Moving away from the geostatistical lamppost: Why, where, and how does the spatial heterogeneity of soils matter?

Philippe C. Baveye, Magdeline Laba

Abstract

Since the late 1970s, thousands of scholarly articles, books and reports have dealt with the application of the mathematical theory of geostatistics to characterize the spatial “variability” of soils, and to produce soil property maps. Insensibly, this application of geostatistics appears to have become an end in itself, and the reasons why one should be concerned about the spatial heterogeneity of soil properties are rarely if ever made clear any more. In this context, the purpose of the present review article is to return to some of the primal questions that motivated this interest in the topic several decades ago. After a brief review of the background behind the application of geostatistics to soils, a number of situations and modeling efforts are described where, even though soils undoubtedly vary spatially, nothing seems to be gained practically by explicitly accounting for their spatial heterogeneity in order to reach a number of management or research objectives. Contrasted, whenever the spatial heterogeneity of soil properties in the field might be relevant, it is shown that very different perceptions about it emerge, depending on the type of measurement that is performed. This suggests that the approach one adopts to characterize spatially-varying soil properties should be dictated by whatever goal one pursues. For example, if the objective is to evaluate the “ecosystem services” of soils in a given region and to reach decisions about them, one should probably first consider the (typically large) spatial scale that is most relevant to the decision-making process, then proceed via a top-down approach to characterize the spatial heterogeneity of soil services, if and when appropriate. In other contexts, it is argued that measurements should be patterned after the behavior of plants or microbes present in soils, relative to which, unfortunately, the macroscopic measurements that are now routinely carried out appear largely irrelevant or misleading. The article concludes with a number of potential lessons learned from the analysis of the research on the spatial heterogeneity of soils, which bear relevance to the broader practice of soil science.

© 2014 Elsevier B.V. All rights reserved.

1. Introduction

In the late 70s, the publication by Journel and Huijbregts (1978) of their landmark textbook on geostatistics greatly contributed to popularize the work carried out earlier by Matern in Sweden, Gandin in the Soviet Union, Krige in South Africa, and Math-eron in France, and described in detail by the latter (Matheron, 1962, 1965), on the statistical analysis of the spatially-varying properties of meteorological events, forests, and geological formations. Soon thereafter, geostatistical techniques were adopted enthusiastically by hordes of ecologists and soil scientists, eager to generate maps of soil properties of interest within agricultural fields, watersheds, ecoregions, or even entire countries. Research in the area took off like a brush fire, rapidly gathering tremendous momentum. According to Google Scholar, an astounding total of 33,700 articles, book chapters, and reports have been written so far on the use of geostatistics to quantify in some way what has come to be described as the spatial “variability” of soils but should probably be referred to more appropriately as their spatial variation or, like in the following, as the spatial heterogeneity of their properties. The Web of Knowledge (Thomson Reuters, Princeton, New Jersey) lists a less impressive, but still commendable, number of about 12,260 articles that use “soils” and either “geostatistics”, “kriging”, or “kried” as descriptive keywords. Judging from the rate at which articles on the topic are being published at present – more than 3000 in 2013 according to Google Scholar – the brush fire...
still appears to burn unabated, more than 35 years after the initial spark.

In view of how hugely fashionable the application of geostatistics has obviously been in the study of soils during the past three decades, one is unavoidably reminded of the very astute observations made by Varn and Bromley (1994) in a context that is slightly different but nonetheless so remarkably relevant that part of their text is worth repeating here in extenso: “The history of science warns us that the mere popularity of a particular epistemological program is not sufficient evidence of its truth content. Nor is popularity a sufficient guarantee that those in a shared pursuit will not lose sight of the larger issues at hand. Indeed, it could very well be that the considerable popularity of a particular research program serves, in a perverse way, to reduce the probability that its ultimate purpose will be kept firmly in view. The very popularity of the research program then becomes self-reinforcing and serves both to envelope an ever larger share of those who might otherwise follow different research programs, and to stifle dissent out of fear of being thought out of the very broad and encompassing ‘mainstream.’ Meanwhile, the research becomes ever more inviolated, and it becomes easier to lose sight of why one began the journey in the first instance. If we may be permitted a nautical metaphor, a long series of technically perfect tackling maneuvers may very well deposit the fastidious crew at a destination quite devoid of virtue.”

Depending on one’s perspective on the merits, or lack thereof, of trends and fashions in scientific research, one can look at this incisive assessment by Varn and Bromley (1994) as either cynical or insightful when it comes to the research involving geostatistics in soil science. Regardless of which description applies, it is of interest to inquire whether the reasons why “one began the journey in the first instance” are still clearly in sight in this field. Evidence suggests that they are not, and that the situation is reminiscent of the popular joke (apparently inspired by the Scottish poet Andrew Lang) about a drunkard looking for his car keys not where he is certain he lost them, but under a lamppost because, as he says, “there is more light here and I can hang on to the lamppost.” Indeed, from the multitude of articles devoted in recent years to what has become termed by some, symptomatically, “soil geostatistics” (e.g., Lark, 2012), a strong impression emerges that the use of geostatistics has in most cases become an end in itself, relegating to the very distant – and obscure – background some of the original questions about soil spatial heterogeneity that may have motivated the authors’ interest in geostatistics in the first place.

In this context, the key objective of the present critical review article is to return to some of these primal questions about the significance of the spatial heterogeneity of soil properties, and about whether it is imperative to be able to describe it quantitatively, using any one of an array of available theoretical frameworks. A useful first step in this analysis is to outline the historical background of the work on the application of geostatistics to soils, in order to better understand what contributed to its being framed the way it has from the outset. Then, we propose a quick overview of a number of situations where spatial variation, even though it is undoubtedly present, is not necessarily relevant to our attempts to describe soil-related processes, or at least does not have to be taken into account explicitly in the manner it has been accounted for in the last few decades. The next step in the analysis deals with situations where local measurements of a spatially-varying soil parameter are obtained and interpolated to address specific purposes, and where it is shown that the volume of soil over which local measurements are made influences significantly the perception one gets of the spatial variation of soils. In that context, the key question, which does not appear to have been asked very much at all in the soil geostatistics literature, is that of determining which measurements are appropriate, at what spatial scale. This question is addressed in the subsequent section in terms respectively of the topical assessment of the services soils provide to human societies, and then of the response of plants and microbes to spatially-varying soil properties. The article concludes with a quick overview and discussion of some of the lessons that the past three decades of research on spatial heterogeneity have taught us. These lessons should guide us in the future, as we finally address some aspects of the topic that have been neglected so far, and they may be applicable as well to other areas of the study of soils.

2. Historical background

Accounts of when exactly humans began to grapple with the spatial variation of soils are lost in the night of times. It is likely that already centuries if not millennia ago, peasants walking alongside their oxen or horses as they worked their fields, or when weeding them by hand, developed a very good feel for the differences among the properties of soils at different locations. In some cases, like Belgium and the Netherlands, this peasant knowledge apparently extended to fatal livestock diseases, e.g., scrapies or “sway-back” (Charlet et al., 2012), associated with individual parcels of land, on which farmers systematically avoided to establish pastures (Voisin, 1959; Joseph Baveye, personal communication, 2002). One likely, but hardly researched, outcome of such daily soil observation over many generations is that it determined the actual size and configuration of agricultural fields to maximize their uniformity, which in turn facilitated their cultivation (Oliver, 2010). Another outcome, much better documented, is that most advanced societies, even those that never developed an alphabet or script, came up with some form of indigenous system of soil classification (e.g., Sandor and Furbee, 1996; Barrera-Bassols and Zinck, 2003; Sampietro Vattuone et al., 2008), and tried early on to understand what made some soils behave very differently than others (Baveye, 2013). Eventually, starting in Russia in the 19th century, this preoccupation led to the launching of extensive soil survey programs around the globe, which have perured to this day.

Agricultural mechanization, especially the adoption of tractors and harvesters with internal combustion engines, and the progressively increasing reliance on fertilizers in the late 19th and early 20th century, changed completely the agricultural landscape and the perception of the spatial variation of soils. Farmers were encouraged and occasionally coerced to consolidate small fields into larger units, which it was considered economical and time-efficient to treat as uniform units. Somewhere along the lines of this evolution, the spatial variation of soils at the field scale became entirely ignored, to the point that many farmers, when they were attempting to get samples analyzed to determine what amendments were needed, often took only one sample per field, regardless of its extent. Some researchers tried early on to argue that this practice was not defendable. For example, Kelley (1922), in Riverside (California), conducted a study where he laid out transects, along which samples were drawn at consistent intervals and were then analyzed for various chemical components. He concluded that “the analysis of a single soil sample drawn from one place within the area studied, has very little value. […] one or more samples from each of several of the experimental points contained practically no alkali salts; other samples contained high concentrations of one or more salts; and still others had a composition intermediate between these extremes. If similar variation characterizes the distribution of salts in alkali soils generally, it may be safely concluded that the analysis of samples such as are commonly submitted by practical farmers is a waste of time. In fact, the conclusions that are likely to be drawn from the analysis of such samples may be so erroneous as to lead to the recommendation of practices the very opposite of those that should be employed.”
In spite of Kelley's prominence in the soil science community, his opinion seems to have been largely ignored, at least for a time. Other aspects of the spatial variation of soils were also discounted. Hursh (1944) pointed out the significance of what eventually became known as "macropores" or "preferential pathways" for the transport of water and solutes in field soils: "A single dead-root channel, worm-hole or insect burrow may govern both the draining of water and escape of air through a considerable block of soil." However, macropores became an object of intensive research only three decades later. Except for those dealing with soil classification and mapping, virtually all soil scientists from the 40s to the 60s carried out their research on soils as if their spatial heterogeneity was not an issue at all. None of the textbooks published toward the end of this period in soil physics (e.g., Baver, 1956; Childs, 1969), soil chemistry (e.g., Bear, 1964), or soil mineralogy (Marshall, 1964) makes any mention of spatial variation or heterogeneity, or analyzes in detail the extent to which field conditions can differ from those in the laboratory. In soil physics, in particular, beside prototypical soils (e.g., the Yolo clay loam), thoroughly sieved and dried prior to experiments, many researchers also viewed uniform columns of glass beads as suitable models of soils (e.g., Biggar and Nielsen, 1964), and used data generated with these systems to test various mathematical theories of water and solute transport that were subsequently applied to field soils, and unfortunately often still are (Beven and Germann, 2013).

In the late 50s and early 60s, attitudes toward soil heterogeneity began to change. Various researchers, in steadily increasing numbers, felt compelled to go outside their laboratories and office buildings to look at soils in the field, and found that they exhibited significant spatial variation (e.g., Ferrari and Vermeulen, 1955; Hammond et al., 1958; Krasil'nikov, 1958). The research effort that this realization stimulated developed along two major directions.

A first avenue of inquiry focused on the estimation of the number of samples necessary to obtain estimates of soil parameters that are representative of entire fields, soil mapping units, or even larger areas. By 1963, the spatial variability of soils in this context was recognized as a "serious problem" (Mader, 1963), urgently requiring attention (Leo, 1963). Work along these lines intensified in the 60s (e.g., Ike and Clutter, 1968; Webster and Beckett, 1968) and culminated in the early 70s with two landmark articles (Beckett and Webster, 1971; Rogowski, 1972) that studied in detail the representativeness of soil parameters estimated over specific areas. A useful outcome of this research was the realization that, in order to represent adequately the properties of soils within any given region, the number of samples one needs to take should be correlated with the heterogeneity of the soils.

This realization paved the way for the second line of inquiry, explored during the 70s and especially after the 80s. Since sizeable numbers of samples often had to be obtained to characterize soils in targeted areas, it seemed wasteful to simply apply to the resulting data a classical statistical treatment, leading to traditional means and variances, but in effect discarding entirely the spatial information (location and position in the landscape) associated with each sample. Instead, it appeared more and more evident over time that methods of analysis that made use of this spatial information would be far more fruitful. In particular, they could provide soil property maps, for which, according to Webster and Beckett (1968), "farmers and engineers have a strong need […] to give them information about the soil on which they are working."

As a matter of fact, in the 70s onwards, the research on soil heterogeneity zeroed in increasingly on techniques of spatial interpolation of punctual data. Webster (1977) reviewed in detail a number of empirical techniques to carry out this type of interpolation. As soon as Journel and Huijbregts' (1978) book got published, most researchers seem to have heeded the message that, as Webster and Oliver (2007, p. 2) put it, "the answer is to use geostatistics". However, within a few years, references to the questions geostatistics might have been answering have in most cases tended to disappear from the horizon. Warrick and Nielsen (1980), for example, started their landmark chapter on geostatistics with the emblematic sentence "Soils vary", made some vague mention about a grower who "needs sound scientific input to support his own good judgment", but otherwise provided precious little discussion of where and why it might be imperative for soil scientists to go through the time-consuming effort needed to quantify the spatial heterogeneity of soils or produce property maps, nor do they discuss what the hypothetical grower might do with the information. Other authors have done similarly. For example, Mulla and McBratney (2002) and Oliver and Webster (2014), very recently, make no mention at all of any reason why the heterogeneity of soils is of practical concern.

Since the late 70s, literally thousands of articles devoted to the application of geostatistics to soils have adopted the same pattern, illustrated schematically in Fig. 1. The spatial variability of a soil in a given context, for example an agricultural field, is considered to be a problem, for reasons that are seldom mentioned. A generally square sampling grid is used to make measurements in the field, or to take samples that are analyzed later in the laboratory. The resulting data are used to evaluate a geostatistical function called the empirical semi-variogram, to which are fitted a number of standard parametric models. These models, in turn, through an interpolation process known as (co)kriging, allow the production of soil property maps, which as Oliver and Webster (2014) write, are generally the "ultimate goals" of many users.

This trend has been questioned by a number of researchers over the years (e.g., Jury, 1982; Letey, 1987; Russo and Jury, 1987; Baveye, 2002). Letey (1987), for example, has argued that the study of the spatial variability of soils should involve “much more than simply applying geostatistical analysis to transect data.” Many of the criticisms made in that context are reminiscent of earlier comments made, in particular by De Bakker (1970), about the pointlessness of soil classification and mapping efforts in the absence of clearly defined theoretical or practical purposes. Some criticisms of the emphasis on geostatistics have zeroed in on technical aspects. A key example in that vein is Russo and Jury's (1987) thorough demonstration that the correlation scale of a spatially variable soil property, estimated on the basis of its semivariogram, depends strongly on the scale of observation, e.g., on the numbers of samples, the grid size, and the location of the sampling points, as well as on the choice of a particular method for the estimation of the semi-variogram. In terms of the location of sampling points, Russo and Jury (1987) show that estimates of the correlation scales based on data along transects may underestimate the correlation scale of the underlying process by a factor of 2 or greater. Nevertheless, in spite of such encouragements to pause, take stock, address fundamental methodological questions, and explore the use of alternative perspectives and techniques when available (e.g., Milne et al., 2005), the application of geostatistics has retained central stage in the research on the heterogeneity of soils over the last three decades, with the same modus operandi.

It should be pointed out that the problems with the pattern illustrated in Fig. 1 are not due to any fault of geostatistics itself. Especially since it has been reformulated in the much sounder Bayesian framework (Omré, 1987; Handcock and Stein, 1993; Pati et al., 2011), but even to some extent in its original frequentist description, the theory of geostatistics is an interesting field of research, and has proven to be a very useful means to an end in a number of contexts where its use is justified and the questions it can help address are carefully posed (e.g., Wang et al., 2011; Houston et al., 2013). As issue in the present article is the fact that in the overwhelming majority of articles describing the application of geostatistics to deal with the heterogeneity of soils, the means has
itself become the end. However, as the next section will illustrate, many soil-related questions can be answered satisfactorily without having to invoke geostatistics at all, nor any alternative form of interpolation of punctual data.

3. Soil heterogeneity does not always have to be quantified and accounted for explicitly

3.1. Transport of solutes in the field

Perhaps one of the clearest examples of a situation where it is not entirely necessary to account for soil heterogeneity explicitly in the description of some soil-related process is provided by Jury's (1982) seminal work on the simulation of solute transport under natural field conditions. In most cases, even the comparatively simpler modeling of water flow in these systems is made very difficult by the extreme lateral and vertical heterogeneity in soil water hydraulic properties due to textural or density differences, as well as imprecise but large flux values caused by channeling through holes or cracks or by instabilities in water flow fronts. Since solute movement predictions, particularly for non-adsorbed chemicals, depend to a great extent on an accurate knowledge of the water flux distribution, models of solute transport in the field suffer from the same problems as water flow models. Jury (1982) pointed out that, in fact, the task can be eased considerably if one realizes that only often an average representation of downward water flow is needed for hydrologic applications. Frequently, the minimum solute residence time is all that should be estimated by a model used in environmental quality projections, particularly of hazardous wastes.

In that context, Jury (1982) recommended abandoning the deterministic approach to modeling chemical transport, which requires a detailed characterization of the spatially varying properties of soils, in favor of a transfer function model, a device commonly used by electrical engineers and hydrologists to characterize systems whose internal mechanisms are unknown or unknowable. The system, in this case the heterogeneous field soil, is characterized entirely in terms of the way it transforms an input function (solutes added to soil surface) into an output function (solutes moving through the soil). By taking such a point of view, Jury (1982) showed that one may derive a simple theory to estimate the average and extreme behavior of solutes moving through a porous medium as a function of a straightforward field-measurable function associated with the distribution of travel times through the porous medium. Various field studies (e.g., Jury et al., 1982; Butters and Jury, 1989; Jury et al., 1990) have shown this model to be highly successful at predicting the transport of solutes (as long as they were not initially present in the soils in appreciable amounts), on the basis of single calibrations to measure their travel time distribution.

3.2. Modeling of watershed-scale processes

At larger spatial scales, several different mathematical models have been developed over the last 30 years to describe the transport of water and solutes (e.g., fertilizers, pesticides) within watersheds. These models range from very detailed (and generally over-parameterized) “distributed, physically-based” models like MIKE SHE, which account explicitly for soil variation within watersheds, to “lumped, conceptual” models like the Birkenes model or BIM, which ignore soil variation and represent watersheds as sets of connected black boxes (Baveye and Boast, 1999). The first category of models corresponds to what has been termed in hydrology a “bottom-up” approach, whereas the lumped models are traditionally associated with that is called a “top-down” approach. From a soil science perspective, the bottom-up approach is, at least at first, significantly more attractive, since it starts with the kind of local, macroscopic measurements traditionally carried out in the field by soil scientists. However, anyone who has tried the bottom-up approach to predict the response to, e.g., a rainfall event, of a watershed of even moderate size, quickly realizes how practically unfeasible this approach is, because most of the data needed are not available, and cannot realistically be obtained. One can in principle use pedotransfer functions to derive some parameter values from more easily measured variables, like the particle size distribution. These empirical equations may provide reasonable approximations as long as one remains within the particular set of soils for which specific pedotransfer functions have been (painstakingly) developed (e.g., Wösten et al., 1990) but they are likely to be inaccurate if they are used outside that set, and therefore they rarely constitute a practical solution.

Basu et al. (2010) provide a vivid example of the magnitude of the problem. They investigated the hydrologic response (hydrograph) of a mesoscale watershed (∼700 km²) in northeastern Indiana. Survey maps indicate that soils in the watershed have

Please cite this article in press as: Baveye, P.C., Laba, M., Moving away from the geostatistical lamppost: Why, where, and how does the spatial heterogeneity of soils matter? Ecol. Model. (2014), http://dx.doi.org/10.1016/j.ecolmodel.2014.03.018
been classified according to 80 soil series and approximately 14,000 soil mapping units, which suggests “the need for detailed, spatially explicit hydrologic models for description of water flow” (Basu et al., 2010). However, a detailed analysis using the Soil-Land Inference Model (SoILM) showed that the available water storage (AWS), a key parameter for the prediction of the hydrograph, was far more uniform within the watershed than the heterogeneity of soil survey maps would have led one to assume. This observation is similar to that, made by Bouma (1989) quite a few years earlier, namely that soils that are pedologically different are not necessarily so from a hydrological point of view. Indeed, Basu et al. (2010) found out that more than 80% of the soils in the watershed had AWS values between 80 and 120 mm, with higher values restricted to soils located along the riparian zones. This relatively uniform distribution of AWS values makes it possible to represent the entire watershed as a single, effectively homogeneous hydrologic unit. Therefore, in spite of significant physical heterogeneity, the watershed considered by Basu et al. (2010) exhibited “functional homogeneity”. Probably because of that, a very simple, parsimonious Threshold-Exceedance-Lagrangian Model (TELM) developed by the authors using a top-down approach to predict the principal moments of event hydrographs was able to perform as well, without parameter calibration, as a complex physically-based hydrologic model (SWAT).

This result is far from unique. There are many more instances in the literature of situations where very simple models have performed as well if not better than very complex, distributed models when the latter were an option, or have fared well when more complex models could not be parametrized (Baveye and Boast, 1999; Basu et al., 2011; Zanardo et al., 2012). For example, Christopherson et al. (1982), using the very simple, lumped BIM model representing a watershed as a series of 3 hypothetical reservoirs, were able to closely reproduce the major features of a hydrograph and elution curves of Al³⁺ and H⁺ in the Birkenes watershed in southern Norway. More recently, Bormann et al. (2009) showed in two case studies, involving three different land use change scenarios, that there was broad agreement among spatiallyexplicit and lumped hydrological models with regard to expected hydrological changes.

These and other similar observations have caused a sea change in the way structural heterogeneity is approached. Early on, the perception that models needed to be made simpler rather than more complex led Rogowski (1972), in his landmark article on soil spatial variability, to identify criteria according to which the soils in a watershed could be assumed uniform, i.e., to determine when their heterogeneity could be ignored. Unfortunately, this perspective, which in retrospect appears eminently sound, was not explored for a few decades to the extent it should have, in large part because of the focus of a lot of the research on a geostatistical viewpoint. However, in recent years, Rogowski’s (1972) message has resurfaced. For example, McDonnell et al. (2007) have pointed out the need to “figure out a way to embed heterogeneity or the consequence of heterogeneity into models in a manner that does not require enormous amounts of generally unavailable data.” One way to do this, according to these authors and in resonance with Jury’s (1982) perspective, is not to concentrate one’s efforts on a detailed characterization of the heterogeneity of the system, but to focus instead on the “properties that emerge with increasing scales, and on their resulting hydrological effects.”

3.3. Risk of groundwater contamination over large geographic areas

A third example of a situation where soil heterogeneity does not have to be accounted for explicitly, or at least where the paradigm of Fig. 1 would not be particularly useful, is related to attempts to predict the risk of groundwater contamination by pesticides. Mathematical models (like Hydrus) are used traditionally for that purpose, and to determine if given herbicides or pesticides should be used in particular states or countries given the nature of the soils. These models require quantitative information about the spatiallyvariable adsorption of pesticides to soils and detailed data about the transport and sorption properties of subsurface materials, all of which are rarely, if ever, available. However, an interesting analysis by Worrall (2001) suggests that this information is not essential. Using observations of pesticide occurrence in 303 boreholes across 12 states in the midwest of the U.S., he shows that the molecular topology of the pesticide molecules themselves is, in and of itself, a good basis to discriminate between polluting and non-polluting pesticide compounds, regardless of the spatial variation of the soils through which the pesticides transit. A logistic regression model involving only parameters related to the topology of the pesticide molecules accounts for 97% of the variation in groundwater contamination data, and applies equally well for both linear and cyclic molecules. This astoundingly high correlation has very significant consequences in terms of the “certification” of categories of herbicides or pesticides in particular states, provinces, or countries. Unlike what is currently assumed via the use of physically-based models to decide whether certain agrochemicals should be allowed or not, Worrall’s (2001) results suggest that this certification process should not be location-dependent: some compounds, because of their chemistry and the topology of their molecule, should probably be banned everywhere, regardless of the nature of the soils.

3.4. “Precision” farming or site-specific management

As a final example, one could consider the case of “site-specific management”, also known often as “precision” agriculture or farming, which in recent years has served increasingly as an implicit justification for the research on soil heterogeneity and in particular for the application of geostatistics (e.g., Oliver, 2010). Schafer et al. (1984) were apparently the first, at least in print, to suggest that soil property maps like that at the right of Fig. 1, could be “used to control fertilizer and pesticide application and tillage operations.” Indeed, at least in theory, if the parameter(s), like available P concentration in the soil or available soil water content, that control(s) the growth of crops in given situations could be identified without ambiguity, if maps of these parameters could be drawn at a suitable resolution, and if practical methods could be devised to deliver the missing ingredients in the exact amount needed by the plants at each location in the field, not only would crop yield be maximized, but also – added environmental benefit – the chances of excess nutrients leaching downward in the soil to the groundwater would be minimized.

What, in principle, appears like a proverbial “win-win” solution may not necessarily be so, even in principle. There is always the possibility that P or N deficiency in one part of an agricultural field be due to a higher leaching rate in the area, leading to an increased rather than decreased risk of groundwater contamination if more fertilizer is applied there. Aside from this potential environmental downside of site-specific management, which is rarely mentioned (John Letey, personal communication, 2014), a number of authors (e.g., Anderson and Bullock, 1998) have also observed that yields are not necessarily higher with variable-than with uniform rate application of fertilizers, but can be equal or even lower. There are several possible reasons for these observations, a key one being that the parameter(s) constraining yields may not have been identified correctly, or that the crops did not respond significantly, or responded negatively, to the parameter(s) that was/were varied in the experiments.

When yields do increase as a result of site-specific management, one could consider that these increases provide ample justification a posteriori for the approach toward the heterogeneity of soils.
illustrated in Fig. 1. Indeed, there is a societal demand to insure food security in the future, and the fact that site-specific management might maximize crop yields is important in that context. Reduced groundwater contamination may also be an attractive objective for society. However, individual farmers, in industrialized countries at least, generally do not care much about either maximizing yields or minimizing fertilizer application to obtain the same yield, unless there are financial incentives from the government to do so. The concern of most farmers is on maximizing profit, which means that they need to subtract from their gross income the cost of sampling the fields before each growing season, of getting the samples analyzed and the results mapped, of purchasing and maintaining new fertilizer spreading equipment to allow variable rate application, of taking the time, if need be, to do multiple passes through their fields instead of just one as they did earlier to apply fertilizers, and, last but not least, of learning how to operate and troubleshoot the sophisticated, computerized control systems on their new equipment. When all these costs are factored in, the outcome in some studies is that site-specific management based on maps generated from soil samples is often not economically appealing to farmers (e.g., Piez et al., 1994; Rejesus and Hornbaker, 1999; Sylvester-Bradley et al., 1999; Yang et al., 2001).

One way to reduce the operational cost, at least partially, would be to eliminate the sampling and map generation steps, i.e., not deal with soil heterogeneity in the manner depicted in Fig. 1, but instead measure directly in the field, with an “on-the-go” sensor, the parameter that is limiting yield, and use the information to adjust the delivery of the corresponding fertilizer. The idea sounds very good and has attracted quite a bit of attention over the last decade, but it hinges on the ability to monitor instantaneously, in the field, parameters that normally require time-consuming wet-laboratory measurements. For some parameters, it has been suggested repeatedly in recent years that visible near-infrared reflectance (VNIR) spectroscopy could be used as a substitute to laboratory methods (e.g., Kinoshita et al., 2012). The VNIR spectral region is dominated by weak overtones and combinations of vibrational bands of light atoms that have strong molecular bonds, for example chemical bonds that contain H attached to atoms such as N, O, or C. As suggested early on by Dalal and Henry (1986), it may, therefore, be possible to measure soil constituents such as water, organic C, and N using the NIR technique, after a site-specific calibration has been carried out (e.g., Ehsani et al., 1999; Adamchuk et al., 2004). However, since VNIR spectra do not contain any feature related to P, K, or most other plant nutrients, the only way VNIR sensors can be used to detect these nutrients is via “surrogate” correlations that may or may not exist, and in any case need to be determined experimentally, between the soil concentration of these chemical compounds and those of soil-borne O-, C-, and N-bearing compounds (Wu et al., 2010). Further drawbacks of VNIR-based sensors are that their readings are strongly affected by the soil water content of the soil (Wu et al., 2009a), as well as by the roughness of the soil surface (Wu et al., 2009b). Because these influences would need to be taken into account to interpret the VNIR signal correctly in the field, the use of VNIR-based on-the-go-sensor is all but straightforward at this stage. Even if progress could be made in this context, VNIR sensing would still suffer from the fact that its penetration depth in soils is extremely shallow (less than a millimeter), a feature that is inherent in the physics of the technique, and cannot be alleviated. This latter shortcoming, of a very localized and likely not very representative measurement, is also shared by other electrochemical sensor concepts that can in principle be used to measure nutrient content, e.g., the ion-selective field effect transistors (ISFET) or ion-selective electrodes (ISE) (Adamchuk et al., 2004).

At present, it still appears somewhat unclear whether variable-rate application systems using on-the-go sensors are economically attractive to farmers (e.g., Maleki et al., 2008; Maleki and Zamiran, 2009; Biermacher et al., 2009). As with the more classical approach, based on soil sampling and property mapping, yields in this case are not always significantly higher than with conventional, uniform application of fertilizers. In addition, there does not seem to be any convincing evidence at this point warranting a high level of confidence in the economic profitability of variable-rate application with on-the-go sensors, once all the costs are factored in. One might conclude from this quick survey that it is not certain at present whether or not farmers benefit from adjusting their practices to take soil heterogeneity into account, if the way to do so follows closely the paradigm illustrated in Fig. 1, or is inspired by it, as with the on-the-go sensors. However, there is another way to approach the question, in line with what was done routinely in the past and with the “niche management” still practiced in various countries (e.g., Tittonell and Giller, 2013). We now have access to satellite images and to aerial pictures that are of far better quality than when the topic of the remote sensing of soils first arose considerable, yet temporary, enthusiasm among soil scientists in the early sixties (e.g., Curtis, 1963; Webster and Beckett, 1964; Simakov, 1964; Cuanalo, 1966). Web sites like Google Earth now provide for free a full coverage of the earth surface with high-resolution satellite imagery, as in Fig. 2. The spatial heterogeneity...
of the soils that these images reveal is sometimes very impressive (and is often taken advantage of by archeologists, with remarkable results). Farmers could use these images as guides to delineate portions of their fields that could be treated uniformly as separate units. Such an approach is clearly adopted in the vineyards on the right of Fig. 2b, where different parcels, separated by hedges, are apparent.

Collaboration of soil scientists with the various companies acquiring satellite images like those in Fig. 2 could insure that they be taken at times (season, but also time of day), when the different hues of the soil surface, or of the vegetation growing on the soils, would be most revealing of the spatial heterogeneity of the underlying ground. Satellite imagery may not be the only game in town either, as aerial photographs can be obtained easily, and more cheaply, by using Unmanned Aerial Systems (UAS), or “drones”, to obtain soil images at minimal cost and at higher temporal frequencies (Emhke, 2013). Several companies, e.g., Precision Drone (http://www.precisiondrone.com) in the US and Airinov (http://www.airinov.fr) in France, have been created recently to provide this service to farmers.

4. Scale of measurements and perception of soil heterogeneity

The previous section illustrated a number of cases where the paradigm of Fig. 1 either was not relevant, not feasible, or not demonstrably helpful. A general observation one may derive from these various examples is that, even if soils vary in a particular context of interest, it is sometimes unnecessary to automatically apply the paradigm of Fig. 1 to them, and it may be far better to explore a different route. Nevertheless, there may be situations where maps generated by the spatial interpolation of punctual measurements (using geostatistical methods or alternatives) might conceivably be useful, for example to eliminate any spurious effect due to soil heterogeneity when trying to decide in an agricultural experiment station whether a particular treatment is more advantageous than another. In these instances, an interesting question is whether each soil property map that is generated is unique, or whether multiple maps are possible, making it necessary for potential users to decide which map is appropriate for the ultimate objective being pursued.

This issue is not abstract or hypothetical, but is very practical and affects every single soil property, static or dynamic, one might want to evaluate spatially. For example, if one wants to measure the infiltration rate of a soil at different locations in a field using a single-ring infiltrometer, one has to decide on the diameter of the ring. Often, that choice is made by the manufacturer of the equipment used, who likely sells only one type of infiltrometer, but in principle, any diameter is acceptable. In a seminal article, Sisson and Wierenga (1981) carried out infiltration rate measurements in the field with 3 different single-ring infiltrometers successively positioned in a nesting pattern (Fig. 3a) in such a way that at 25 locations in the field, marked by arrows in Fig. 3, the three rings were concentric. Comparison of the values obtained at these points shows that in some cases (e.g., third position along the transect in Fig. 3b, starting from the left), the large-ring value is the highest of all, whereas sometimes (fourth position), it is the lowest. There are other differences between the infiltration profiles obtained with the three rings but probably the most striking is how different the overall perception of soil heterogeneity is with each of them. In the case of the small rings, the infiltration rate varies by more than 2 orders of magnitude along the transect shown in Fig. 3b. With the intermediate ring, the range of variation between highest and lowest value is considerably narrower, and only represents about a 6-fold increase. For the large rings, the highest infiltration rate value found is only 21% higher than the lowest value, and overall, along that transect at least, the field does not appear very heterogeneous at all with respect to the infiltration rate, on the contrary. Therefore, clearly, how spatially variable a soil is in the field is related as much to the perspective of the observer as to the soil itself.

Sisson and Wierenga’s (1981) observations, and similar results obtained by others as well (e.g., Bouma, 1992) stimulated further theoretical research in the years that followed the publication of their work (e.g., Baveye and Sposito, 1984, 1985; Cushman, 1988; Beckie, 1996; Mayer et al., 1999; Baveye et al., 2002), but surprisingly, only a few articles attempted to investigate in more detail how the observer-dependence of measurements affected the perception of spatial heterogeneity of soil properties, and the application of geostatistics. Notable among them is the series of articles by Starks (1986), Zhang et al. (1990), Shouse et al. (1994), Ellsworth and Boast (1996), and Warrick et al. (1999). Shouse et al. (1994) measured steady-state infiltration rates in a 4.0 by 4.0 m field plot at 3 different scales and observed, like Sisson and Wierenga (1981), a decrease in mean infiltration rate with decreasing measurement scale. Maps that Shouse et al. (1994) generated using the infiltration rates obtained at the two smallest measurement scales (0.25 by 0.25 m and 1.0 by 1.0 m squares, respectively) exhibited very marked quantitative differences (Fig. 4), which the authors attributed in part to scale-dependent edge artifacts associated with the infiltration measurements themselves. In addition, Shouse et al. (1994) showed how, as suggested early on by Matheron (1962), experimental semi-variograms derived from measurements made with a specific geometrical support could be deconvoluted to
produce a “point” semi-variogram, corresponding in principle to punctual measurements. In a later study of the transport of bromide, nitrate, and chloride in a field soil, Ellsworth and Boast (1996) also demonstrated that point semi-variograms could be derived relatively easily.

From these point semi-variograms, one could in principle produce by convolution the semi-variogram associated with any given geometrical support. More interestingly, one can also generate, via simulation, the frequency distribution of data associated with geometrical supports of arbitrary size. Warrick et al. (1999) obtained such frequency distributions of P concentration at a field site in Illinois where, generally, fertilization is recommended if P concentration is below a threshold of 40 units. Warrick et al. (1999) showed that if the entire field was chosen, the mean P value was around 50, above the threshold, and no fertilization was needed. By contrast, if the point support was chosen, approximately 38% of the total area was below the threshold value. Basically any percentage of the surface between 0 and 38% could be obtained by selecting smaller or larger geometrical supports for the measurement of P.

In other words, depending on how local measurements are made in the field, one can end up with very different recommendations on how to manage the field optimally.

The same situation arises when one tries to determine whether a contaminated soil in a given area is below or above regulatory standards with respect to a particular contaminant. If the geometrical support of measurements is small, many contamination hotspots may be found and would need to be remediated, whereas if one takes a support that is even slightly bigger, the hotspots get “diluted” by less contaminated soil and pretty much disappear from the picture. Such changes in geometrical support have been suggested candidly as a valid engineering solution to, in essence, make problems disappear at no cost (e.g., Brakewood and Grasso, 2000).

The various articles mentioned in this section all suggest that, in more ways than one, the magnitude of the spatial heterogeneity of soil properties is very much in the eyes of the observer. Without a specific purpose when envisaging the quantification of this spatial heterogeneity, one would have no way of deciding which perspective, among many possible, is most appropriate. Fortunately, as shown in the following section, in many situations there are clear indications about the “proper” scale to adopt to carry out measurements. At the same time, unfortunately, it is also clear that this proper scale often does not correspond at all to the geometrical support generally considered in the research on the spatial variability of soils, or by on-the-go sensors developed in recent years for site-specific management of agricultural fields.

5. Significance of soil heterogeneity at scales other than those of traditional measurements

5.1. Soil “ecosystem services”

A first area where researchers are likely to have to deal with spatial and temporal scales that differ from those traditionally considered in soil science involves the multifarious services that soils render to human societies, a topic that has a long history (Baveye et al., 2013) but for various reasons has become very popular since the publication of the Millenium Ecosystem Assessment report in 2005. What is known for better or worse as the “ecosystem services” (ES) of soils has since been the object of a large body of research over the last few years (e.g., Dominati et al., 2010; Robinson et al., 2013), fueled by sizeable interest on the part of regulators and policy makers in many countries.

Among the many ecosystem services of soils, provisioning services like the contribution of soils to the production of biomass can be envisaged at local scales, e.g., within a field plot or at the level of a farm, but other services necessarily require broader spatial and temporal scales to be accounted for adequately. Regulating services, like the filtering of water percolating in the subsurface before it eventually discharges into streams, typically take place over large distances and long timescales. Therefore, it seems intuitively clear that, progressively, regulatory decisions based on consideration of the ecosystem services of soils will have to gravitate toward a range of spatial scales that are significantly larger than typical agricultural fields or farms (Smith et al., 2012; Dale et al., 2013), and will need to adopt an integrated and comprehensive perspective (e.g., Zhang et al., 2013).

At present, it remains largely unclear which spatial scale(s) should be envisaged in any particular situation. However, regardless of the scale that is eventually considered most suitable to describe in detail the various ecosystem services of soils in a particular region, some researchers will undoubtedly be tempted to estimate these ecosystem services spatially using a variant of the
paradigm depicted in Fig. 1. Work along those lines has already begun. On the basis of information provided by classical or digital soil survey maps, various authors have recommended using geo-

statistical methods to quantify ecosystem services over catchments or regions of various extents (Jiang et al. 2007; Peukert et al. 2012; Glendell et al. 2014; Menon et al. 2014). However, this approach is fraught with difficulties. Data gathered as part of soil survey efforts, even after application of pedotransfer functions, are not necessarily relevant to the description of soil ecosystem services. For example, the granulometry of soils, the composition of the minerals they contain, or a nominal description of their architecture (e.g., platy, columnar), are typically not of much use to describe the capacity of soils to transmit rainwater to subsurface strata under conditions where a sizeable component of water transport is controlled by wormholes or preferential pathways resulting from root decay, or to predict solute transport below the root zone when water move-

ment is significantly affected by plant growth (e.g., Pang and Letey, 1999). In fact, soils that are classified in different categories and mapped as distinct units can very well have similar properties relative to one or more ecosystem services, and should therefore be grouped together, in some fashion, instead of being dealt with sep-

erately.

This situation is in many ways reminiscent of that observed by Basu et al. (2010) in their description of the response of mesoscale watersheds, and by many other authors who have been confronted over the years with the huge data requirements of bottom-up, dis-

tributed hydrological models. Therefore, it is very reasonable to consider that, in the case of soil ecosystem services as well, a top-
down approach would be far preferable to a bottom-up perspective, whether or not it relies heavily on geostatistical interpolation (McDonnell et al. 2007). A first step in a top-down approach would consist of first identifying and classifying ecosystem services at the spatial scale that one deems most relevant to the particular question that is being asked. Indeed, as argued in detail by Fisher et al. (2009), “any attempt at classifying ecosystem services should be based on both the characteristics of the ecosystems of inter-
est and a decision context for which the concept of ecosystem services is being mobilized. Because of this there is not one clas-
sification scheme that will be adequate for the many contexts in which ecosystem service research may be utilized.”

Once the decision context, a suitable spatial scale, and the dif-

erent ecosystem services at play have been identified, the next logical step should consist of determining what soil properties are most directly influencing the ecosystem services provided by soils in the region of interest, and how reliable information about these soil properties can be obtained. In specific cases, some of the field data accompanying survey maps may provide part of the needed information, but in general one should expect that some other information, not typically considered in soil mapping efforts, is likely to be required. Therefore, instead of trying to bend over back-
ward to find a way to make existing and potentially misleading survey data conform to what is needed so that geostatistical meth-

ods can be applied, following the traditional pattern, it would be most useful to first pause, think the problem through, and deter-

mine what is really needed to assess soil ecosystem services in a meaningful way. Then, and only then, should we determine which mathematical technique ought to be implemented to handle data optimally.

5.2. Plant response to root zone soil heterogeneity

One of the foremost reasons one might be interested in the spa-
tial heterogeneity of soil properties is to determine how plants, be they trees in a forest or crops in agricultural fields, are likely to grow, and to assess what amendment and fertilizers, if any, are needed to optimize their yield. In that context, it seems intuitively obvious that measurements should be carried out in geometrical supports that are as close as possible to the volume of soil that plants explore and in which they take up water and nutrients, at different stages in their growth.

The literature devoted to this question is surprisingly scanty, at least compared to the striking mass of articles on soil variability in general. Almost 80 years ago, Weaver (1926), in his remark-

able atlas of the “Root development of field crops”, still frequently cited at present, illustrates the fact that the volume of soil that is probed by plant roots varies greatly from one plant species to another, and from one stage of development to another. In the case of winter wheat, for example, Weaver (1926) observes that thread-like roots penetrate vertically downward (Fig. 5a). Some run obliquely downward but seldom reach a lateral spread of more than 0.2 m. The maximum rooting depth extends astoundingly to 1.89 m. In the case of corn at maturity, the pattern is very different (Fig. 5b). Some roots have a nearly horizontal position throughout their entire course. Others run at various angles for between 10 cm and a meter, then turn downward either abruptly or with a gentle curve. Some roots go as deep as 2.4 m. Weaver (1926) estimates that over 5.66 m² of soil and subsoil are “quite thoroughly drawn upon for water and nutrients by the roots of a single plant.” Sun-

flower roots (Fig. 5c), by contrast, remain within 0.45 m of the soil surface. The taproot gives off so many laterals and tapers so rapidly that, at a depth of 0.20–0.025 m, it is only 4–5 mm, i.e., of roughly the same size as some of the major branches. Some of the lateral roots, extending in the surface 15 cm of the soil, reach as far as 1.67 m from the base of the plant. White sweet clover (Fig. 5d) has a taproot, which in some cases extends to 1.95 m. Rarely more than two large roots branch off the taproot, all of them clothed by many small laterals.

Other authors have observed similar rooting patterns for various crops, and in particular they have described how a given plant can adapt its rooting pattern depending on the possibly heterogeneous conditions that it finds in the soil. For example, root systems of some plants are plastic enough that they can systematically avoid volumes of soil that are less favorable to growth (e.g., Bingham and Garber, 1970; Linkohr et al. 2002), because of nutrient depletion or the presence of some toxicant. The recent article by Liu et al. (2013) shows how stupendously different the rooting pattern, in particular the lateral root elongation, of a phreatophytic perennial desert plant can be under a number of irrigation patterns. Some of this information has also been encapsulated into several computer models of root development (e.g., Doussan et al. 2003; Dunbabin et al. 2013), which produce outcomes that are remarkably similar to field observations, for example those of Weaver (1926).

An unescapable conclusion of this research on plant rooting pat-

terns is that the measurements of nutrient content traditionally made in very small soil samples (e.g., slightly less than 5 × 5 × 5 cm in size) are directly relevant to plant growth only for a very short period of time, from immediately after the young seedlings have exhausted the nutrient resources contained in the seeds, until their roots begin to extend outside the sampling volume. For some plants, this time may be extremely short, or even not exist at all. The situation is even worse for the measurements made with VNIRS-based on-the-go sensors for site-specific management, whose depth of penetration extends downward at most to a mm below the soil surface. In most cases what happens or does not happen within this very thin skin at the surface of the soil is utterly inconsequential in terms of the growth of plants.

The extent of the different root geometries depicted in Fig. 5 makes it abundantly clear that it is unfeasible to routinely mea-

sure anything in heterogeneous soils within a geometrical support anywhere close to that explored by plant roots, and this observa-
tion is of course even more true for trees and shrubs in semi-arid environments, whose roots often extend to depths of 40 or 50 m

Please cite this article in press as: Baveye, P.C., Laba, M., Moving away from the geostatistical lamppost: Why, where, and how does the spatial heterogeneity of soils matter? Ecol. Model. (2014), http://dx.doi.org/10.1016/j.ecolmodel.2014.03.018
(Hilgard and Loughridge, 1898). Change of support and block kriging techniques that are part of the geostatistical toolbox may enable observers to change the lateral extent of the measurement geometrical support to approximate the root exploration pattern, however the problem is not resolved at all by that approach, in that information is still only related to the surface horizon, and is lacking about soil conditions in depth.

A possible alternative option to obtain meaningful data about the soil condition relevant to crop growth would be to stop thinking that we absolutely need to quantify the spatial heterogeneity of soil properties directly, but to consider instead that we could use individual plants as integrators or indicators. Instead of measuring, e.g., the P concentration in as big a volume of soil as feasible, it might be preferable to try to measure the P content of the plants themselves, and use that value to determine whether they are able to find in the soil the amount of P that they need. To do so may be tricky, since P uptake by plants may be influenced by disease, parasites, and a number of other effects, which may be difficult to tease apart from strict nutritional considerations. Nevertheless, in the long run, this approach may be able to provide answers that are more meaningful. Similarly, assessment of the degree to which a metal-contaminated soil has been reclaimed and is now suitable for agriculture could be made, not by sampling exhaustively the soil in the root zone, but by looking directly for any sign of metal-induced vegetation stress (e.g., Rathod et al., 2013). Both for P concentration and metal-induced stresses in plants, new types of satellite or aerial imagery have become commercially available in recent years that make it possible to monitor plant status over large areas, at spatial resolutions that were not imaginable even a few years ago.

5.3. Microbial processes and soil heterogeneity

Soil microorganisms, bacteria, archaea, or fungi, pause a number of challenges to researchers that are very different than those...
associated with plants. Traditionally, microbial numbers have been assessed within macroscopic samples, via extraction and either direct- or plate counting. More recently, various types of molecular biology techniques have been developed to obtain quantitative information about the abundance of some strains. Regardless of the enumeration method one uses, the resulting population density data can be compared with bulk chemical or physical measurements made on the same macroscopic samples. From the perspective of the paradigm illustrated in Fig. 1, one could in principle perform some sort of kriging on these various measurements, produce maps of microbial activity and soil organic matter (SOM) content, and with these maps try to predict, for example, how global climate change might affect soil organic matter mineralization in large parts of the world.

It has been clear for a while, however, that such an effort would be largely pointless. Indeed, microbiologists and biochemists have observed experimentally that the interpretation of strictly macroscopic measurements relative to microbial processes presents daunting challenges (e.g., McConnaughey and Bouldin, 1985; Arah and Vinten, 1995). Even when, in a given soil, microbial growth should in principle be inhibited by unfavorable bulk chemical conditions (of pH or redox potential), or the presence of contaminants (e.g., Cu, Ag) at bulk concentrations exceeding that normally tolerated by microorganisms, often nothing of the kind occurs, and microorganisms appear to thrive (Yamamoto et al., 1985). In other cases, although macroscopic measurements in a soil indicate that microorganisms are abundant, and that an electron donor that the microorganisms are in principle able to metabolize is also present, growth on the electron donor does not ensue over periods of time that may be extremely long (>2500 years), but the mere addition of a small amount of glucose or sucrose is all it takes to jump start the process (Fontaine et al., 2007). Clearly, these observations demonstrate that traditional macroscopic measurements are unable to capture, even qualitatively, the key features of microbial activity in soils, that crucial aspects of the intricate conformation of soils at sub-macroscopic scales are being missed, and therefore that it makes no sense to interpolate macroscopic values and map them over large areas. What is needed is to characterize the spatial heterogeneity of soil properties at the micrometric scale that is typical of many bacteria and archaea. This need has been acknowledged for more than 50 years (e.g., Alexander, 1964), but could not be acted upon until very recently because of a lack of appropriate technologies. For many years, the only equipment that was of any help in this respect was the scanning electron microscope, which could be used to visualize microbial micro-environments in soils (Foster, 1988). Unfortunately, the remarkable pictures this technique often generates (e.g., Vandevivere and Baveye, 1992) could not be correlated to local chemical or physical measurements in the vicinity of the microorganisms, because such measurements were for the most part not feasible.

This situation has changed dramatically in the last few years. Significant technological advances have provided soil researchers with routine access to X-ray computed tomography systems, which, as methodological roadblocks (Baveye et al., 2010) are being resolved (e.g., Hapca et al., 2013), increasingly provide reliable information about the geometry of pores and solids in soils at resolutions as small as 0.5 μm. Concomitant progress in near-edge X-ray spectroscopy (NEXAFS), synchrotron X-ray absorption spectroscopy, and synchrotron-based microfluorescence spectroscopy of thin sections of soils has led to observations of sharp spatial heterogeneity in chemical composition of soil organic matter (Schumacher et al., 2005) and in the accumulation of trace metals (Jacobson et al., 2007; Hesterberg et al., 2011) over minute distances, respectively of the order of nanometers to micrometers. Significant advances related to biological markers now allow specific bacteria to be identified in soils and their spatial distribution at micrometric scales to be determined in thin sections (Eickhorst and Tippkötter, 2008a,b), and this information can be translated into 3-dimensional distributions using recently developed statistical algorithms (Hapca et al., 2011). In addition, very efficient modeling tools, like the Lattice-Boltzmann model, allow the description of transport and physico-chemical processes occurring in soil pores at scales directly relevant to microorganisms (e.g., Vogel et al., 2005; Falconer et al., 2012; Genty and Pot, 2013), whereas individual-based models, also developing rapidly (Gras et al., 2010), can describe the dynamics of small groups of microorganisms inhabiting the pore space (e.g., Garnier et al., 2008).

These different micro-scale techniques as well as modeling frameworks enable researchers to analyze the spatial heterogeneity of soil properties at the microscopic scale, suited to the study of microbial activities. In the next few years, as research in this area intensifies, a clearer picture will certainly become available of how micro-scale features of soils, in particular the geometry of their architecture and the topology of the pore space, cause properties to “emerge” at the macroscopic scale the way they do. Dubious concepts like the “sequestration” of contaminants in soils or the “recalcitrance” of SOM, which were invoked in the past to account for experimental observations, will be replaced by sound mechanistic representations of micro-scale processes. In addition, as soon as adequate descriptors are developed for the microscale spatial heterogeneity of soils, it will be interesting to find out if some of these descriptors lend themselves to a form of spatial interpolation similar to that illustrated in Fig. 1, which could then lead to a better understanding of the possible effect of global climate change on the fate of soil carbon over large regions. However, the research is not at that stage yet, and the need is to first find out how the current microbially-related macroscopic measurements we routinely carry out on soils at the moment, which we know are largely meaningless, can be replaced by a new generation of macroscopic measurements that convey useful information.

6. Conclusions and take-home message for soil science

The implicit premise of much of the enormous literature on the application of geostatistics to characterize the spatial heterogeneity of soil properties is that this heterogeneity needs to be quantified and mapped using the paradigm of Fig. 1, from the bottom upwards, as a prerequisite to attempts to describe processes at field- or larger scales. The analysis carried out in this article argues that this viewpoint should not be taken for granted at all, and that, even when it might in principle be applicable, attention needs to be paid to details of its implementation. Several examples are presented to illustrate the fact that there are quite a few exceptions to the basic premise, i.e., situations that can be described very well without invoking at all the spatial heterogeneity of soils, or at least by approaching it in a way that is very different than that embodied in the geostatistical paradigm. In addition, in situations where mapping soil properties may be useful, a handful of articles have demonstrated that measurements can be performed on soils in different ways, leading to sharply contrasting perceptions on the level of their heterogeneity. This raises the question of which one of these perspectives is the most adequate, a question that can be answered only by carefully considering the objective(s) one pursues. Finally, we have argued that among all the perspectives that are possible, some, which are most closely attuned to the properties of plants and microbes, require measurements to be performed over volumes of soils that are very different than those associated with the measurements currently carried out. This realization should cause us to pay closer attention than in the past to the link between the
measurements we carry out of various soil properties and the purposes for which we perform them.

These conclusions might "raise the hackles of some and [the] temperature of others", but hopefully "a few will recognize some truth in what has been said". Thomas (1992) wrote these comments more than 20 years ago, in a gem of a short note that should be assigned to every soil science graduate student to meditate over. It starts with words that are applicable as well to the senior author (PCB) of the present article: "Old age and a certain lack of patience have prompted me to write a few thoughts on our practice of science both for the edification and the annoyance of my colleagues and friends. I do not undertake this out of pique alone, although there certainly is some of that. More importantly, I wanted to register some objections to what I see as more a concern with style than with substance in environmental and soils research. The blame for this must be shared by the granting agencies and the editors and other reviewers of scientific journals who demand and generally obtain adherence to the prevailing fashion of science. Nevertheless, more blame attaches to us, the people who do the work, for succumbing without a fight to whatever is popular at the moment."

One can argue that in the voluminous literature on the spatial heterogeneity of soil properties, the undue concern with style over substance about which Thomas (1992) was lamenting is manifested by an extreme fascination for the intricacies and potentialities of the theory of geostatistics, by a comparatively limited emphasis on the purposes that motivated its use in the first place, and, at the same time, as other authors have also pointed out (e.g., Lin et al., 2005), by a severe lack of attention to the underlying causes of soil heterogeneity and its manifestations.

As an example of this neglect, while tens of thousands of articles have demonstrated how geostatistics could be applied to countless soil properties in myriads of locations around the world, Beven and Germann (2013) concluded recently that preferential flow, arguably one of the most significant causes of spatial heterogeneity of the physical, chemical, but also biological properties of soils, “has still not received the attention that its importance deserves”. Indeed, even though some progress has been made (e.g., Bouma, 1991; Radulovich et al., 1992; Boottink and Bouma, 2002), many aspects of preferential flow in soils are still as obscure now as they were 25 years ago. For example, whereas the role of earthworms and soil mesofauna in the development of macropores/mesopores in soils is well recognized, our ability to predict it quantitatively in a given soil is still in its infancy. Similarly, we cannot predict quantitatively at this point how much of a biological “hotspot” effect (e.g., Bundt et al., 1991) is likely to result from the presence of macropores in soil profiles. Sadly, one could continue with an almost endless list of aspects of macropores whose surface has been barely scratched over the last 35 years.

There is a lesson in all this that is relevant to the practice of soil science in general. It is understandable that after a group of researchers decides to find out whether a given mathematical theory, or a novel technology, might be the key to solving a particular question related to soils, some time be spent on getting deeper into the theory or technology, in particular to perfect or fine-tune it to enhance its usefulness. As this fundamental understanding is attained, one may very well reach the conclusion that the new tool or process, in which high hopes may have been vested, does not meet expectations. To some extent, one could say that this has happened in the past with the application of fractal geometry to soils (e.g., Ogawa et al., 1999) and is happening at the moment with VNIR spectroscopy sensing of soils. Regardless of the final verdict, however, there is a danger that, as they actually attempt to learn in depth about the new theory or technology, researchers lose track after a while of the objective they pursued initially, and which they thought the theory or technology in question could help them reach.

This progressive infatuation for tools is obvious in the case of the application of geostatistics to the spatial heterogeneity of soil properties. In a closely related area, the same might be said about the large effort currently underway to produce a digital soil map of the world, in which much attention is devoted to how this map can be generated, but the questions of how it is going to be used practically, and therefore – as De Bakker (1970) argued already 45 years ago – of what information it should contain to serve any useful purpose, seldom seem to surface. One could argue that similar infatuations with tools, and concomitant downplaying of ultimate objectives, are also affecting to some extent the application of, among others, X-ray microtomography, synchrotron-based XANES, and “shotgun” DNA sequencing to soils, leading to significant amounts of research being method-based rather than goal-driven. Aside from tools, an obsessive fixation on a topic, progressively more and more disconnected from its original motivation, as the current biochar craze appears to be, can be equally damaging in the long run. In this context, the “take-home” message of the present article, based on a quick overview and analysis of 35 years of research on the spatial heterogeneity of soil properties, is that it is crucial never to lose focus of whatever objective is being pursued in any research endeavor. As a reviewer of this article suggested, every effort ought to be made time and time again to ask ourselves, at each step along the way, what should be the first question on any researcher’s checklist: “Why are we doing this?”, or, formulated slightly differently, “Is what we are doing helping us address the crucial question(s) we need to answer?”

Acknowledgements

The senior author (PCB) wants to express his gratitude to Professors John Letey, William A. Jury (University of California at Riverside), Charlie Boast, and Tim Ellsworth (University of Illinois at Urbana-Champaign) for numerous stimulating discussions on soil heterogeneity in the 80s and 90s. Lively arguments with the late Professor Jeff Wagenet (Cornell University), until his untimely death in 1997, contributed greatly to sharpening some of the questions addressed in the present article. Professors Johan Bouma (Wageningen, The Netherlands) and John Letey have kindly provided very insightful comments on an initial draft of this article. Last, but not least, heartfelt gratitude is expressed to Professors Claire Chenu (AgroParisTech, France) and Wilfred Otten (SIM-BIOS Centre, Abertay University) for sharing on multiple occasions, under the auspices of the project MEPansom (funded by the Agence Nationale de la Recherche (Grant No. ANR-09-SYSCOMM-MEP-SOM), France), their deep grasp of the practical implications of soil microheterogeneity.

References


Dominati, E., Patterson, M., Mackay, A., 2010. A framework for classifying and quantifying the natural capital and ecosystem services of soils. Ecol. Econ. 69, 1119–1125.

Please cite this article in press as: Baveye, P.C., Laba, M. / Ecological Modelling (2014) xxx–xxx


New York.


