, soil color and soil properties (measured or inferred). Soil profiles can be plotted in a given order, e.g. by the type of landform from which they have been sampled.

Soil property data organised according to genetic horizons are difficult to process due variable horizonation depths. A solution to this problem is offered in AQP by normalizing a collection of horizons, irrespective of the horizon type, according to a common system of ''slices''. Essentially, each soil property (from each soil profile) is aligned to a common depth basis. With this new data structure it is possible to plot, aggregate, map, or compute numerical measures of similarity by slice. The use of aggregate depth functions could support a fundamental shift in how soil survey is presented: from the concept of a ''modal profile'' (i.e. a single pedon) to a collection of ''representative depth functions''. Representative soil property depth functions would give users a continuous estimate of soil properties and fulfil a long-standing criticism of soil survey regarding the current lack of uncertainty estimates for soil property data.

In addition, the normalization of soil profile horizons allows to run similarity analysis on any given collection of soil profiles. Such similarity measures can then be used in numerical soil classification by leveraging the important base of classification methods published within the extensible and open source R framework. Alternative classification schemes could also be generated from the same underlying data, but directed towards specific goals, by selecting which variables and dissimilarity metrics are used.

Functions in the aqp package have been successfully applied to studies involving several thousand soil profiles. AQP is an open source project. Its scriptable nature allows pedometrics research to be reproducible, but also embeddable in other tools, such as the SoilWeb mobile app.

The stable version of the aqp package is hosted on CRAN, and the development version is hosted on R-Forge. A recent presentation on the "aqp family" of R packages can be found here.

2. Lark, R.M., Lapworth, D.J., 2013. The offset correlation, a novel quality measure for planning geochemical surveys of the soil by kriging. Geoderma 197–198, 27–35.

One of the useful things about geostatistical prediction is that if you know the variogram of a soil property then you can compute the mean-squared error of the kriging prediction for any location relative to some hypothetical sampling grid. This means that you can find a sample grid that will allow you to map the property with adequate precision and to avoid over-sampling. In 1981 Alex McBratney, Richard Webster and Trevor Burgess wrote a paper in Computers and Geosciences where they pointed this out and described a computer program to do it. This is a simple but elegant approach, and should appeal to the practitioner.

This method has been useful in practice, but it sometimes runs into problems of communication. We have experienced these problems in dealings with government, management, colleagues, farmers and advisors. Even if the manager or official who makes the decision on funding for a survey understands variances the mean square error of predictions is not always useful for planning general baseline surveys with many possible end-users.

The geochemist, or indeed the farmer or other environmental manager, is very aware of the existence of spatial variation. In our experience this sometimes makes them sceptical of the kriged map. "Aha," they say, "but if this sample point had been 100m away in the next field then the pH would have been much lower because they never lime that one." Our proposed criterion builds on this entirely sensible intuition.

Consider a region sampled on a 500-m square grid. We collect the data, analyse them, and produce a map. Now, what would happen if another team, using all the same methods and equipment, sampled at exactly the same intensity, but with their points all 250 m north and 250 m east of the original grid? The new map will not be identical, but just how different will it be? How sensitive, in short, is our overall procedure (including the grid spacing) to an arbitrary shift in the origin of the sample grid? It seems reasonable to propose that a robust sampling scheme to map a spatial variable should not be sensitive to this offset. How sensitive it is will depend on the spatial variability and the grid spacing.

Just as with the kriging variance one can compute, from the variogram alone, the correlation between predicted values on two maps made with the same grid density but a half-grid offset of the origin. We call this the offset correlation. It is a bounded measure of the consistency of the map under arbitrary shifts of the origin, potentially easier to explain intuitively to the data user than is a variance. In our paper we show some hypothetical and real examples of the offset correlation both for ordinary kriging and factorial kriging, considering geochemical data from the East of England and comparing the designs of two national-scale soil sampling schemes from the UK.

3. Malone, B.P., McBratney, A.B., Minasny, B., 2013. Spatial Scaling for Digital Soil Mapping. Soil Science Society of America Journal 77, 890.

A remarkable growth in the application of Digital Soil Mapping (DSM) is currently being experienced around the world. It is being used to address important environmental issues over a range of spatial extents — fields and farms, landscapes and regions, countries and continents, and importantly, globally. There is potentially a significant amount of comprehensive spatial soil information throughout the world. Yet, what we have recognised, and what could potentially be an operational hurdle further into the future, is the inequality between the 'scale' of the digital soil information which is available, and that which is required to address a particular issue or question. For example, a soil organic carbon map produced at the national extent (perhaps for a national soil carbon accounting purpose), may be inappropriate at the field extent. Alternatively, existing digital soil information may be available for points, but is required over small areas i.e. each prediction represents an integral of the soil attribute of interest across the spatial dimensions of the area. The problems may be addressed through implementation of spatial scaling methods. Our paper examines this issue and provides a broad overview of spatial scale concepts and spatial scaling procedures that are specifically relevant for DSM.

First, we clarify some fundamental concepts of scale. Essentially, digital soil maps have three spatial scale entities: extent, resolution, and support. Map extent is the areal expanse or coverage of a mapping domain. Resolution is the grid-cell spacing or pixel size of the raster. While support is likened to a volume or area.

Secondly we set about describing a suite of pedometric techniques that could be used for spatial scaling. One may think of these as upscaling or downscaling methods, where upscaling may involve an increase in extent, support, or grid cell resolution size, which could either be modified conjunctively or just focusing on one or two spatial entities only. Downscaling is essentially the opposite process of upscaling. However, we describe spatial scaling for DSM with due reference to the scale entities (extent, resolution, support). Such that fine-gridding, deconvolution, and disseveration are different variants of downscaling operations. While coarse-gridding, convolution, and conflation are variants of upscaling operations. We provide some of the theory of each of these operations, and provide examples of their usage either from the literature or from our own data.

One important discovery from this paper is that we can efficiently, without changing the grid cell resolution; obtain areal or block averages from point support maps. This is likened to going from a source digital soil map with 20m point support to a destination map where the predictions represent averages on supports of 20m × 20m. This can be achieved with block kriging. Another useful application of spatial scaling for DSM is the downscaling from coarsely resolved to finely resolved maps.

We feel we have thrown light on the issue, provided some new solutions, and are very optimistic about future developments which inevitably will help to solve pressing soil and environmental issues around the world.

4. Meerschman, E., Van Meirvenne, M., Van De Vijver, E., De Smedt, P., Islam, M.M., Saey, T., 2013. Mapping complex soil patterns with multiple-point geostatistics. European Journal of Soil Science 64, 183–191.

The commonly used variogram function is incapable of modelling complex spatial patterns associated with repetitive, connected or curvilinear features, because it is a two-point statistic. Because this was strongly limiting to petroleum- and hydrogeologists, they developed multiple-point geostatistics (MPG), an approach that replaces the variogram by a training image (TI). However, soil scientists also face complex spatial patterns and MPG might be of use to them as well. Therefore, this paper aims to introduce MPG to soil science and demonstrate its potential with a case study of polygonal subsoil patterns caused by Weichselian periglacial frost cracks in Belgium.

A high-resolution proximal soil sensing survey provided a reference image from which a continuous (655 sensor data) and a categorical (100 point observations) dataset were extracted. As a continuous TI, we used the geophysical data of another part of the field, and as categorical TI we used a classified photograph of an ice-wedge network in Alaska. The resulting MPG maps reconstructed the polygonal patterns very well and corresponded closely to the reference image. Consequently, we identify MPG as a promising technique to map complex soil patterns and suggest that it should be added to the pedometrician's toolbox.

5. Mulder, V.L., Plötze, M., de Bruin, S., Schaepman, M.E., Mavris, C., Kokaly, R.F., Egli, M., 2013. Quantifying mineral abundances of complex mixtures by coupling spectral deconvolution of SWIR spectra (2.1–2.4 µm) and regression tree analysis. Geoderma 207–208.

Soil mineralogy is an important indicator for soil formation and parent material characterization. In environmental and geological studies, the characterization (and quantification) of soil mineralogy is typically achieved using X-ray diffraction (XRD). Visible Near Infrared and Shortwave Infrared (VNIR/SWIR) spectroscopy has proven to be an efficient alternative for the determination of various soil properties. In this paper we propose and demonstrate its use for simultaneous quantification of mineral abundances from complex mixtures. Detection of minerals having absorption features within the 0.350–2.500 µm range have been successfully obtained using linear spectral unmixing techniques. However, these analyses were limited to estimating the main component within a sample having the most distinct absorption feature. Hence, reflectance spectra of mixtures are typically a complex result from the combinations of the spectral characteristics of the constituents. Depending on the composition, the abundance and the spatial arrangement of the minerals, the total reflectance resulting from the scattering of the minerals within the intimate mixture produces positional shifts, changes in intensity, disappearance of absorption features or changes in their shape.

Hence in this work we aimed to quantify mineral abundances using spectral deconvolution (SD) followed by regression tree analysis (RT). SD involves modelling the total reflectance and the inference of absorption components within complex features by fitting (modified) Gaussian curves to the absorption features and absorption components. Next, mineral abundances were predicted by RT using the parameters of the fitted Gaussians as inputs. The approach was demonstrated on a range of prepared samples with known abundances of kaolinite, dioctahedral mica, smectite, calcite and quartz and on a set of field samples from Morocco.

Cross validation showed that the prepared samples of kaolinite, dioctahedral mica, smectite and calcite were predicted with a root mean square error (RMSE) less than 9wt%. For the field samples, the RMSE was less than 8 wt% for calcite, dioctahedral mica and kaolinite abundances. Smectite could not be well predicted, which was attributed to spectral variation of the cations within the dioctahedral layered smectites. Substitution of part of the quartz by chlorite at the prediction phase hardly affected the accuracy of the predicted mineral content; this suggests that the method is robust in handling the omission of minerals during the training phase. The degree of expression of absorption components was different between the field sample and the laboratory mixtures. This demonstrates that the method should be calibrated and trained on local samples. Concluding, our method allows the simultaneous quantification of more than two minerals within a complex mixture and thereby enhances the perspectives of spectral analysis for mineral abundances.

Noteworthy Articles

• *Performing Kriging with Privacy*

Most of us know that we need data with geographical locations to be able to perform kriging. Researchers Bulent Tugrul and Huseyin Polat from Turkey recently posed the problem of data privacy. Their paper was published in International Journal of Innovative Computing, Information and Control, August 2013, and again in Knowledge-Based Systems in May 2014. An online service provider has data of soil measurements from various locations within a region, and a client wish to predict soil property in a location. Because of privacy, neither the client nor the server wants to reveal their (geographical coordinates) information to each other. The researchers proposed a scheme, which helps the clients and the servers perform Kriging interpolations while protecting their confidentiality. The authors first outline a naïve scheme which is simply the client incorporates fake locations into her data to send out to the server. The authors then proposed an improved scheme which involves encrypting the distances between the server data and client. The authors concluded that "the proposed scheme protects privacy and it does not cause any accuracy losses. We also analyze it with respect to inevitable additional costs, which do not affect online performance." With the increase in crowd-sourcing data, and the creation of global soil databases, this issue of geographical location privacy will become important, and we may need to resort to this algorithm.

• *Comparing traditional and digital soil mapping approach*

Romina Lorenzetti together with Edoardo Costantini, Maria Fantappiè, Roberto Barbetti and Giovanni L'Abate of CRA-ABP (Florence, Italy) recently published a paper in *Geoderma* on: "Comparing data mining and deterministic pedology to assess the frequency of WRB reference soil groups in the legend of small scale maps". This is one of the few studies that compare traditional and digital technqiues.

The authors compared 1:5,000,000 map of Italian soil regions and digital soil mapping at 1 km grid spacing in predicting the WRB reference soil groups (RSGs). They selected 5 of the 10 soil regions in Italy to cover half of the Italian land area. The soil regions range from 16,000 to 47,000 km2 with profile desnity between 0.05 to 0.24 profiles/km2, and each region comprises 18 to 25 RSGs. Data mining techniques were tested against traditional approach in estimating the frequency of WRB classes in the legend of the 1:5,000,000 soil region map. 10% of the profiles in each region were set aside as a "test". They used a Bayesian validation approach to calculate the positive predictive and negative predictive indices based on "test" soil profiles in each of the region. The positive predictive value expresses the probability that an event occurs when the model estimates it to occur; and vice versa for negative predictability.

Their results showed that Support Vector Machine (SVM) performed best and better than the traditional maps. The mean of positive predictive value is 0.442 for SVM, while traditional maps only achieve a value of 0.220. The authors concluded that "this work suggests that the SVM method is better than the traditional approach. A future challenge is to test the suitability of data mining to estimate soil class frequency in the legend of maps produced at more detailed scales or with other spatialization methodologies."

Noteworthy Articles

• *How to effectively bend wires across a landscape*

Uta Stockmann and colleagues recently asked the question, can we perform proximal soil sensing of gamma radiometrics for a large region. As we know, proximal soil sensors such as gamma spectrometers can be attached to an ATV and collect data on the go. Most applications are within a field (up to 100 ha), where the data is collected with a line spacing of 10 to 30 m. If we want to survey a large area say of several hundreds of km^2, it is not feasible to drive at such a small spacing. In addition, most areas within a region will be constrained by fences and roads.

Thus the authors came up with an algorithm called WIRES to undergo Wide-Ranging Exploratory (soil) Surveys. The idea is to first place “sampling points” throughout the survey area using an established design such as the conditioned Latin Hypercube Sampling (cLHS) based on ancillary environmental information. Rather than driving the vehicle on defined, parallel line spacing, a figurative meaning of the algorithm is the arrangement and bending of wires across a landscape to connect the established “sampling points”. The authors tested this approach across an area in the Hunter Valley, NSW, Australia. Based on driving along these WIRES transects, they generated radiometric maps using a regression kriging modelling approach. They concluded that “surveying a location of interest using optimal wide-ranging transects (≥ 100 m transect width) is sufficient enough to capture the soil's (heterogenic) radioelement concentration, and that a detailed fine-scale, high resolution survey (30 m transect width) may not be required to be able to measure the variability of soil attributes.”

***Reference*:**

Stockmann, U., Malone, B. P., McBratney, A. B., & Minasny, B. (2015). Landscape-scale exploratory radiometric mapping using proximal soil sensing. *Geoderma*, 239, 115-129.

• *Collaborative research results in higher citations rates*

Vincent Larivière et al. from Montreal recently provides the first historical analysis of the relationship between collaboration and scientific impact from papers published from 1900 to 2011. The authors analysed data from over 28 million publications from Natural and Medical Sciences (NMS) and 4 million papers from Social Sciences and Humanities. They analysed types of collaboration in terms of: co-authorship (number of authors), inter-institutional collaboration (number of addresses), and international collaboration (number of countries). Not surprisingly they found that the number of single authored papers decrease linearly over time from 90% in the 1900 to 7% of papers in 2011 (for NMS). It is rare to find a single authored paper nowadays. The most frequent numbers are: 4-5 and 6-10 authors. A similar trend is also observed for the number of addresses and number of countries represented in the article. And most interestingly, the more authors in a paper, the higher number of citations it will get. So start collaborating!

Reference:

Larivière, V., Gingras, Y., Sugimoto, C. R., & Tsou, A. (2014). Team size matters: Collaboration and scientific impact since 1900. *Journal of the Association for Information Science and Technology*.

Noteworthy Articles

• *Kriging and Pedotransfer functions*

Generally we view kriging as a spatial interpolation technique and pedotransfer functions as empirical soil relationships. Tiago Ramos and colleagues from Portugal recently proposed estimating soil water retention based on kriging technique. In essence it is predicting water retention using sand, silt, and clay content. Rather formulating an empirical formula, the authors used kriging to interpolate existing observations in a texture triangle. The available water content of Portuguese soils was then derived from interpolated volumetric water content at -33 kPa and -1500 kPa in ternary diagrams. The authors concluded that “The hydraulic ternary diagrams may thus serve as simplified tools for estimating water retention properties from particle size distribution and eventually serve as an alternative to the traditional statistical regression and data mining techniques used to derive PTFs.”

Reference:

Ramos, T. B., Horta, A., Gonçalves, M. C., Martins, J. C., & Pereira, L. S. (2014). Development of ternary diagrams for estimating water retention properties using geostatistical approaches. *Geoderma*, 230, 229-242.

Biopedometrics

by Alex McBratney

There is always a question about what we should work on.

For about 35 years pedometricians have worked on soil classification, spatial analysis of soil properties, digital soil mapping, monitoring and applied areas such as precision agriculture and soil remediation assessment. Most of this work has focused on chemical and physical soil properties.

More recently there has been a huge upsurge in biological soil studies.

This has moved from single organism studies that control biochemical processes to wider studies that look at biodiversity and food webs. Many of these studies have ecology as a disciplinary basis. Ecology has iTs own statistical methodologies as exemplified in the excellent book by Legendre and Legendre (2012). A good example is the recent study by de Vries et al. (2013). In this we see the authors use 'spatial filters' that relate soil properties, land use, and soil food web characteristics. These filters have an interesting effect on the conclusions to the study.

In these kinds of study there seems little cognisance of pedometric approaches. I see pedometrics can bring soil science together by linking the new soil biology with soil physics, chemistry and pedology. We need a small but eager cohort of biopedometricians and projects to link pedometrics with quantitative soil ecology. Key questions for this new area are: what is the nature of the function that relates soil functionality with soil biodiversity? Is it simply linear as much of the work implies? Linear functions are rare in ecology! To what extent and how does soil diversity control soil biodiversity (Ibáñez and Feoli, 2013) ?

Good examples of BIOPEDOMETRICS are the papers by Griffiths et al. (2011) and Ranjard et al. (2013) which show strong spatial patterns of soil biodiversity which question the old ideas that any organism could pop up anywhere with equal probability.

Let's get stuck in.

References

de Vries, F.T., Thébault, E., Liiri, M., Birkhofer, K., Tsiafouli, M.A., Bjørnlund, L., et al., 2013. Soil food web properties explain ecosystem services across European land use systems. *Proceedings of the National Academy of Sciences*, 110(35), 14296-14301.

Ibáñez, J. J., Feoli, E., 2013. Global relationships of pedodiversity and biodiversity. *Vadose Zone Journal*, 12 (3), doi:10.2136/vzj2012.0186.

Legendre, P., Legendre, L.F., 2012. *Numerical Ecology* (Vol. 20). Elsevier, Amsterdam.

Griffiths, R.I., Thomson, B.C., James, P., Bell, T., Bailey, M., Whiteley, A.S., 2011. Bacterial diversity of British soils. *Environmental Microbiology,* 13 (6), 1642–1654.

Ranjard, L., Dequiedt, S., Chemidlin Prévost-Bouré, N., Thioulouse, J., Saby, N.P.A., Lelievre, M., Maron, P.A., Morin, F.E.R., Bispo, A., Jolivet, C., Arrouays, D., Lemanceau, P., 2014. Turnover of soil bacterial diversity driven by wide-scale. *Nature Communications* 4, Article number: 1434. doi:10.1038/ncomms2431

Biopedometrics

Spatial Filtering

A method used in ecology to deal with the spatial structure of the species data is called principal coordinates of neighbour matrices (PNCM) or also called Distance-based Moran's eigenvector maps (dbMEM). The 3 simple steps are as follows (Borcard and Legendre, 2002):

(1) Compute a pairwise Euclidean distance matrix between the n sampling locations (**D**).

(2) Choose a threshold distance value *t* and construct a truncated distance matrix D* as follows:

if $d_{ij} \leq t$, $\mathbf{D}^* = d_{ij}$, else$\mathbf{D} = 4t$

the use of 4t was justified as Borcard and Legendre (2002) "observed that beyond a factor of four times the threshold for the 'large' distances, the principal coordinates remain the same to within a multiplicative constant"

(3) Perform principal coordinate analysis on the truncated distance matrix D*.

"The method basically consists of diagonalizing a spatial weighting matrix, then extracting the eigenvectors that maximize the Moran's index of autocorrelation. These eigenvectors can then be used directly as explanatory variables in regression or canonical models."

Pierre Legrende noted that he principal coordinates "represent a spectral decomposition of the spatial relationships among the study sites"

"If the sampling design is regular, they look like sine waves; this is a property of the eigendecomposition of the centred form of a distance matrix. If the design is irregular, the sine waves are distorted."

References

Borcard, D., Legendre, P., 2002. All-scale spatial analysis of ecological data by means of principal coordinates of neighbour matrices. Ecological Modelling, 153(1), 51-68.

Dray, S., Legendre, P., Peres-Neto, P. R., 2006. Spatial modelling: a comprehensive framework for principal coordinate analysis of neighbour matrices (PCNM). Ecological Modelling, 196(3), 483-493.

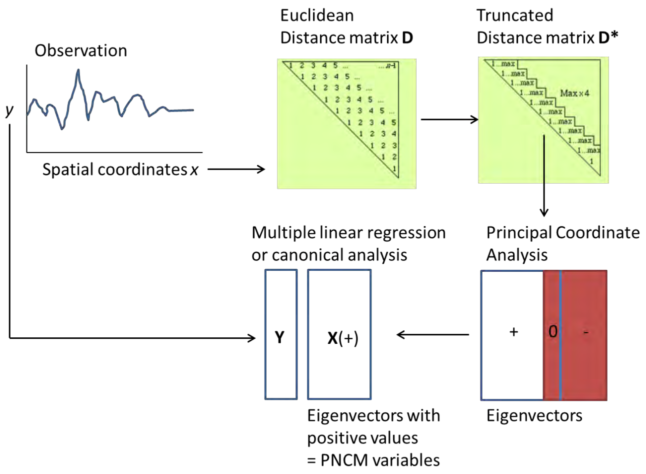


Figure 1. Principle of the principal coordinate analysis of a truncated matrix of Euclidean distances among observation sites (figure based on Borcard and Legendre, 2002).

Expert knowledge as prior information for spatial disaggregation using Bayesian area to point kriging

by Phuong N. Truong

What is ATP kriging?

In many cases we needed a fine resolution spatial data, however the only available data is a coarser resolution map. Thus we need spatial disaggregation to bring the available data to a finer level of spatial detail. Area to point kriging (ATP kriging) is one of the methods that can be used for spatial disaggregation (Kyriakidis, 2004). ATP kriging follows the principle of classical kriging in geostatistics and makes predictions of an attribute at point support (PoS) from block support observations (BSO) (i.e. data at coarse spatial resolution) of the same attribute. This is called *deconvolution* according to Malone et al. (2013) where the disaggregation produces finer grid spacing at a point support.

Let us first summarise the ATP algorithm. Assuming the variable of interest z to be a realisation of a second-order stationary Gaussian random function Z and let $\bar{\mathrm{z}}(\mathrm{B_i}) = \frac{1}{|\mathrm{B_i}|}\int_{\mathrm{s}\in\mathrm{B_i}} \mathrm{z}(\mathbf{s})\mathrm{d}\mathbf{s}$ be the value of z at block support, where $\mathrm{z}(\mathbf{s})$ is the value of z at point location $\mathbf{s}$ and $|\mathrm{B_i}|$ is the area of a block B indexed by i. Because the arithmetic averaging is linear in its argument, the random process at block support is also a Gaussian process.

Let $\boldsymbol{Z}_{\mathrm{p}} = (Z(\mathbf{s}_1), \ldots, Z(\mathbf{s}_M))^{\mathrm{T}}$ and $\bar{\boldsymbol{Z}}_{\mathrm{B}} = (\bar{Z}(\mathrm{B}_1), \ldots, \bar{Z}(\mathrm{B}_N))^{\mathrm{T}}$ denote vectors of Z at point and block support, then their joint probability distribution is jointly Gaussian:

$$\begin{bmatrix}\boldsymbol{Z}_{\mathbf{p}}\\ \bar{\boldsymbol{Z}}_{\mathbf{B}}\end{bmatrix} \sim \mathrm{N}\left(\mu\begin{bmatrix}\mathbf{1}_{\mathrm{M}}\\ \mathbf{1}_{\mathrm{N}}\end{bmatrix}, \begin{bmatrix}\mathbf{C_{pp}} & \mathbf{C_{pB}}\\ \mathbf{C_{Bp}} & \mathbf{C_{BB}}\end{bmatrix}\right) \quad (1)$$

where μ is the constant spatial mean of Z, $\mathbf{1}_{\mathrm{M}}$ and $\mathbf{1}_{\mathrm{N}}$ are M and N vectors of ones, $\mathbf{C}_{\mathrm{pp}}$ is the M×M variance - covariance matrix of $\boldsymbol{Z}_{\mathrm{p}}$, $\mathbf{C}_{\mathrm{BB}}$ is the N×N variance-covariance matrix of $\bar{\boldsymbol{Z}}_{\mathrm{B}}$, $\mathbf{C}_{\mathrm{pB}}$ and $\mathbf{C}_{\mathrm{Bp}}$ are the variance-covariance matrix between $\boldsymbol{Z}_{\mathrm{p}}$ and $\bar{\boldsymbol{Z}}_{\mathrm{B}}$ and vice versa.

Because their joint distribution is normal, the optimal predictor of Z_{p} given $\bar{\boldsymbol{Z}}_{\mathrm{B}}$ is a linear combination of the BSO:

$$\hat{Z}_{\mathrm{p}} = \mu\mathbf{1}_{\mathrm{M}} + \mathbf{C_{pB}}\mathbf{C_{BB}^{-1}}(\bar{\boldsymbol{Z}}_{\mathbf{B}} - \mu\mathbf{1}_{\mathrm{N}}) \quad (2)$$

The variance-covariance matrix of the prediction error, called $\mathbf{C}(Z_{\mathrm{p}} - \hat{Z}_{\mathrm{p}})$, is given by:

$$\mathbf{C}(Z_{\mathrm{p}} - \hat{Z}_{\mathrm{p}}) = \mathbf{C_{pp}} - \mathbf{C_{pB}}\mathbf{C_{BB}^{-1}}\mathbf{C_{pB}^{T}} \quad (3)$$

Why do we need expert knowledge for ATP kriging?

ATP kriging requires that the spatial structure at PoS, which is measured by the variogram, is known. Previous research estimates the PoS variogram only from the BSO using iterative deconvolution process (Pardo-Igúzquiza & Atkinson, 2007; Goovaerts, 2008). The optimisation condition is that the derived PoS variogram is the one that minimised the difference between the theoretically regularised variogram model and the model fitted to the BSO. As a result, several PoS variograms can satisfy this condition, which can cause uncertainty on the disaggregation outcomes. Moreover, the nugget component of the PoS variogram was often dismissed and assumed to be zero as it was thought that the BSO retain little information to infer the nugget component (Nagle et al., 2011).

Bayesian A to P Kriging

Acknowledging the value of experts in the field, we proposed using expert knowledge on the PoS variogram as prior information to constraint the estimation outcomes. We are also aware that expert knowledge can be uncertain, and this uncertainty should be considered. Thus we developed a Bayesian framework for ATP kriging (Truong et al., 2014) because Bayesian approach is a formal framework for combining prior information with observations and can quantify the uncertainty on the inference of the PoS variogram parameters.

How can we extract knowledge from the experts?

Extracting knowledge from the experts can be done by using statistical elicitation techniques, with the expectation of getting the right knowledge. The right knowledge here means the knowledge that reflects the experts' true knowledge and opinions. We have carried out the necessary steps of a formal expert elicitation process as recommended in literature (e.g. Knol et al., 2010 or Kuhnert et al., 2010). These steps are shown in Figure 1.

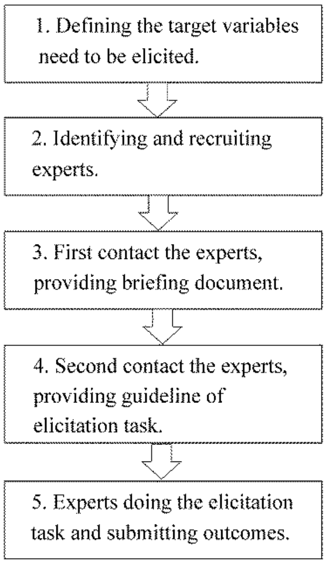


Figure 1: Formal statistical expert elicitation procedure

Our case study involves disaggregating MODIS satellite image of the air temperature over the Gelderland province in The Netherlands (Figure 2) from a 5 km resolution image to 1 km resolution. We invited three experts who have expertise in meteorology and geostatistics to get their opinion on the variogram of the air temperature. The target variables of the elicitation are probability distribution functions (pdf) of the parameters of the Matérn variogram model for PoS. The experts were informed about the case study and the requirements of the elicitation task through a briefing document. Several statistical elicitation techniques have been developed and well documented in literature (O'Hagan et al., 2006). Furthermore, it also has been implemented as web-based tools that are free to access and to use by researchers. In this study, we utilised the MATCH Uncertainty Elicitation Tool (http://optics.eee.nottingham.ac.uk/match/uncertainty.php#).

The outcomes from the elicitation task were the summary statistics (maximum, minimum, mean and standard deviation) of the pdfs of the point support variogram's parameters. The pdfs were fitted to these summary statistics (Figure 3). The last but not least step is combining the three experts' judgments into single judgement by probabilistic averaging, i.e. linear opinion pooling (O'Hagan et al., 2006) of many quantiles generated from the fitted probability distributions of all experts. The averaged quantiles plotted in (Figure 3) were used as informative priors in the Bayesian ATP estimator after they were fitted to a multivariate kernel density.

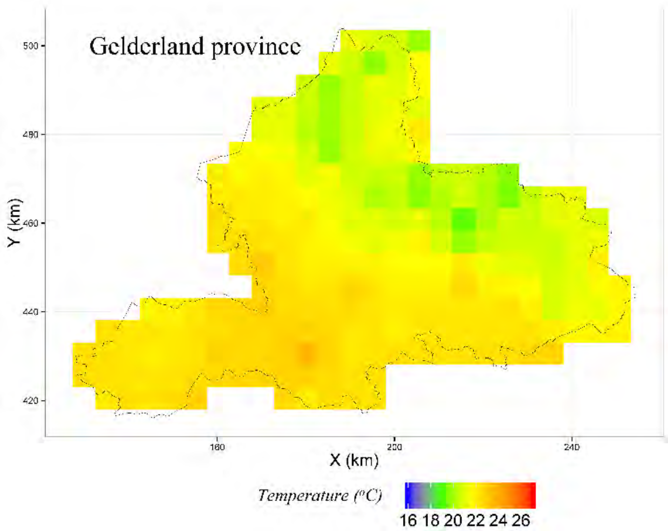


Figure 2: MODIS dataset (MOD07_L2) of Gelderland province

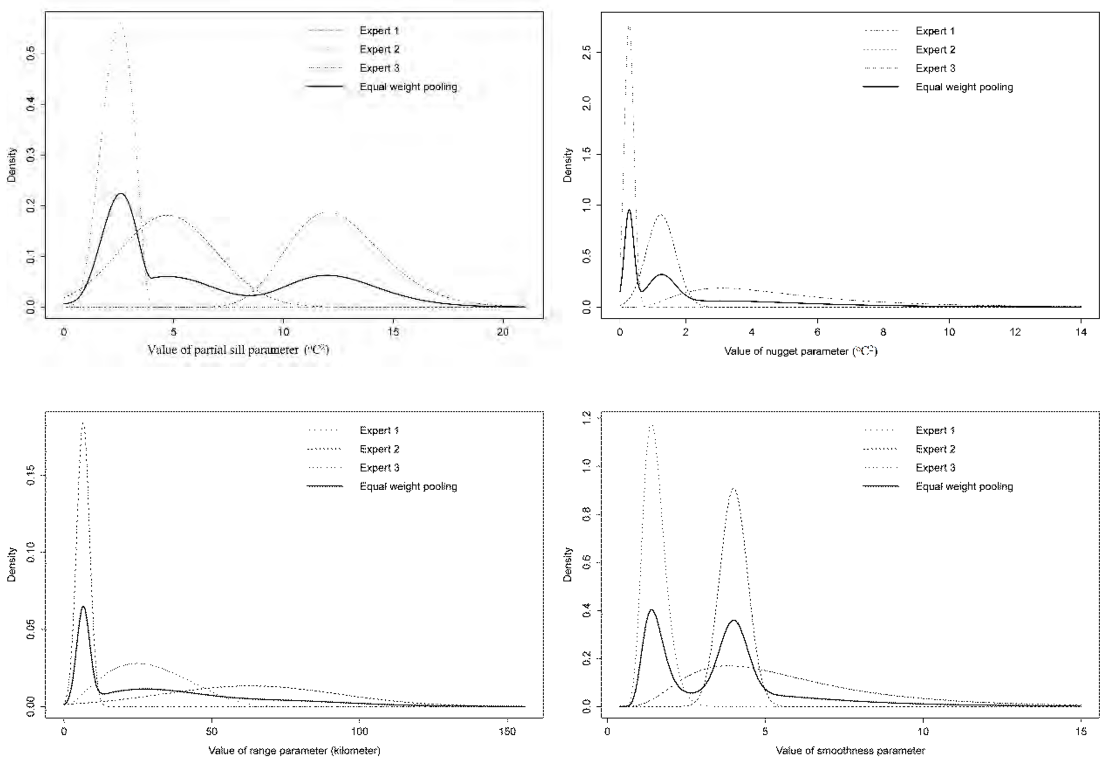


Figure 3: Probability density functions of the PoS variogram's parameters from three experts and their equal weight poolings

Bayesian A2P Kriging

How does the Bayesian ATP kriging work?

Figure 4 shows the three main steps for Bayesian ATP kriging. We have described step 1 as above.

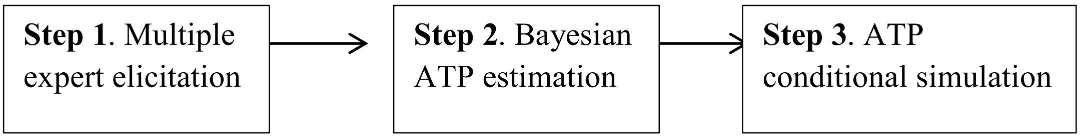


Figure 4: Three main steps of Bayesian area-to-point method

In step 2, the well-known Markov Chain Monte Carlo method was applied, where the 'Metropolis within Gibbs' or hybrid MCMC algorithm that simultaneously uses both Gibbs sampling steps and Metropolis-Hastings steps was used to simulate the posterior distribution of the spatial mean and the joint posterior distribution of the point support variogram's parameters. Recall that the prior distribution for the point support variogram model was the joint distribution of the pooled pdfs of all parameters that was derived by fitting a multivariate kernel density to all pooled pdfs. In step 3, ATP conditional simulation enables to generate realisations of Z at PoS conditional on BSO $\overline{\mathbf{Z}}_{\mathrm{B}}$. We applied the stochastic conditional simulation by first generating unconditional simulation and then conditioning it to BSO by kriging the differences between the BSO and simulated block arithmetic averages (Defouquet, 1994).

The striking result of the Bayesian estimation of the PoS variogram parameters is that the posterior distribution of the nugget parameter is the same as its prior distribution. This result confirms that expert opinion indeed play a significant role here. Without (sufficient) observations at PoS, expert knowledge is the best or perhaps the only source of information available about the nugget effect. Figure 5 shows the three quantiles of the simulation of disaggregation outcomes at 1km resolution compared to its BSO at 5km resolution (Fig. 2).The maps of the 5th (lower limit) and 95th quantile (upper limit) values at each simulation node shows the uncertainty of the disaggregation outcomes.

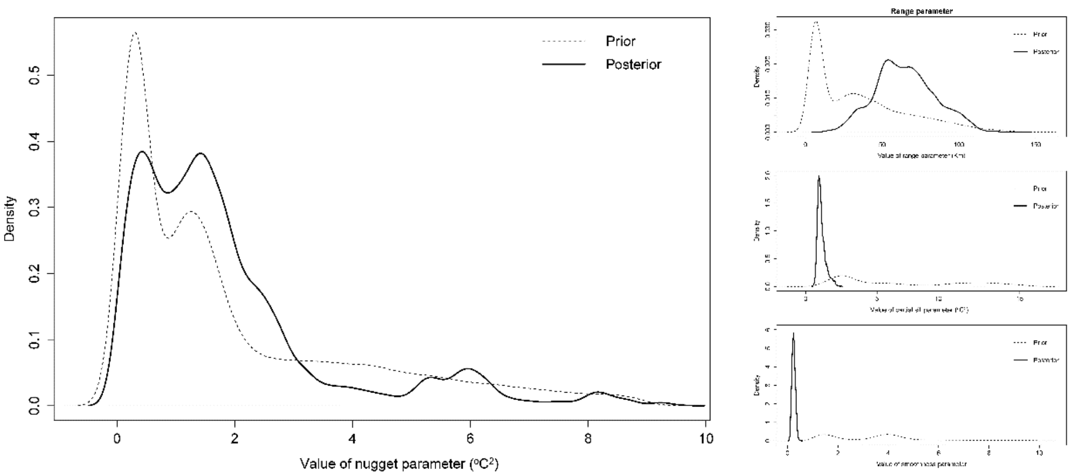


Figure 5: Results of Bayesian estimations of the PoS variogram parameters.

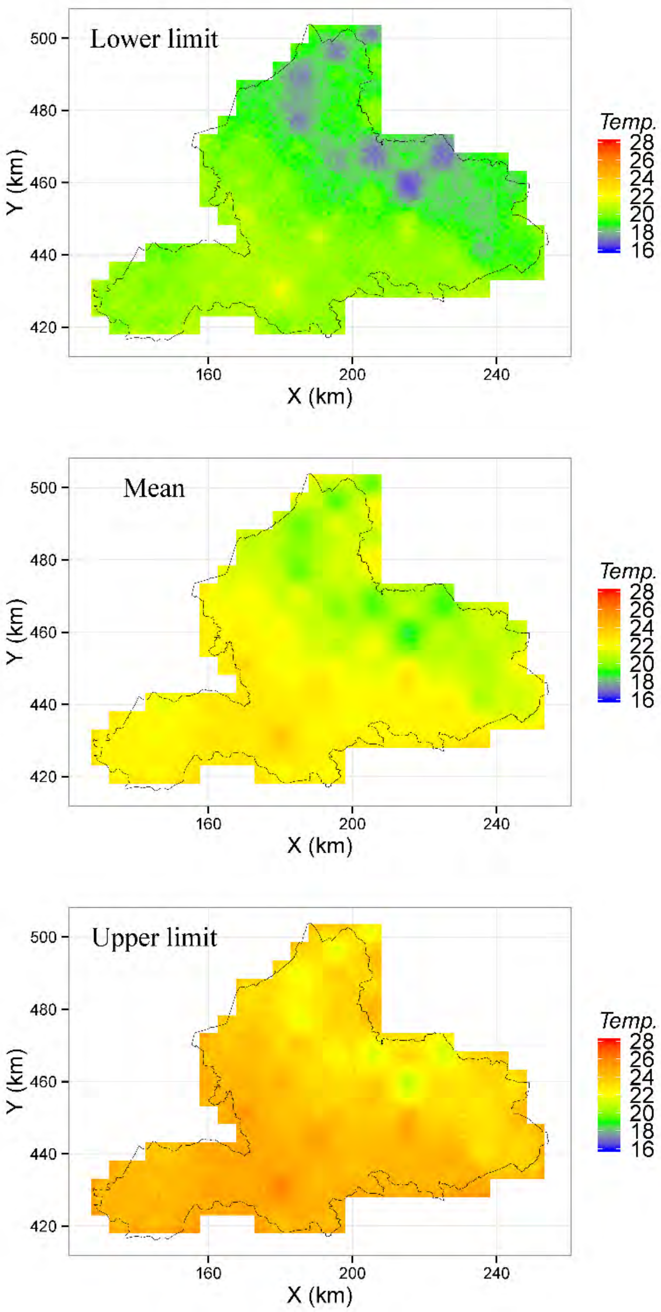


Figure 6: Bayesian ATP conditional simulation outcomes (temperature at 1 km resolution in C°), the 5th (lower limit), mean and 95th quantile (upper limit).

Another interesting result is that the variance of the ATP simulations was much larger than the ATP kriging variance with a fixed PoS variogram that used the modes of the joint posterior of the variogram parameters (Figure 7). This shows that uncertainty about the PoS variogram parameters can make a substantial contribution to the uncertainty of prediction.

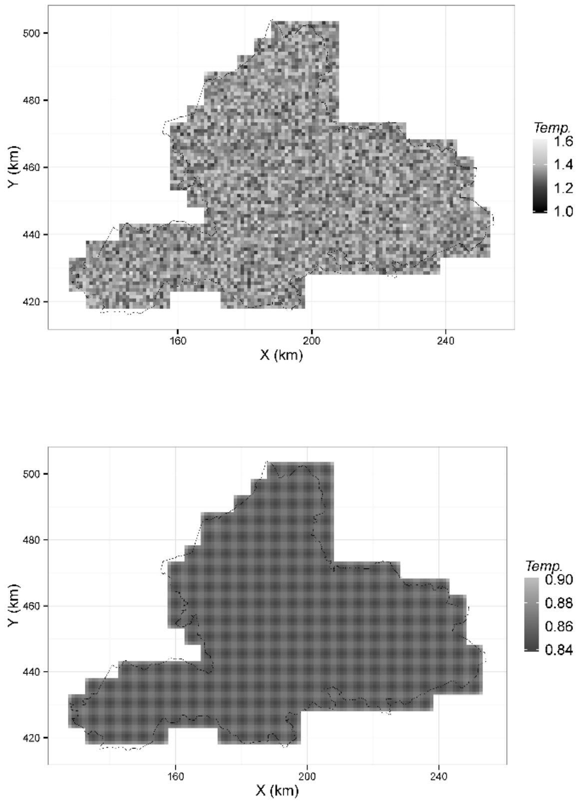


Figure 7: Standard deviation of ATP conditional simulations (top) and that of ATP Kriging with a modal variogram (bottom) (°C)

Summary

We have demonstrated how to use expert knowledge to help us disaggregate spatial information using Area to Point kriging. Although our case study is for air temperature, we believe that this method is readily applicable for soil data as well.

Phuong N. Truong (truongngocphuong@gmail.com)

Phuong holds a MSc degree in Geo-Information science from Wageningen University. She recently obtained her PhD from Wageningen University in June 2014 with a thesis on incorporating expert knowledge in geostatistical inference and prediction. Her research interests are in quantitative analyses for spatio-temporal phenomena using both measurements and expert knowledge.

References

Defouquet, C., 1994. Reminders on the conditioning kriging, in: Armstrong, M., Dowd, P.A. (Eds.), *Geostatistical Simulations: Proceedings of the Geostatistical Simulation Workshop 7*, Fontainebleau, pp. 131-145.

Goovaerts, P., 2008. Kriging and Semivariogram Deconvolution in the Presence of Irregular Geographical Units. *Mathematical Geosciences*, *40*(1), 101–128. doi:10.1007/s11004-007-9129-1.

Knol, A., Slottje, P., van der Sluijs, J., Lebret, E., 2010. The use of expert elicitation in environmental health impact assessment: a seven step procedure. *Environmental Health*, *9*(1), 19.

Kuhnert, P. M., Martin, T. G., Griffiths, S. P., 2010. A guide to eliciting and using expert knowledge in Bayesian ecological models. *Ecology Letters*, *13*(7), 900–914.

Kyriakidis, P. C., 2004. A Geostatistical Framework for Area-to-Point Spatial Interpolation. *Geographical Analysis*, *36*(3), 259–289. doi:10.1111/j.1538-4632.2004.tb01135.x

Malone, B. P., McBratney, A. B., Minasny, B., 2013. Spatial scaling for digital soil mapping. *Soil Science Society of America Journal*, 77(3), 890-902.

Nagle, N. N., Sweeney, S. H., Kyriakidis, P. C., 2011. A Geostatistical Linear Regression Model for Small Area Data. *Geographical Analysis*, *43*(1), 38–60.

O'Hagan, A., Buck, C., Daneshkhah, A., Eiser, J., Garthwaite, P., Jenkinson, D., Rakow, T., 2006. *Uncertain Judgements : Eliciting Experts' Probabilities*. *Statistics in practice* (Vol. 35, p. 321). Chichester : John Wiley & Sons, Ltd.

Pardo-Igúzquiza, E., & Atkinson, P. M., 2007. Modelling the semivariograms and cross-semivariograms required in downscaling cokriging by numerical convolution-deconvolution. *Computers & Geosciences*, *33*(10), 1273–1284.

Truong, P. N., Heuvelink, G. B. M., & Pebesma, E., 2014. Bayesian area-to-point kriging using expert knowledge as informative priors. *International Journal of Applied Earth Observation and Geoinformation*, *30*, 128–138. doi:10.1016/j.jag.2014.01.019

It's the accuracy, stupid

by Gerard Heuvelink

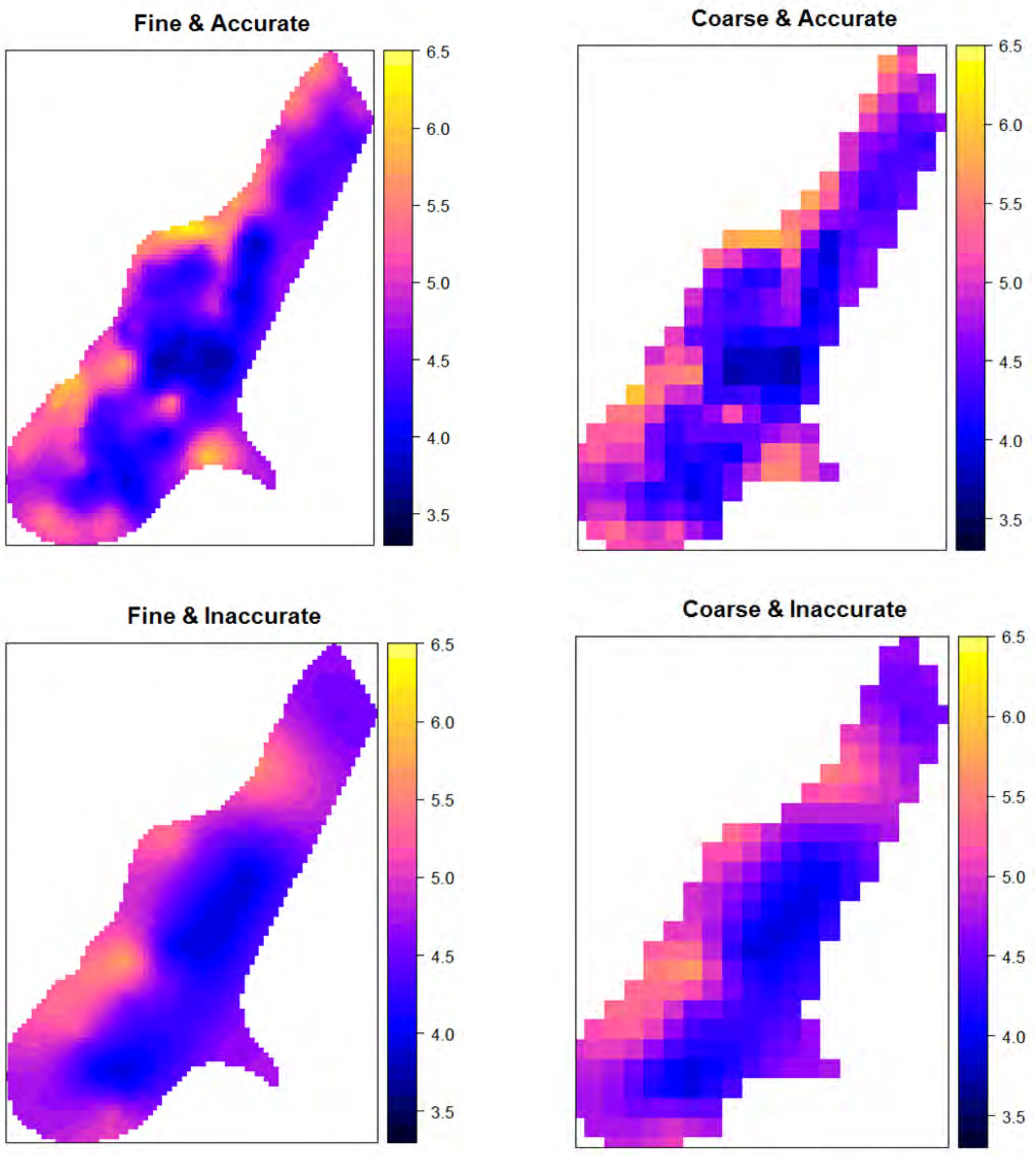
More than twelve years ago I was asked to advise in a debate between a Dutch research institute and one of their clients. The institute had prepared groundwater table maps but the client was not satisfied. The maps had been made using regression kriging and were very different from maps derived using conventional survey. This was one of the client's concerns, but the client also pointed to various places on the map where the spatial patterns and predictions were very wrong. The institute argued that given the available budget, one could not expect more accurate maps. They also mentioned that the client could have known about this because the uncertainty of the spatial predictions had been quantified by means of separate standard deviation maps. The client then took a look at the uncertainty maps and noticed that the prediction intervals were indeed very wide. The client had expected much more accurate maps, and one of the reasons why they expected high thematic accuracy was that in the contract the institute had agreed to deliver the maps at scale 1:10,000. With this, the institute merely referred to the cartographic scale at which printed maps would be completed, while the client interpreted it as a sign of high accuracy. They were used to maps at scale 1:10,000 being much more accurate than maps at scale 1:25,000 or 1:50,000.

In conventional soil mapping, map scale is indeed in effect a measure of map accuracy, because mapping at fine scales involves many more man hours and borings per unit area than mapping at coarse scales. As long as everybody knows and agrees about this it is okay, although it would obviously be better to simply quantify the accuracy of a map directly (for instance, it could have avoided the misunderstanding between the research institute and client mentioned above). It may be implicit that a conventional soil map at scale 1:10,000 has higher (thematic) accuracy than a 1:50,000 scale map, but exactly how accurate and how much more accurate is rarely quantified. So even in conventional soil mapping it is not ideal to interpret map scale as a synonym for map accuracy, but in digital soil mapping it is even more so.

We all know that cartographic scale is no longer relevant in the digital world because we can freely zoom in and out on a digital screen. Indeed, it would be pure coincidence if 1 cm on the screen agrees with 100 m in the real world when a 1:10,000 soil map is displayed on the screen. In digital soil mapping, we therefore rarely refer to the 'scale' of a map (although some hardliners still do!) and use the term 'resolution' instead. That is all very fine, as long as we do not make the same mistake again by confusing resolution with accuracy. Resolution and accuracy are very different things. For instance, a high-resolution soil property map may well have very poor accuracy because of high short-distance spatial variation combined with low sampling density. Also, a low-resolution map may be very accurate because the mapped soil property has little to no short-distance spatial variation, while the sampling density is high and/or there is strong correlation with available covariates.

To illustrate that high-resolution maps can be inaccurate and coarse resolution maps can be accurate, consider the four kriged maps below. The high-accuracy maps were obtained by ordinary kriging of the log-transformed topsoil lead concentration using all 155 observations of the meuse dataset (cross-validation residual variance = 0.16), while the low accuracy maps used only 25 observations, randomly selected from the 155 (cross-validation residual variance = 0.25). The fine resolution map grid cells are 40 by 40 meter, while the coarse resolution map grid cells are 160 by 160 meter (but note that in either case the predictions are at point support: grid cell colours refer to the predicted log-transformed lead concentration at the centre locations of the cell).

It's the accuracy, stupid



Which of these four maps do you prefer? Clearly, this will be the high-resolution, high-accuracy map. I can also imagine that in some cases the high-resolution, low-accuracy map will be considered second-best, although most users should benefit more from the low-resolution high-accuracy map. Coarse resolution and low accuracy is the least attractive.

The points that I am trying to make here are that:

1. we should not confuse spatial resolution with accuracy;
2. producing high-resolution maps is much easier than producing high-accuracy maps.

It’s the accuracy, stupid

Somehow, confusing resolution with accuracy has entered the digital soil mapping community. For instance, the GlobalSoilMap consortium has set as its main goal to map the world at 100 m resolution. This is thought to be a true challenge, while in fact, it is very easy. For instance, in the sp package of R, it is done with a simple parameter statement ‘cellsize = 100’ in the makegrid function. All that needs to be done is to define a fine prediction grid and run the mapping algorithm. Well okay, perhaps ‘very easy’ is not entirely fair. It does not do justice to the many technological problems that have to be solved to run digital soil mapping algorithms for the billions of one hectare pixels on the globe. I have much respect for my colleagues who address these technological challenges and I am impressed by their clever solutions, but these are not pedometric problems. The real challenge is not in increasing the resolution but in increasing the prediction accuracy.

I think that the digital soil mapping community should be much more concerned about the accuracy of the maps that are produced than about their resolution. Luckily, this is now happening in the GlobalSoilMap community, because recently the GSM specifications were extended with accuracy levels (A, AA, AAA). A table defines these accuracy levels for each soil property (for instance, for pH the width of the 90% prediction intervals are 1.5, 1.0 and 0.5 pH unit for the A, AA and AAA accuracy levels, respectively). The idea is that in future any map may receive the A, AA, or AAA quality stamp, depending on how accurate it is.

The accuracy levels table is still in development – led by Dominique Arrouays - and much work needs to be done to complete it and get it accepted and used by the soil science community, but I think it is the right way forward. One of the important issues that we must then also address is how to establish the accuracy of a soil map, because it may not always be acceptable to trust the kriging standard deviation map or derive the accuracy from cross-validation statistics. Best is to collect a sufficiently large, independent validation set using probability sampling. This requires that we encourage our clients that part of the budget reserved for map making is put aside for validation, and also here we still have a lot of convincing to do (not only our clients, but sometimes also ourselves!).

Impact Factor & h index

by Budiman Minasny

Thompson Reuters published its 2013 Journal Impact Factor (IF) last July. As we know, IF is a metricthat is widely used to rank journal's performance, including in soil science. As a reminder the 2013 IF is calculatec from: the number of citations in 2013 to articles published in the journal the past 2 years (2011 and 2012 divided by the number of articles published in 2011 and 2012. It can be interpreted as the average number o citations of an article (published in that journal). At the same time, Google also published itsmetrics called h5-index and h5-median. h5-index is the h-index of the journal for the past 5 years (the 2014 index is data 5 years prior to June 2014). The h-median of a journal is the median of the citation counts in its top cited h articles.

In reality, IF and h5-index are quite correlated (Figure 1. Spearman's Rank Correlation ρ = 0.90). However as a prestige, the order or ranking of a journal can become quite important. We can see some journals tha ranked higher in one metric can be worse in another. E.g. Soil Sci Soc Am J ranked #13 in IF but ranked #6 ir h5-index. A small difference in the IF value (e.g. 0.1) can cause quite a shift in the order (Figure 2).

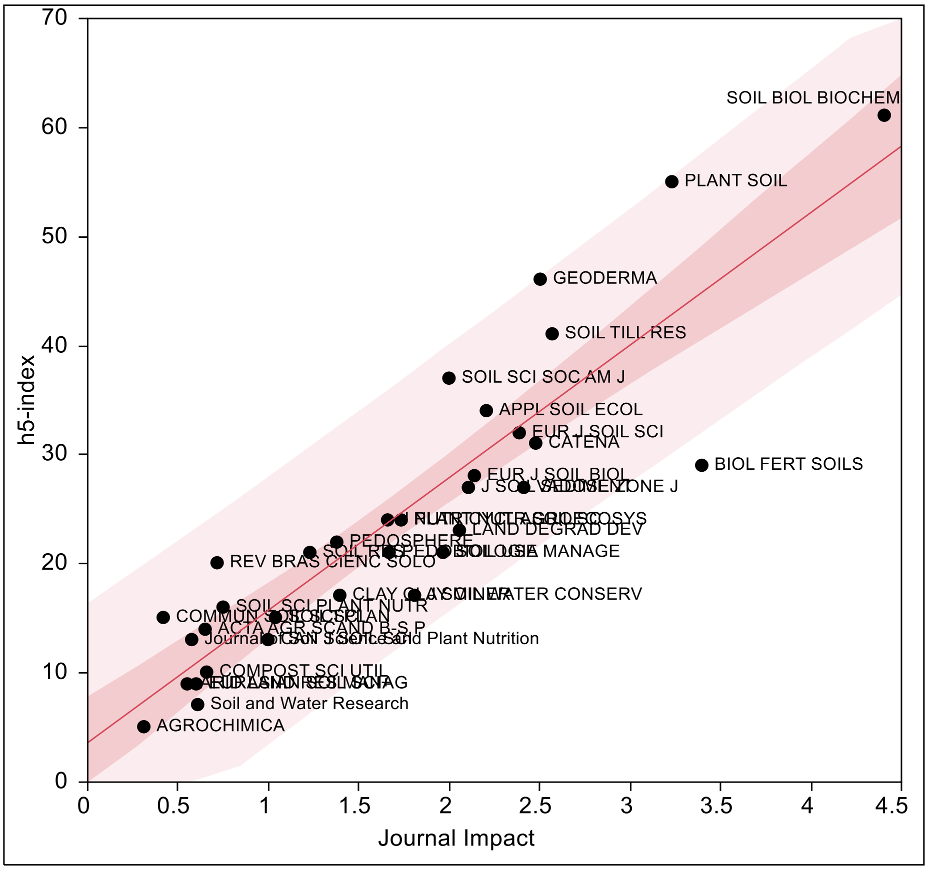


Figure 1. The relationship between journal impact factor and h5-index.

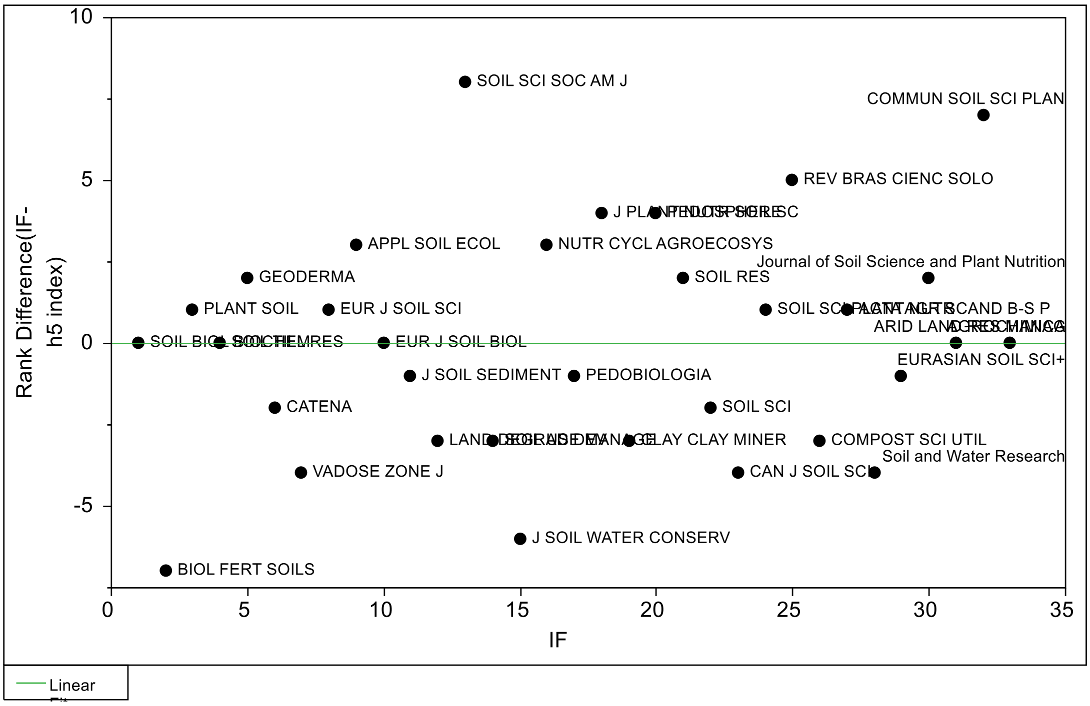
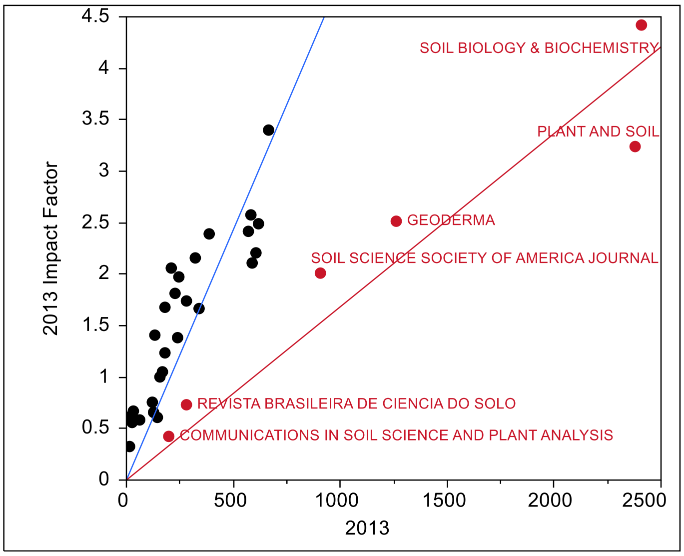


Figure 2. Rank Difference between Impact Factor and h5 index.

We previously showed that the h5-index is a more robust index than IF and it is highly related the number o citations (Minasny et al., 2013). IF, which is calculated with a high precision to 3 decimal place, can be misleading.

As the formula said, Impact Factor is related not only to the number of citations, but also the number of papers. We can see two trends here (Figure 3), for journals that published less than 300 papers in the past : years, the rate of IFs increase with number of citations is about 3 times higher when compared to journal that published more than 300 papers. Thus, an effective way to increase keep a journal's IF high is to keep th number of papers low. And IF only requires few papers to be highly cited. Meanwhile the h-index is only related to the total number of citations.



Journals with < 300 papers Impact Factor = 0.0049*Cites

Journals with > 300 papers Impact Factor = 0.0017*Cites

Figure 3. Number of citations and impact factor.

Impact Factor & h index

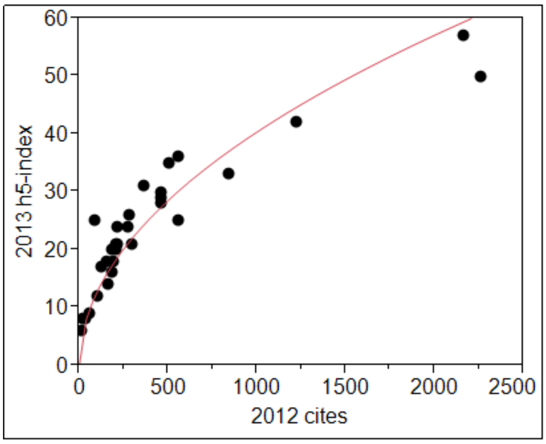
The Square-root function

We compared the 2012 and 2013 IF citation data (from Web of Knowledge) and h5-index for 2013 and 2014 (from Google Scholar). Although Google Scholar and Web of Knowledge are 2 different databases and covered different periods of citations, the Google h5-index has a consistent relationship with the number of cites from Web of Knowledge:

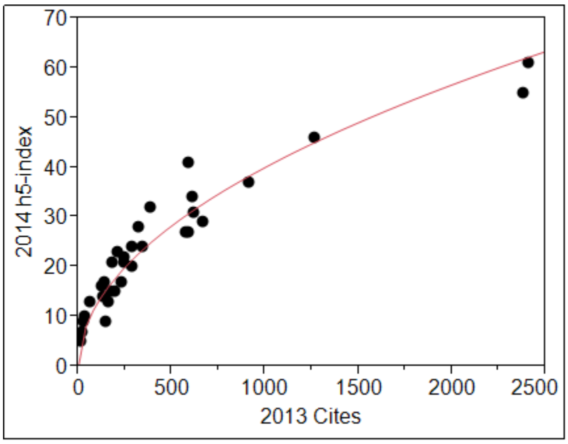
$$\text{h5-index} = 1.27\ \text{Sqrt(Cites)}$$

where cites is the number of citations in that year for papers that were published 2 previous years

This implies that the pattern of a journal's citation and its h5-index is quite predictable. As we may experienced, despite of its impact factor, we all have an idea on the quality of the papers in a journal and the journal's reputation.



h5-index = 1.278 *Sqrt(2012 cites)



h5-index = 1.267 *Sqrt(2013 Cites)

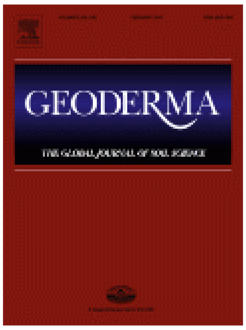
Summary

In general, IF is quite sensitive to the number of papers and citations. A simple way of achieve a high IF is to maintain a low number of papers. The h5-index on the other hand depends mostly depends mostly on citation. Although Google Scholar's citation can sometimes be incorrect, it correlates quite well with Web of Knowledge. The relationship between citations and its h5-index (or ranking) in soil science journals are usually quite predictable, and any annual fluctuations in the Impact Factors should be treated cautiously.

Reference

Minasny, B., Hartemink, A. E., McBratney, A., & Jang, H. J. (2013). Citations and the h index of soil researchers and journals in the Web of Science, Scopus, and Google Scholar. PeerJ, 1, e183.

Geoderma Announces Best Paper Award 2013



Geoderma recently announces its Best Paper Award for 2013. And the winner is a pedometrics paper by Titia Mulder et al. on quantifying mineral abundances in soil using SWIR spectra. This paper was also nominated for the 2013 Pedometrics Best Paper Award. Congratulations to all the authors!

The article is available as Free Article online until December 2015

Mulder, V.L., M. Plötze, S. de Bruin, M.E. Schaepman, C. Mavris, R.F. Kokaly, M. Egli, Quantifying mineral abundances of complex mixtures by coupling spectral deconvolution of SWIR spectra (2.1–2.4 µm) and regression tree analysis, Geoderma, Volumes 207–208, October 2013, Pages 279-290.

And here is the summary of the article:

Soil mineralogy is an important indicator for soil formation and parent material characterization. In environmental and geological studies, the characterization (and quantification) of soil mineralogy is typically achieved using X-ray diffraction (XRD). Visible Near Infrared and Shortwave Infrared (VNIR/SWIR) spectroscopy has proven to be an efficient alternative for the determination of various soil properties. In this paper we propose and demonstrate its use for simultaneous quantification of mineral abundances from complex mixtures.

Detection of minerals having absorption features within the 0.350–2.500 µm range have been successfully obtained using linear spectral unmixing techniques. However, these analyses were limited to estimating the main component within a sample having the most distinct absorption feature. Hence, reflectance spectra of mixtures are typically a complex result from the combinations of the spectral characteristics of the constituents. Depending on the composition, the abundance and the spatial arrangement of the minerals, the total reflectance resulting from the scattering of the minerals within the intimate mixture produces positional shifts, changes in intensity, disappearance of absorption features or changes in their shape.

Hence in this work we aimed to quantify mineral abundances using spectral deconvolution (SD) followed by regression tree analysis (RT). SD involves modelling the total reflectance and the inference of absorption components within complex features by fitting (modified) Gaussian curves to the absorption features and absorption components. Next, mineral abundances were predicted by RT using the parameters of the fitted Gaussians as inputs. The approach was demonstrated on a range of prepared samples with known abundances of kaolinite, dioctahedral mica, smectite, calcite and quartz and on a set of field samples from Morocco.

Cross validation showed that the prepared samples of kaolinite, dioctahedral mica, smectite and calcite were predicted with a root mean square error (RMSE) less than 9 wt%. For the field samples, the RMSE was less than 8 wt% for calcite, dioctahedral mica and kaolinite abundances. Smectite could not be well predicted, which was attributed to spectral variation of the cations within the dioctahedral layered smectites. Substitution of part of the quartz by chlorite at the prediction phase hardly affected the accuracy of the predicted mineral content; this suggests that the method is robust in handling the omission of minerals during the training phase. The degree of expression of absorption components was different between the field sample and the laboratory mixtures. This demonstrates that the method should be calibrated and trained on local samples. Concluding, our method allows the simultaneous quantification of more than two minerals within a complex mixture and thereby enhances the perspectives of spectral analysis for mineral abundances.

Upcoming Meetings

- Spring School on **Soil Mapping and Soil Assessment**, ISRIC – World Soil Information, Wageningen, The Netherlands, May 18-22, 2015. http://www.isric.org/content/isric-spring-school-2015

- **Communication of uncertain information in the earth sciences** at the EGU General Assembly, 12th to 17th April 2015 in Vienna, Austria. The deadline for abstract submission is 7th January 2014, and abstracts may be submitted at http://meetingorganizer.copernicus.org/EGU2015/sessionprogramme

- **Large scale digital soil mapping: challenges and opportunities in delivering high-resolution soil information** at the EGU General Assembly, Vienna. More information can be found at www.egu2015.eu, through the Programme Group Soil System Sciences (SSS) under – Material and Methods in Soil Science (SSS12). The deadline for abstract submission is 7 January 2015. Organised by: Titia Mulder, Bas Kempen, Dominique Arrouays and Luca Montanarella

- **Proximal sensing of soils – spectroscopy, morphometrics and other observational tools** at the EGU General Assembly on 12-17 April 2015. Organised by Matt Aitkenhead.

- **Global Soil Security Symposium**, Texas A&M University, College Station, Texas, May 19-21 2015. https://globalsoilsecurity.tamu.edu/index.html

- **The 4th Global Workshop on Proximal Soil Sensing**, Sensing Soil Conditions and Functions, Hang Zhou, China, May 12-15 2015. http://www.gwpss2015.com/

- **Global Workshop on Digital Soil Morphometrics**, University of Wisconsin, Madison, June 1-4, 2015. http://digitalsoilmorphometrics.org/

Session C1.5-1 at the 20th WCSS: Validation of Soil Carbon Sequestration

by Sabine Grunwald

The theme of the WCSS in Jeju, Korea (June 8-13, 2014) was “Soils Embrace Life and Universe”. Indeed the Congress embraced the breath of soils applications and research cutting across spatial and temporal scales, geographic settings around the world, and a variety of themes featured by Divisions, Commissions, and Working Groups of the International Union of Soil Science (IUSS). Jeju Island features the 7 wonders of nature – one more spectacular than the other with a wide variety of volcanic topography and geologic formations, among them the UNESCO Biosphere Reserve, UNESCO World Natural Heritage, UNESCO Global Geopark – e.g. Mt. Hallasan, Youngcheon, Daepo Columnar Joint Lava, Seogwipo shellfish fossil layers, and Cheonjiyeon Waterfall. The organizing ceremony of the WCSS featured a striking sand art life presentation – worth to watch! Go to: Opening Ceremony – sand art performance: http://www.20wcss.org/xe/movie/463475 or visit ‘Let it go (sand art)’.

The Pedometrics Commission 1.5 IUSS organized the session “Validation of Soil Carbon Sequestration” with the invited keynote talk “Spatial Stratification in Design-based Sampling for Soil Carbon Auditing” given by Jaap de Gruijter, Alex McBratney and Budiman Minasny. Other talks in the session provided a journey to different geographic settings across the world (U.S., South Africa, Brazil, and Philippines) where soil carbon has been studied using a variety of approaches considering relationships with environmental co-variates, specifically land use/land cover and climate. The oral session sponsored by the Pedometrics Commission 1.5 “Quantification and Application of Uncertainty in Pedometrics” featured Keith Shephard as keynote speaker talking about “How much Soil Spatial Information do We Need to Address Critical Uncertainties in Development Decisions”. His talk was complemented by speakers addressing uncertainty from different perspectives (non-probabilistic uncertainty assessment, epistemic uncertainty, uncertainty directed digital soil mapping (DSM), and Bayesian geostatistical modeling). The Working Group Digital Soil Mapping organized the session “Progress in Digital Soil Mapping and Global Soil Map” representing the state-of-the art in DSM. And Working Group Proximal Soil Sensing organized the session Proximal Soil Sensing featuring sensing methods applied in soil mapping applications. Besides the oral sessions numerous posters focused on pedometrics and DSM themes providing an exciting spectrum of applications. The interdisciplinary Symposiums “Soil Security”, “Soil-Plant Welfares for Human”, and “IUSS for Global Soils: Future Nexus” provided a platform for reflection, discussion, and engagement in contemporary hot topics reminding us what role soils play within the larger global context. In totality, it was an exciting WCSS in a beautiful setting, welcoming and ever smiling and cheerful hosts, a high-quality scientific program and technical tours, a rich cultural program integrating soil science and art that made the congress a memorable experience.

Session C1.5-2 at the 20th WCSS: Quantification and Application of Uncertainty in Pedometrics .

by Lin Yang

The session was held with oral and poster presentations chaired by Prof. A-Xing Zhu on June 10 during the 20th WCSS. The development of pedometric techniques has vastly improved our ability to map soil information digitally at various scales. One of key research topics is the provision of and the application of uncertainty information associated with the information products from these techniques. The purpose of this session was to bring together scientists involved in the quantification and application of uncertainty to exchange research findings and to create a synergy on the topic.

Five colleagues from Kenya, Australia, USA, and China contributed to this session with their studies ranging from the conceptual to the applied. Dr. Xiong Xiong from University of Florida presented a Bayesian geostatistical approach to characterize the spatial variability of SOC with uncertainty assessment, and Prof. Lin Yang from Chinese Academy of Sciences presented a non-probabilistic approach to estimate prediction uncertainty with sparse ad hoc samples for Jing Liu from University of Wisconsin-Madison, who couldn't attend the conference because of visa problem. Prof. A-Xing Zhu, the former chair of Pedometrics Commission talked about the uncertainty-directed digital soil mapping. Keith D. Shepherd, a Principal Scientist and the Science Domain Leader of Land Health Decisions department in World Agroforestry Centre, ICRAF, talked about how to build up a bridge between gathering soil data and improving development decisions by uncertainty reduction. A review of data requirements for stakeholders in African agriculture revealed a lack of alignment between perceived data needs, uncertainties, and data gathering efforts. He and his colleagues thus proposed a decision analytic approach whereby researchers first model the uncertainty in key decisions they seek to improve before designing measurements. Also David Rees from Australia talked about changing epistemic uncertainties in soil classification and digital soil mapping.

It was a stimulating and interesting Symposium. Colleagues from different countries asked interesting questions and discussed not only on details of the presented uncertainty estimation methods but also the effects of using uncertainty in soil mapping and other applications.

Nine posters were presented in this session dealing with several issues including:

- Uncertainty assessment in soil mapping (Laura Poggio, The James Hutton Institute, United Kingdom, Spatial uncertainty in 3D modelling of soil properties; Joulia Meshalkina, Moscow Lomonosov State University, Russia, uncertainties Assessment of Semivariogram Parameters and Maps Comparison for Soil Properties with Different Nugget Effects),
- Sampling design (Abdur Rab, Victorian Government Department of Environment, Australia, Sampling design and the predictive accuracy of pedotransfer functions; Lin Yang, Institute of Geographical Sciences and Natural Resources Research, Chinese Academy of Sciences, A multi-grade representative sampling strategy using auxiliary environmental variables for regional soil mapping: a case study in Anhui, China),
- Digital soil mapping (Thomas Orton, University of Sydney, Australia, Analysis of the spatial and depth-wise variation of soil properties based on horizon-sampled data; Ricardo Simao DinizvDalmolin, Universidade Federal De Santa Maria, Brazil, Prediction of soil organic carbon and texture in complex areas using Vis-NIR spectroscopy; Qianlong Wang, Zhejiang University, China, Grey incidence analysis (GIA): A new local method for modeling Chinese soil vis-NIR spectral library to predict soil total nitrogen; Istvan Sisak, Univer-

sity of Pannonia Georgikon Faculty, Hungary, Probability-based harmonization of digital maps to produce conceptual soil maps),

- Digital soil mapping prototype system (Jingchao Jiang, Institute of Geographical Sciences and Natural Resources Research, Chinese Academy of Sciences, CyberSoLIM: An Easy and Fast Online Digital Soil Mapping Prototype System).

Participants held heated discussions on the above issues. The CyberSoLIM, which is an easy and fast online digital soil mapping prototype system, raised great interests. The best poster presentation was awarded to Jingchao Jiang after voted by several delegates chosen from the participants.

This session provided a good opportunity for the international Pedometrics (including Digital Soil Mapping and Assessment) community to meet and exchange ideas, research, and perspectives on how to quantify the associated uncertainty and actually use uncertainty in improving development decisions. The Jeju Island served as an ideal host, promoting the international research and friendship of the globally oriented Pedometrics community. Expectations are high for the Pedometrics 2015 in Córdoba, Spain. See you there!

6th Global Workshop on Digital Soil Mapping (Nov.11-14, 2014) Nanjing, China

by Titia Mulder

This workshop was organized to bring together scientists working in the field of digital soil mapping and related disciplines. The organization was in the hands of the State Key Laboratory of Soil and Sustainable Agriculture, Institute of Soil Science, Chinese Academy of Sciences. Prof. Zhang Ganlin and his colleagues were excellent hosts and their hospitality was highly appreciated. The participation of the workshop was successful, considering the contributions of 120 attendees originating from 15 countries from all continents, having 58 talks and 17 posters.

The major theme of the workshop was "digital soil mapping across paradigms, scales and boundaries", addressing digital soil mapping, -modelling, -applications and -resource inventory. Also, recent findings were presented on the use of legacy data, soil sampling, covariates, soil spectroscopy and 3D modelling in DSM. Especially, sampling and the uncertainty assessment of DSM products was found to be a major topic, which should be accounted for in future research. The challenging theme of this workshop provided the base for interesting discussions about the perspectives and prospects of soil mapping, including the present and future prospects of e.g. GlobalSoilMap and SoilGrids1km.

Highlights:

- Over the years, the discipline of DSM has substantially matured and we have reached a major consensus concerning suitable mapping and modelling procedures.
- Future research may focus on the improved use of covariate data and appropriate use of legacy or sampling data, in such a way that it suits the scientific objective.
- In addition to the frequently used data mining techniques, alternative modelling approaches slowly emerge into DSM. At this workshop, they demonstrated their efficiency in sampling strategies, modelling soil properties and related uncertainties for DSM. Examples of such approaches are Gaussian simulation and Bayesian modelling, among others.
- Advanced methods for efficient and effective data collection were presented. In the future, we may see an increase in soil and environmental information delivered by community-based data collection, through the use of e.g. special applications on mobile telephones.

Fieldtrip:

On the 3rd day of the conference there was a field trip scheduled, visiting the Jiangsu province in the Yangtze river delta. On this trip, we had the opportunity to study two different paddy soils, classified as an Albic Fe-accumuli-Stagnic Anthrosol and a Calcaric Ochri-Aquic Cambosol. Paddy soils are characterized by prolonged and intensive human activities. In China, the human influence involves rice cultivation, already for more than 7000 years. In addition, the hydrothermal regime of the soil is usually subject to change by land levelling or terracing, and the alteration of oxidation and reduction which takes place is due to seasonal irrigation. The Albic Fe-accumuli-Stagnic Anthrosol was located in the rice fields at the FACE (Free-Air CO2 Enrichment) research site. The Rice/Wheat FACE is a platform used to study the response of a rice/wheat crop rotation system to enriched CO2. This experiment allows assessing changes in the agro-ecosystem to climatic changes. The beautiful weather added to the already good spirit of the group and fuelled the discussions around the soil pits. The trip was closed with a delicious dinner and tour around a beautiful ancient street in Yangzhou city.

Outlook:

It may have become time to find ways to integrate DSM into a wider context, as was indicated by various keynote presentations. Recent advances indicate that DSM is slowly shifting into a new phase, where we use a high diversity of methods and approaches for integrating data and knowledge and for understanding about soil and ecosystems and soil functions. Thereby, DSM is becoming more and more inter- and multidisciplinary, which gives DSM challenging research prospects. Finally, the scientific output of this workshop shows that we are indeed crossing paradigms, scales and boundaries. The coming years we will be able to further develop DSM, employing the emerging models and prospects for DSM. Since advances in DSM go quickly, we do need workshops and conferences for exchanging our findings. Therefore, we want to welcome everybody to present their work at the 7th Global Workshop on Digital Soil Mapping, which will most likely take place in Aarhus, Denmark, 2016.

Welcome Address by Ren-Fang Shen

Groups photo

Albic Fe-accumuli-Stagnic Anthrosol

Discussion around the soil pit at the FACE research site

FACE research site

Dinner in the center of Yangzhou

P.H.T. Beckett. 1928–2014.

by Murray Lark and Richard Webster

Philip Beckett, who died earlier this year, was formerly lecturer in soil science at Oxford University and fellow of St Cross College there. He had wide interests, and in one respect sparked a revolution in pedological thought and application from which pedometrics emerged.

Philip read chemistry at Oxford, but his principal enthusiasm as an undergraduate was for exploration. In 1947 he was chemist on a University expedition to Iceland. By his account this experience taught him that the expedition chemist's life was one of drudgery: analysing samples while colleagues from other disciplines monopolized exploration of the landscape and also expected the chemist to do the cooking. For this reason he equipped himself with a copy of The Study of Soil in the Field by G.R. Clarke, then University Reader in Soil Science, and led the next expedition as a soil surveyor.

That expedition was to Kerman, Iran, and is documented by Anthony Smith in the book Blind White Fish in Persia (1952). Smith wrote: 'Philip dug holes wide and deep, made a cloud of dust and wrote lengthy notes in his book.' In due course Philip's findings on the soil were published in the Journal of Soil Science. Other papers described the historical geography of the region, its climate, agriculture, the distribution of blood groups among the population and the qanats (underground tunnels built to channel water from upland aquifers).

After graduation Philip began research under Walter Russell's supervision. On Russell's departure to East Africa, however, he was left without a supervisor and was called up for national service in the army's Intelligence Corps. One of his tasks was to investigate the bogging of British armour in North Germany towards the end of the second World War: why had tanks sunk into what was expected to be firm ground? The reason was that commanders were relying on inference drawn from geological maps without knowing what lay on top of the rocks; they did not know in advance what the soil was like.

Philip returned to Oxford to continue research, but now supervised by R.K. Schofield and on a fresh topic, namely the thermodynamics of exchange reactions of potassium in soil. His papers on that work in the Journal of Soil Science brought him world-wide recognition. He was also appointed to the university staff.

In an article in New Scientist magazine[1] Philip wrote 'In research, as in life, the most seminal ideas often arise before the mind and imagination have settled into a rut'. He took this principle seriously, resisting narrow specialization. So, while he continued research on exchange processes he was ready to apply himself as a scientist to any interesting problem.

One such came to light because of his ongoing attachment as a Territorial officer to a unit of the British Army's Royal Engineers in which he rose to the rank of major.

Soft going on the Oxford Clay: Philip Beckett, left, with geologists of the Royal Engineers

That unit identified the need for information on the soil to predict going conditions for vehicles and for makeshift roads and airfields, and it persuaded the government to fund research into the matter for both military and civil purposes. The received wisdom at the time was that all could be achieved with conventional soil maps. Philip was unconvinced, and for two reasons: (a) conventional mapping was too slow, and (b) it did not and almost certainly could not represent all the variation present. The first problem could be addressed by physiographic mapping from air photography, which Philip and his team demonstrated with land-system atlases of Uganda, Swaziland and western Kenya. The second could be placed on a statistical footing so that predictions could be accompanied by assessments of uncertainty. That was the revolutionary spark, and it led to what we now know as pedometrics.

This new branch of research was developed by a series of doctoral students. The first of these was Richard Webster who assessed the predictive value of land facets delineated on photographs by estimating and comparing the within- and between-facet components of variance of soil properties . Beckett and Webster also wrote a review of the literature on soil variability which set out much of the agenda for subsequent work[2]. Peter Burrough examined multivariate methods for soil classification and survey and studied the costs and utility of soil maps made by different methods and intensities of survey in the Vale of the White Horse. Stein Bie and Philip, with Philip's family and Stein's wife, undertook a year-long tour of Australia to examine how soil surveys were made and used. They went on to publish what may be the first paper in a soil science journal to cite Matheron on geostatistics[3]. Raul Ponce Hernandez and Jesus Viloria, from Mexico and Venezuela respectively, worked on the use of geostatistics in soil information systems, and a paper with Francis Marriott from Oxford's statistics department launched the still live question of how to reconstruct continuous models of soil variation down the profile from data on mean values for successive horizons[4]. Philip's last student was Murray Lark, who worked on statistics of remote sensor data and examined the spatial variations of categorical soil information. This led to Philip's one nomination for the Pedometrics Commission's "best paper" prize[5].

Philip was also editor in chief of a series of twelve books, Monographs on Soil Survey, which was published by Oxford University Press between 1977 and 1986 and in which an international authorship wrote authoritatively about, inter alia, statistics, soil classification, land evaluation, soil description, geographical information systems and soil survey for forestry and engineering.

Philip down below taking a monolith in the Oxford Clay at Wytham

In the mid 1970s Philip turned his attention to the problems created when sewage sludge containing heavy metals is applied to land. He realized that the regulations then in place in different countries implied different mathematical forms for the joint effects of metals on the soil system: a set of ready-made hypotheses. Along with his students he also examined the critical concentrations of heavy metals in plant and animal tissue, and the chemistry of heavy metals in digesters, soil organic matter and the soil solution. A characteristic innovation, when computer graphics were still rudimentary, was a cube formed of layers of perspex, on which it was possible to plot a joint response surface to two variables for examination in three dimensions.

Although exchange processes, soil variability and heavy metals were the three principal themes of Philip's research they by no means exhaust his output. An article on Walther Penck's Aufbereitung concept, given a pedological interpretation, hints at ideas for developing a basic theory of soil in the landscape which has yet to be realized[6]. He also wrote about the logistics of agricultural extension, yield variation, land values and the aesthetic appreciation of landscapes.

Philip in retirement

Philip Beckett retired in 1995. He was the last of the soil scientists sensu stricto in the university, and to mark the occasion he held in his college a wake attended by many of his former students and colleagues. We remember him with affection for his guidance early in our careers, his good humour, stimulating discussions and care for our well-being.

References

(1) Beckett, P.H.T., Bie, S.W. 1972. Diminishing returns in research. New Scientist, 56, 517–519.

(2) Beckett, P.H.T., Webster, R. 1971. Soil variability: a review. Soils and Fertilizers, 34, 1–15.

(3) Beckett, P.H.T., Bie, S.W. 1976. Reconnaissance for soil survey. II. Pre-survey estimates of the intricacy of the soil pattern. Journal of Soil Science, 27, 101–110.

(4) Ponce-Hernandez, R., Marriott, F.H.C., Beckett, P.H.T. 1986. An improved method for reconstructing a soil profile from analyses of a small number of samples. Journal of Soil Science, 37, 455–467.

(5) Lark, R.M., Beckett, P.H.T. 1998. A geostatistical descriptor of the spatial distribution of soil classes, and its use in predicting the purity of possible soil map units. Geoderma, 83, 243–267.

(6) Beckett P. H. T. 1968. Soil formation and slope development; 1, A new look at Walther Penck's Aufbereitung concept. Zeitschrift für Geomorphologie, 12, 1–24.

From Soil Change Matters Workshop,

March 24-27, 2014, Bendigo, Australia