Uncertainty assessment of spatial soil information☆

Gerard BM Heuvelink[a] and Richard Webster[b], [a]ISRIC – World Soil Information and Wageningen University, Wageningen, The Netherlands; [b]Rothamsted Research, Harpenden, United Kingdom

Introduction 1
Defining and quantifying error and uncertainty in soil data 2
Single soil measurements 2
Statistical model of measurement error 2
Estimating the measurement error standard deviation from replicates 4
Proximal soil sensing and model errors 4
Spatial soil information 5
Statistical model of errors in spatial soil variables 5
Geostatistics and machine learning 5
Multivariate distribution of soil map errors 6
Cross-correlation, positional uncertainty and uncertainty of categorical soil variables 8
Positional uncertainty 8
Errors and uncertainty in soil class maps 8
Change of support 9
Spatial support and resolution 9
Uncertainty propagation 9
Uncertainty propagation example 10
The added value of uncertainty information 10
Conclusion 12
References 12

Key points

- All information about the soil embodies error.
- Acknowledgement of error translates into users being uncertain about the true values of soil properties and classes.
- Uncertainty can best be quantified by probability distributions.
- Geostatistics provides the technology to quantify uncertainty in spatial soil information.
- The propagation of uncertainty in soil information through agronomic and environmental models can be traced with Monte Carlo simulation.
- Communication of uncertainty to end users and how uncertainty can be incorporated in decision making needs more attention than it has received.

Introduction

Uncertainty is present in our daily lives. It affects our decisions on what to do. The weather forecast might tell us that there is a 60% chance that it will rain: we take umbrellas. If it says that the chance of rain is only 10% we might decide to leave our umbrellas at home and risk getting wet. More seriously, farmers want to know the likelihood of disease in their crops and the deficiencies in plant nutrients in the soil. These are matters that affect profit and loss of farm business. Agencies responsible for public health and environmental protection need to weigh the risk of doing nothing in the face of uncertain threats against the cost of acting unnecessarily to counter them when the threats are almost non-existent. There are many examples of decision making problems involving uncertain soil information. They include the remediation of polluted soil, the prevention of soil erosion, and the mitigation of pesticide leaching. They are practical matters, not purely academic exercises in statistics.

All measurements of soil properties (and other environmental variables) contain error in the sense that they depart from the true values. That error arises from imperfections in the analytical instruments, from the people who use them and from errors that occur during the processing of the recorded data to make them suitable for storage in information databases. Short-range spatial variation is another source of error, given that soil samples are never returned to where they were taken and sampling locations have positional error. Soil taken from location $\mathbf{s}$ and analyzed in the laboratory might differ substantially from the soil at location $\mathbf{s} + \mathbf{h}$,

☆*Change History*: February 2023. GBM Heuvelink prepared the update. Update consisted of: (1) uploading high-resolution figures; (2) providing publisher information for a few references; (3) correcting section header hierarchy; (4) minor corrections to the text.

even if $|\mathbf{h}|$ is as small as a few decimeters. Composite soil sampling can diminish these differences, but some error inevitably persists because even such a composite is still only a sample of all the soil at that site. All this means that we can never be sure about the true state of the soil: we, the producers and users of soil information, are to some extent uncertain.

Uncertainty tends to increase when measurements of basic soil properties are used to obtain derived ones via pedotransfer functions or mechanistic models of dynamic soil processes, for example. Interpolation from measurements to create maps of soil properties adds to the errors of measurement and so too increases uncertainties. We must conclude that considerable uncertainty is often associated with the information that is stored in soil databases and presented in various forms, including maps. This does not mean that the information is of no value; uncertainty is not the same as ignorance. In many cases we do know a great deal about the soil, but we must also acknowledge that the information is not perfect.

Some numerical expression of the uncertainty is important because it is needed to determine whether the information is sufficiently accurate for the purpose that a user has in mind. Soil data of too poor a quality might lead to flawed decisions with serious undesirable consequences, both economic and environmental. For instance, the European legislation on the use of pesticides in agriculture depends crucially on the leaching potential of these substances to the ground- and surface-water, which in turn depends importantly on soil properties. In these circumstances users should be aware of the quality of the soil information so that they can be sure that it is sufficiently reliable for their purposes. Ideally they should account for the uncertainty of the information when making their decisions.

This chapter (i) provides a statistical definition of uncertainty in soil information; (ii) extends this definition to uncertainty in spatial soil information; (iii) reviews methods that are used to quantify uncertainty in soil information, while paying attention to different sources of uncertainty; (iv) shows how uncertainty in soil information propagates through subsequent analyses; and (v) explains how uncertainty information can be used in decision making. It focuses on the quantification of uncertainty of soil properties that are measured and recorded on continuous scales: properties such as pH, particle-size distribution, and soil organic matter content. The chapter also addresses uncertainty of categorical variables, such as soil type and diagnostic properties recorded as present or absent, i.e. binary variables. It begins with defining uncertainty in a single soil measurement.

Defining and quantifying error and uncertainty in soil data

Single soil measurements

The concept of error and uncertainty is perhaps best introduced with an example. Consider a soil sample that is analyzed for clay content in the laboratory. Suppose the outcome of the analysis is 38.4%. Because of imperfections in technique, equipment or operating conditions, however, the measurement is subject to error, and the true value almost certainly differs more or less from the measured value. Perhaps it is 36.1%, or perhaps 40.0%: we cannot ascribe a single value to the true clay content. We are in effect uncertain about the true value. But we are not completely ignorant: we have a good idea of the clay content. Given the measurement of 38.4% we should consider it highly unlikely that the true clay content is greater than 60% or smaller than 20%. With this in mind we should be able to list a collection of possible values of the true clay content, and to associate a probability with each of them. In other words, we should be able to characterize the true clay content or the measurement error of it with a probability distribution.

For a continuous variable such as clay content one can usually characterize the measurement error by a normal (Gaussian) distribution (Fig. 1). It is a remarkable fact that measurement errors in many fields of endeavor have this distribution, first recognized by Gauss early in the 19th century, though the distribution itself had been formulated earlier by de Moivre and Laplace. It seems to arise because measurement error comprises numerous apparently independent random effects, which when summed give rise to the normal distribution, a result known as the Central Limit Theorem. In the measurement of clay content, for example, many disturbances contribute to the error: drying, grinding, sieving, pipetting or laser diffraction, instrumental fluctuation and so on. Together these lead to an approximate normal distribution, which one can verify by plotting histograms or plotting quantiles of the observed values against those of a normal distribution (Q–Q plots), based on repeated measurements of the clay content of the same soil sample, such as is done in proficiency testing. There are also many variables for which the measurement errors are better represented by lognormal distributions; for these it seems that individual sources of error are multiplied rather than summed. Again, one can see this by plotting histograms and Q–Q plots of repeated measurements.

Statistical model of measurement error

Representing an uncertain soil property by a probability distribution effectively defines a statistical model. This model states

$$Z_{\mathrm{T}} = z_{\mathrm{M}} + \varepsilon, \tag{1}$$

where Z_{T} is the true soil property value, z_{M} a measurement of it, and ε the measurement error. Note that Z_{T} is written in upper case (capital letter) and z_{M} in lower case. This is to distinguish the stochastic Z_{T} (a random variable with a probability distribution) from the deterministic quantity z_{M} (the measured value). The measurement error ε is also stochastic, and here assumed normally distributed. Its probability density p is given by

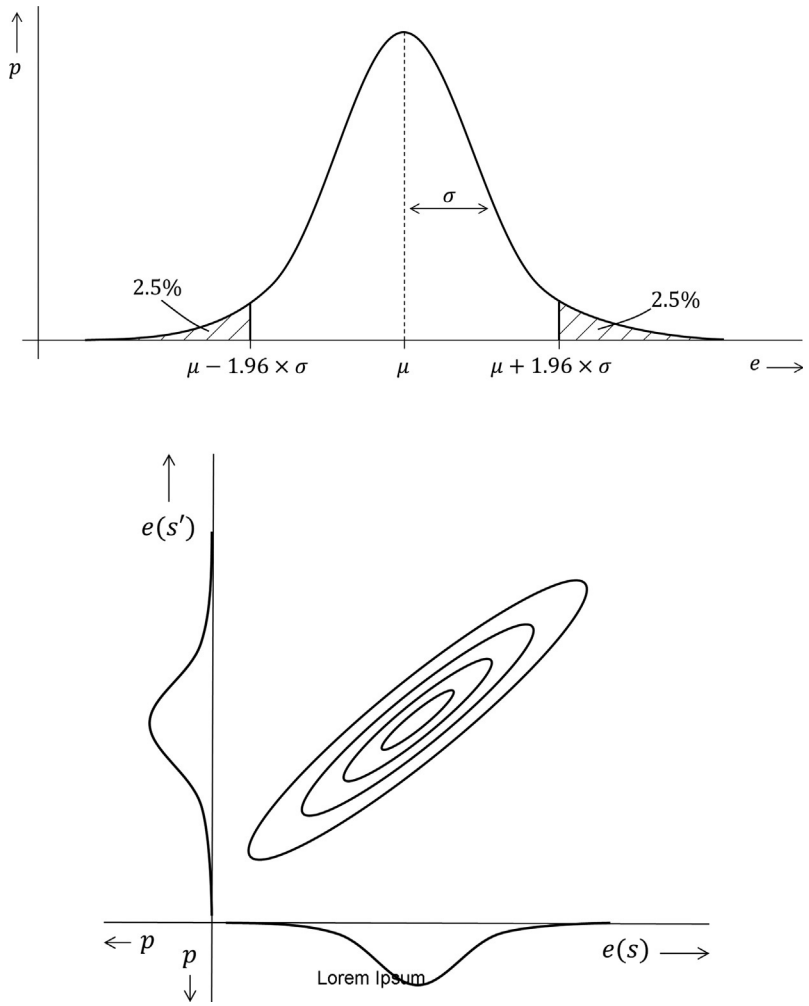

Fig. 1 Probability density of the normal distribution representing measurement or prediction error. Top: univariate case, with μ the mean (systematic error) and σ the standard deviation (measure of random error). Lower and upper limits of a 95% confidence or prediction interval are also indicated. Bottom: contour lines of a bivariate probability distribution of errors at locations **s** and **s**′. Marginal densities projected on the axes. Note that the rotated ellipses indicate a positive correlation between the errors at the two locations.

$$p(e) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{ -\frac{(e-\mu)^2}{2\sigma^2} \right\}. \tag{2}$$

The normal distribution is fully characterized by two parameters: its mean μ, which in this case signifies the systematic measurement error, and its standard deviation σ, which signifies random measurement error (Fig. 1). The square of the standard deviation, i.e. σ^2, is the variance.

A well managed laboratory aims to avoid bias, and instruments are calibrated and laboratory staff trained to achieve that. A consequence is that the mean of ε may be assumed to be zero, and its standard deviation may be estimated from replicated measurements on the same soil sample. If the same soil sample is analyzed by several laboratories then ε may be decomposed into multiple components, such as the batch, laboratory and residual error. For instance, using data from the Wageningen Evaluating Programmes for Analytical Laboratories (WEPAL, 2022), Van Leeuwen et al. (2021) estimated the standard deviation of the measurement error of total organic carbon as 0.64%, with between-laboratory variation contributing more than variation within laboratories.

We do not know what the true value in Eq. (1) is. All we can say is that it is one of an infinite number of values represented by a probability distribution; we treat it as a random variable and denote it as Z_T. It is not that we believe that the world is stochastic (Webster, 2000). In other words, we propose a statistical model of a reality that in itself is deterministic, and this model provides us with an estimate of our uncertainty about the true value. A narrow distribution (i.e. small standard deviation) conveys the notion that we are fairly certain about the true value, whereas a wide distribution (large standard deviation) means that our uncertainty is large. In the pathological case of an infinitely wide distribution (i.e. one that has infinite variance), we should be ignorant.

Estimating the measurement error standard deviation from replicates

A complete description of a measurement error requires specification of its probability distribution, but if it is assumed that the error is normally distributed about zero, i.e. has no systematic error, then all that is needed to characterize the error is its standard deviation σ. As mentioned above, σ can be estimated from replicated measurements. Fig. 2 shows a histogram of 58 measurements of total organic carbon (TOC) of the same soil sample (WEPAL, 2022). The mean is 35.58 g kg^{-1} and the standard deviation is 6.60 g kg^{-1}. Thus, if we assume that this standard deviation is representative of the random laboratory error of measuring TOC, and if for another soil sample we had only one measurement, say 50.92 g kg^{-1}, then our estimate of the TOC of that soil sample would be its measured value, with an uncertainty expressed by a standard deviation of 6.60 g kg^{-1}. If we further assume that the measurement error is normally distributed then we can state with 95% confidence that the true TOC value is in between $50.92 - 1.96 \times 6.60 = 37.98$ g kg^{-1} and $50.92 + 1.96 \times 6.60 = 63.86$ g kg^{-1}. Here, the value 1.96 is taken from the standard normal distribution, and we assume that the standard deviation of the measurement error is constant. In practice we often find that it is proportional to the measured value, meaning that σ is larger in case of a larger measured value. A decision between constant and proportional errors may be based on repeated measurements of multiple soil samples and plots of the standard deviation for each sample against its average measured value.

We can diminish uncertainty about the true TOC value of the soil sample by taking the average of replicated measurements. For instance, in this specific case where we had 58 independent measurements of the same soil sample, we are 95% confident that the true TOC lies between $35.58 - 1.96 \times 6.60/\sqrt{58} = 28.88$ g kg^{-1} and $35.58 + 1.96 \times 6.60/\sqrt{58} = 32.27$ g kg^{-1}. This is because from sampling theory we know that the average of n independent measurements has standard deviation $\sigma/\sqrt{n}$; this is the standard error of the mean. So, to double accuracy (effectively to halve the uncertainty) we need to quadruple the number of measurements. In the example above with a standard deviation of 6.60 g kg^{-1}, the standard error would be 2.09, 1.04 and 0.52 g kg^{-1} for samples of size 10, 40 and 160, respectively. This 'square root law' relation between the standard deviation and standard error might seem somewhat daunting to investigators when they plan sampling and measurement campaigns, but that is how it is.

Proximal soil sensing and model errors

In recent years the use of proximal soil sensing (Viscarra Rossel et al., 2010) has grown rapidly because it is much cheaper to measure soil properties this way than by wet chemistry. However, the techniques come at a price: these measurements tend to be more uncertain. This is because the statistical methods used to convert spectral signals to values of soil properties introduce additional uncertainty. Perhaps the most often used conversion method is partial least squares regression. The residual variances of this method should be added to the variances of the wet chemistry measurements to quantify the uncertainty of the proximal soil sensing estimates (Takoutsing et al., 2022). This additional error might be small for some soil properties and sensors, but it can be much bigger for others. This additional uncertainty may be termed 'model uncertainty', though it results from both an imperfect model connecting the signals from the sensors to the soil and imperfections in the sensors themselves. Next to measurement uncertainty, model uncertainty is an important source of uncertainty in soil data. It also occurs when properties that are hard or time-consuming to measure are predicted from properties that are straightforward and relatively cheap to measure with pedotransfer functions (Stenemo and Jarvis, 2007; Deng et al., 2009) or when soil quality indicators are derived from basic soil properties

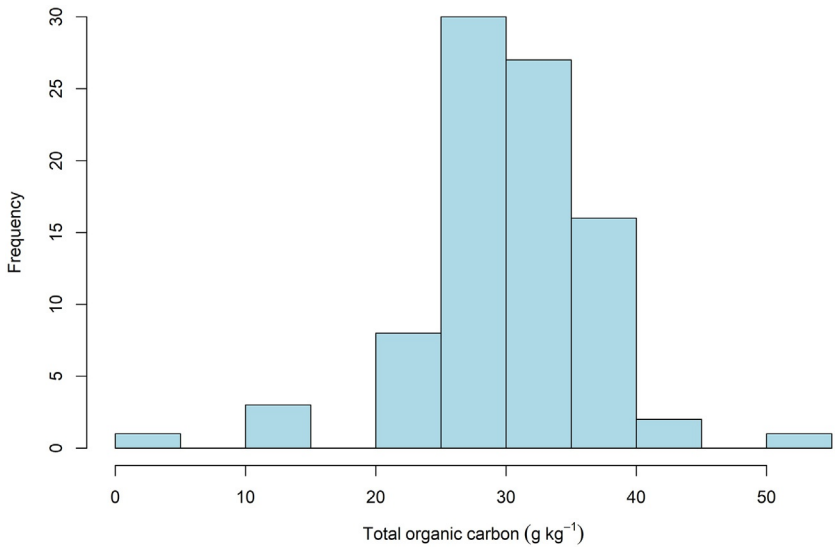

Fig. 2 Histogram of 58 replicated measurements of total soil organic carbon (g kg^{-1}) of the same soil sample. Source: WEPAL (2022). International Soil-Analytical Exchange Programme. https://www.wepal.nl/en/wepal/Home/Proficiency-tests/Soil/Proficiency-tests/ISE.htm.

and classes (Nosrati, 2013). Mechanistic soil process models also have model uncertainty. One might quantify model uncertainty in particular circumstances, but this might lack generality. For instance, one would expect a model of soil erosion constructed for the USA and calibrated with data from the USA to serve well in the USA. But it might perform less well, even badly, in the different circumstances of, say, China.

It is beyond the scope of this chapter to provide a full treatment of model uncertainty. Interested readers are referred to Brown and Heuvelink (2005).

Spatial soil information

Farming technology has advanced apace in recent years. Machines can now be programed to vary the amounts of fertilizer in response to known or estimated local concentrations of nutrients, provided their locations are known. This is an important development in precision agriculture (Oliver, 2010). In the same vein, an environmental protection agency might want to remove soil polluted by a toxic metal if it knows where the soil's load of that metal exceeds some statutory limit, or alternatively forbid cultivation or grazing there. In principle, both farmers and protection agencies want maps of soil conditions. Neither farmers nor protection agencies can measure the soil everywhere; they can at best sample on small supports, typically cores a few cm in diameter or small areas within which cores of soil are bulked. Yet they wish to know the concentrations everywhere, so that they can treat the soil in accord with those concentrations. They therefore turn to mapping, and for this they may use various approaches, as outlined in other chapters of the Encyclopedia. The focus of this and subsequent sections is on quantifying the uncertainties that are introduced by mapping. But before we address this we must first extend the statistical model of uncertainty of a single soil measurement to a model for uncertain spatially distributed soil properties.

Statistical model of errors in spatial soil variables

The extension of the statistical model presented in the previous sections to spatially distributed soil properties starts by our recognizing that Eq. (1) may be copied to all locations $\mathbf{s}$ in a region D:

$$Z_{\mathrm{T}}(\mathbf{s}) = z_{\mathrm{P}}(\mathbf{s}) + \varepsilon(\mathbf{s}) \quad \text{for all } \mathbf{s} \in D. \tag{3}$$

Here the measurement z_{M} in Eq. (1) has now been replaced by a soil prediction z_{P} to acknowledge that in this case we compare a predicted or mapped soil property to a true value. The mapped property is typically not directly measured but obtained by prediction, be it human (i.e. what was formerly convention) or statistical (i.e. digitally). This implies that spatial interpolation error will add to the uncertainty. Thus, ε is no longer solely a measurement error but rather a spatial prediction error and is likely to have a larger variance than the measurement error of Eq. (1).

Geostatistics and machine learning

Eq. (3) shows that quantification of the map error requires the specification of the probability density of the prediction error for all locations in the region. If we invoke the assumption of a normal distribution again then we shall need the mean $\mu(\mathbf{s})$ and standard deviation $\sigma(\mathbf{s})$ for all $\mathbf{s} \in D$. This might seem a formidable task, but it turns out to be a routine by-product of kriging, i.e. of mapping geostatistically. It is explained in some detail in the geostatistics chapter of the Encyclopedia and fully in text books (e.g. Webster and Oliver, 2007; Chilès and Delfiner, 2012). To krige a soil property one needs a geostatistical model, for which certain assumptions are made; these usually comprise second-order stationarity and isotropy. If the assumptions are satisfied then kriging produces unbiased predictions while minimizing $\sigma(\mathbf{s})$, the kriging standard deviation.

There are various forms of kriging that depend on the specific assumptions about the variables being predicted and mapped. They include lognormal kriging, which is used to predict variables that appear to have lognormal distributions. Indicator kriging was developed to predict and map variables that have distributions not readily parameterized; the procedure is somewhat cumbersome, and the prediction errors are no longer normally distributed (Webster and Oliver, 2007). These disadvantages are overcome by the technically more advanced disjunctive kriging. Ordinary and simple kriging are based on models in which there is no spatial trend. Trend, a fixed effect, can be accommodated in universal kriging, regression kriging and kriging with external drift, where the latter two take into account fixed effects of environmental covariates (Hengl et al., 2004). Nevertheless, whichever kriging method is used the uncertainty of the predictions is quantified. Typically, the prediction errors are small in densely sampled regions and large in sparsely sampled ones and close to the borders of the regions being mapped. In other words, the sparser is the sampling the greater is the uncertainty. The magnitude of uncertainty also depends on the spatial correlation of the target variable, as quantified by the variogram. The geostatistics chapter provides details.

If soil maps are derived by machine learning instead of geostatistics then it is more difficult to quantify the prediction error at each location in a region. The most common approach for this in digital soil mapping is Quantile Regression Forests (Meinshausen, 2006; Vaysse and Lagacherie, 2017; Poggio et al., 2021). This method requires no assumption of a normal distribution; instead it computes the quantiles of the probability distribution in a non-parametric way. All quantiles together characterize the probability distribution of the prediction error. GlobalSoilMap standards (Arrouays et al., 2014) request the 0.05 and 0.95 quantiles, which are the lower and upper limits of a 90% prediction interval, respectively. Fig. 3 shows an example.

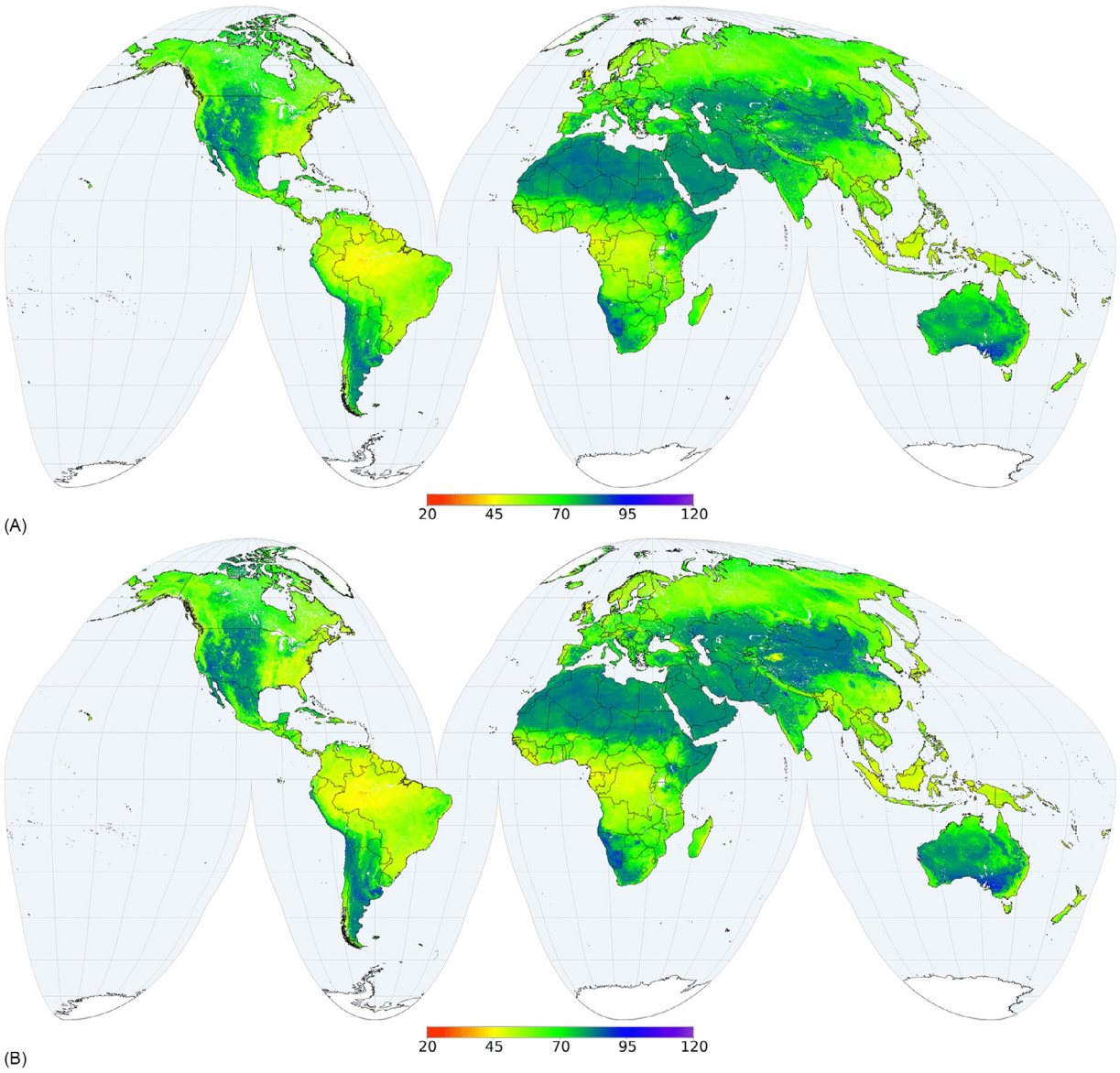

Fig. 3 Lower (A) and upper (B) limits of a 90% prediction interval of the SoilGrids 2.0 global map of pH_{water} (units in 10pH) in the 60–100 cm depth interval, as obtained using Quantile Regression Forests. Note that differences between the two maps are small, indicating high map accuracy. From Poggio L, de Sousa LM, Batjes NH, Heuvelink GBM, Kempen B, Ribeiro E, and Rossiter D (2021) SoilGrids 2.0: Producing soil information for the globe with quantified spatial uncertainty. *Soil* 7: 217–240.

Multivariate distribution of soil map errors

A comprehensive definition of soil map uncertainty also requires that the multivariate or joint distribution of map errors at multiple locations is specified. The bivariate probability distribution specifies the joint distribution of map errors for pairs of locations $\mathbf{s}$ and $\mathbf{s}'$ in D. An example of this bivariate distribution is given in the bottom panel of Fig. 1, where it is portrayed by 'contour' lines of equal probability density. Map errors are often positively correlated if the distance between $\mathbf{s}$ and $\mathbf{s}'$ is small, as is illustrated by the rotated ellipses in Fig. 1. Fig. 1 also shows that the marginal, univariate distributions of the errors at the two locations are obtained by projection of the bivariate distribution on the two axes.

If we invoke the multivariate normal distribution and assume that the map errors at all locations in the region are jointly normally distributed then all that is needed to characterize this distribution is a vector of means and a variance–covariance matrix. These so-called two-point statistics are enough because for the multivariate normal distribution the higher-order moments are completely determined by them.

Kriging implicitly quantifies the correlation between map errors at any pair of locations; that is, it quantifies the correlation between the kriging prediction errors $\varepsilon(\mathbf{s})$ and $\varepsilon(\mathbf{s}')$. They are usually not explicitly computed, but are taken into account when

values from the joint distribution are obtained by conditional spatial stochastic simulation (Webster and Oliver, 2007, Chapter 12). Fig. 4 shows an example. These simulations or 'possible realities' are created with the help of a pseudo-random number generator. In principle, an infinite number of possible realities can be generated. Each of these could be the 'true' reality, but we do not know which one it is; we are uncertain. Thus, the differences among them express our uncertainty about the true situation at any location. In fact, the spread (represented by the standard deviation) computed from very many possible realities will equal the kriging standard deviation. Their average equals the kriging prediction, which implies that maps of the prediction errors (i.e. the differences

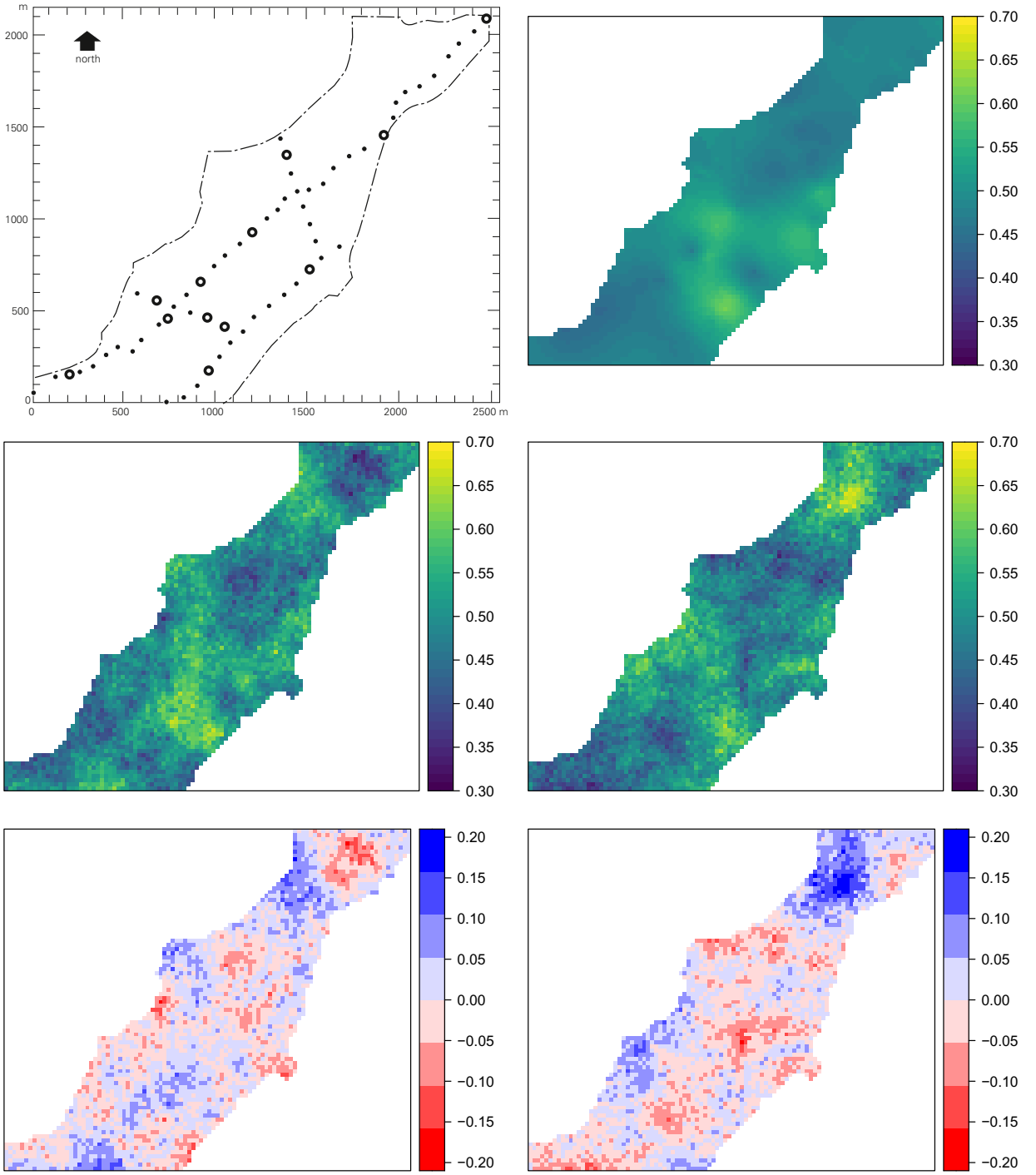

Fig. 4 Allier study area in France. Top left: stars show sampling sites of moisture content at field capacity and porosity; circled sites are those where in addition moisture content at wilting point was measured. Top right: kriging map of topsoil porosity. Middle row: two realizations of topsoil porosity, obtained using conditional spatial stochastic simulation. Bottom row: error maps corresponding with the two realizations, obtained by subtracting the kriging map from the realization. See Heuvelink (2018) for details.

between the true values and the kriged predictions, see bottom row of Fig. 4) average to zero over many simulations. This is a consequence of kriging's unbiasedness. Spatial stochastic simulation and the possible realities that it produces are essential inputs for the propagation of uncertainty through Monte Carlo methods, as described below.

Machine learning methods cannot compute the joint distribution of map errors at multiple locations because they are essentially non-spatial and do not account for residual spatial correlation (Heuvelink and Webster, 2022). If one requires maps of prediction errors that take into account spatial correlation, for computing spatial averages and their uncertainty, for example, then a possible solution is to krige the residuals of the machine learning model (e.g. Szatmári et al., 2021).

Cross-correlation, positional uncertainty and uncertainty of categorical soil variables

Environmental scientists and agencies often want to consider several soil properties together and the errors arising from their analysis. If two properties are cross-correlated then one can use a similar approach to that above to characterize the joint distribution of $\varepsilon_1(\mathbf{s})$ and $\varepsilon_2(\mathbf{s})$, or that of $\varepsilon_1(\mathbf{s})$ and $\varepsilon_2(\mathbf{s}')$ (Heuvelink et al., 2007). For example, cadmium and lead often occur together as pollutants and are correlated in the ordinary Pearson sense. If they are kriged to map the pollution their kriging prediction errors are likely to be positively correlated also, if only because a hot spot of high pollution may have been missed in sampling. Moreover, the interpolation error of cadmium at location $\mathbf{s}$ will probably be positively correlated with that of lead at location $\mathbf{s}'$, if the distance between $\mathbf{s}$ and $\mathbf{s}'$ is smaller than the variogram's range. All these correlations can be quantified using cokriging (Webster and Oliver, 2007, Chapter 10).

If measurement error is the only source of uncertainty then the correlation between errors of different soil properties may safely be assumed zero. For instance, even though organic carbon and total nitrogen in soil are usually positively correlated, their measurement errors are likely to be uncorrelated. This is because the two properties are measured in different ways and the measurement errors do not influence each other. Compositional data, such as particle-size distributions, are an exception. The proportions of sand, silt and clay measurements will be constrained to sum to 1 or to 100%. Their measurement errors are likely to be negatively correlated, therefore, because a positive measurement error for one proportion must be compensated by a negative measurement error for another.

Positional uncertainty

So far we have considered uncertainty only in the measured or predicted values of soil properties. In geographical information science this is known as 'attribute uncertainty'. It is different from 'positional uncertainty', which refers to uncertainty about the geographic position of spatial objects, such as sampling locations. Positional uncertainty of soil data can be substantial, particularly with legacy data arising from sampling campaigns of many years ago. One can define positional uncertainty in a similar way to attribute uncertainty using a bivariate probability density, representing uncertainty in latitude and longitude. Positional uncertainty can be translated into attribute uncertainty if short-range spatial variation is known. When a perfect measurement at location $\mathbf{s}$ is taken to represent a soil property at $\mathbf{s} + \mathbf{h}$, with $\mathbf{h}$ a possible positional error, then this short-range variation means that the property at $\mathbf{s}$ almost certainly differs from that at $\mathbf{s} + \mathbf{h}$. Therefore, even though the measurement at $\mathbf{s} + \mathbf{h}$ is free of error, it nevertheless represents the property's value at $\mathbf{s}$ with error.

Errors and uncertainty in soil class maps

For many years spatial variation in soil was represented solely by maps of soil classes, and those classes were defined by the soil in a profile regardless of any lateral variation. Errors were recognized, either because the boundaries between classes were drawn in the wrong places or because the parcels delineated contained profiles that did not belong to classes depicted, or usually both. In other words, pedologists and the clients for whom they mapped the soil were aware of the uncertainty of these maps, although it was rarely quantified.

Uncertainty in categorical soil variables such as soil type obviously cannot be characterized by models such as Eq. (1) and Eq. (3) nor can they be normally distributed. But we can form a discrete probability distribution by listing all possible values of a variable of interest at each location $\mathbf{s}$ and their associated probabilities. Clearly, these probabilities must all be non-negative and add to 1. While the definition is easy, acquiring them in practical situations for all locations in a region is more difficult. Indicator geostatistics may be used (Webster and Oliver, 2007, Chapter 11), or perhaps more advanced statistical methods such as the generalized linear geostatistical model (Diggle and Ribeiro, 2007; Kempen et al., 2012). Quantifying spatial dependence of uncertain categorical soil variables is especially difficult. Bivariate distributions are not enough, and so-called 'multi-point' statistics are required, using techniques such as multi-point geostatistics (Meerschman et al., 2013).

As a last resort one may rely on expert judgment to quantify the uncertainty of categorical soil variables (O'Hagan, 2012; Lark et al., 2015). For example, to assess uncertainty of a soil class in a pit, several experts might be asked to classify the soil independently. If they all agree then the uncertainty will be small or absent, but if they do not then the distribution of the designated classes may be used to represent the uncertainty (Van Leeuwen et al., 2018). In fact, one can also use the judgment of experts for assessing uncertainty of properties measured and recorded on continuous scales, particularly when there are no replicates of field and laboratory measurements and when soil maps are made by non-statistical methods (Truong and Heuvelink, 2013).

The uncertainty of a map of soil classes as a whole can also be represented in a confusion matrix. If there are m classes then the confusion matrix, $\mathbf{C}$, will have m rows, one for each of the mapped classes, and m columns, again one for each class. Each cell $C[i, j]$

of the confusion matrix contains the number of instances that the soil class at a location in the region that is mapped as class j is in fact class i. In practice, the confusion matrix is estimated from independent validation data, which are preferably placed randomly within the mapped region. The counts for each combination of map class and true class are then entered into the matrix. Counts in the diagonal of the matrix represent agreement; those in off-diagonal elements disagreement or error. The purity of the soil map may be expressed as the proportion of instances that the map was correct:

$$\text{purity} = \frac{1}{n}\sum_{i=1}^{n} C[i,i], \tag{4}$$

where n is the total number of observations. Note that the confusion matrix provides only a summary measure of the map accuracy. It is not spatially explicit. In contrast, indicator kriging and the generalized linear geostatistical model specify the map uncertainty for each and every location separately.

Change of support

Many users of soil maps are not interested in predictions of soil properties at points; they are much more concerned with averages of soil properties over larger areas of land. Such averages can be predicted by block kriging, as explained in the geostatistics chapter of the Encyclopedia. The blocks might be rectangular, even square, but they may take irregular shapes. They may even be as large as the entire mapped region. The geostatistics chapter also explains that the block kriging standard deviation is typically much smaller than the point kriging standard deviation, especially when the short-range variation is large. The reason for this is that within-block variation averages out when block means are predicted, and hence predicted block averages are more accurate than point predictions. So, the uncertainty of soil maps depends strongly on the support, where 'support' is defined as the area or volume over which a measurement or prediction is made. If the mapping did not include a change of support then the prediction support will be the same as the measurement support, which is usually treated as point support and can actually be either that of a single soil sample or a composite of several spread over some small area. Since uncertainty depends on the support, any database that reports the uncertainty of its soil maps must record the support of the map predictions as metadata.

In some circumstances users require predictions on smaller supports than those of the measurements, and if they do they may resort to area-to-point kriging (Kerry et al., 2012). This is rare for soil properties that are observed directly because the measurements are made on small supports. When properties are recorded or estimated by remote sensing, however, the supports can be substantially larger than users' requirements; so too can those associated with the polygons of legacy soil maps (Kerry et al., 2012). Another example is the use of area-to-area kriging in the vertical dimension, to predict soil properties for layers that differ from those from which the soil sample was taken (Orton et al., 2016). In contrast to block kriging, area-to-point kriging leads to larger standard deviations because it makes predictions over smaller supports than those of the measurements.

Spatial support and resolution

Spatial support is often thought to be the same as spatial resolution. That is unfortunate; they are not; they differ in principle. For instance, the pH maps shown in Fig. 3 have a spatial resolution of 250 $m \times 250$ m, but because the methods used to create these maps did not incorporate a change of support, the support of these maps is the same as that of the training data, that is the point support of the soil samples. The prediction interval limits shown in Fig. 3 thus refer to predictions at point support. The prediction interval would probably be narrower (smaller standard deviation) if the pH were predicted over a 250 $m \times 250$ m support, but this can be done only if the variogram of the map error is known and a geostatistical approach is used. One possibility is to apply block kriging to the residuals of a machine learning model. Szatmári et al. (2021) took this approach to quantify uncertainty of the change of soil organic carbon in Hungary between 1992 and 2010. They showed that uncertainty on large supports was substantially smaller than on small supports. Because of this they could show that the concentration of organic carbon in the soil had changed significantly between 1992 and 2010 for counties and the entire country, whereas for square blocks of 1 km $\times$ 1 km and 10 km $\times$ 10 km most predicted changes were not statistically significant, and on the point supports they found none.

Uncertainty propagation

We mentioned above the use of pedotransfer functions to predict soil properties that are difficult or time-consuming to measure from simpler basic ones. These are examples where soil information is used to compute derived properties. Soil maps are frequently used in this way—to predict by modelling crop yield, the emission of greenhouse gases, water infiltration, pesticide leaching and erosion, for example. Maps of basic soil properties are also used to derive more complex soil properties, such as soil functions and soil threats. The outputs of these models and analyses suffer from two main sources of uncertainty. One is model uncertainty, which may be split in uncertainty concerning the form of the model and uncertainty of the actual parameters of the model (e.g. Chapagain et al., 2022). Here we focus on the second one, namely uncertainty of the input variables. If the input to a model is uncertain then that uncertainty propagates into the model's output.

There are two main methods for analyzing the propagation of uncertainty through environmental models. The Taylor series method can be used for fairly simple models that are described by a continuous-differentiable mathematical function. It approximates these models locally by a linear or quadratic function, after which the propagation of uncertainty can be computed analytically (Heuvelink, 2018). The Monte Carlo method is more flexible, but it is also computationally demanding. It requires the joint probability distribution of all uncertain inputs, from which realizations or 'possible realities' are generated with a pseudo-random number generator. For spatially distributed inputs one makes use of spatial stochastic simulation, as discussed above. By repeatedly sampling from the input distribution and running the model for each realization, one obtains an equivalently sized sample from the model's output distribution. Thus, Monte Carlo simulation involves the following steps.

1. Repeat N times (typically $100 \leq N \leq 10{,}000$):
   (a) Generate a realization of each uncertain model input.
   (b) Run the model with this set of realizations and store the model output.
2. Compute parameters of the distribution of the N outputs and use these to characterize the uncertainty of the model output (e.g. mean, standard deviation, percentiles).

The Monte Carlo method is flexible in that it imposes no restrictions on the model. It can also deal with categorical uncertain inputs, whereas the Taylor series method cannot. The main difficulties are the computational burden and the need to characterize input uncertainty with a probability distribution, which can be a complex task for spatial inputs, as mentioned above. The Monte Carlo method involves an approximation error, but one can diminish this error by increasing the number, N, of Monte Carlo runs.

There are numerous applications of Monte Carlo uncertainty propagation in soil science. Some examples are Bishop et al. (2006), Van den Berg et al. (2012), and Benke et al. (2018). One can use it also to assess the contribution of individual uncertain inputs to the overall output uncertainty. One repeats the uncertainty propagation analysis with all uncertainty sources present except that for which the uncertainty contribution is required. This input is deemed certain and set to a representative value (usually the mean of the probability distribution). One then assesses the contribution of this input to the model's output uncertainty by evaluating the reduction of uncertainty that is caused by setting it to its representative value.

Uncertainty propagation example

Fig. 5 shows an example where model uncertainty and input uncertainty were propagated by the Taylor series method. The example refers to the Allier region in France shown in Fig. 4. A simple pedotransfer model that took the shape of a multiple linear regression model and that predicts the soil moisture content at wilting point from soil moisture at field capacity and soil porosity was first calibrated with data from 12 locations in the region. Model uncertainty was quantified by the variances and covariances of the regression coefficients and the residual variance; input uncertainty was derived by cokriging. Fig. 5 shows that the standard deviation of the model's output (i.e. of the soil moisture at wilting point) is smaller than that of the predicted output, although there are locations where the ratio of the two exceeds 0.5. This occurs mainly at places far from where the model inputs were measured (see top left panel in Fig. 4). The bottom panel of Fig. 5 shows that in this case the contribution of input uncertainty was much larger than that of model uncertainty. This indicates that one could obtain a more accurate map of soil moisture content at wilting point by taking more measurements of soil moisture at field capacity and soil porosity. It is remarkable that model uncertainty had a small contribution, given that the model was calibrated on only 12 paired observations.

The added value of uncertainty information

We explain above how uncertainty in soil measurements and soil maps can be quantified and used in analyses to propagate uncertainty. The quantification and propagation of uncertainty makes substantial demands for data and is not easily done. So, is it worth the effort? What is the added value of an uncertainty assessment of spatial soil information? Below we list several purposes that it serves.

1. First and foremost, any self-respecting investigator should want to check the quality of his or her results before they are made public. We should not publish soil data that have large errors nor publish poor soil maps. If they are the only data available, however, we should be honest about that and report their accuracy.
2. Clients and end users must know the quality of soil information so as to judge whether the information is good enough for their intended use of it. For example, there is serious regard for the soil's ability to store more carbon and mitigate climate change (Amelung et al., 2020). There have been claims of success in sequestering carbon, but these claims must be verified. One can do that by estimating the increase in organic carbon in a region between the start and end date of a project, though one would want to test the increase for statistical significance to be confident that the increase was real. And that would necessitate the quantification of the uncertainty.
3. Quantified uncertainty also helps producers of spatial soil information because it allows them to compare the performances of two or more mapping techniques, to determine which performs best, and to decide which technique to use in any particular case. For example, do many cheap, but inaccurate, observations from proximal sensors provide more accurate maps than few expensive measurements by wet chemistry? Does it pay to replace a simple empirical erosion model by a complex mechanistic

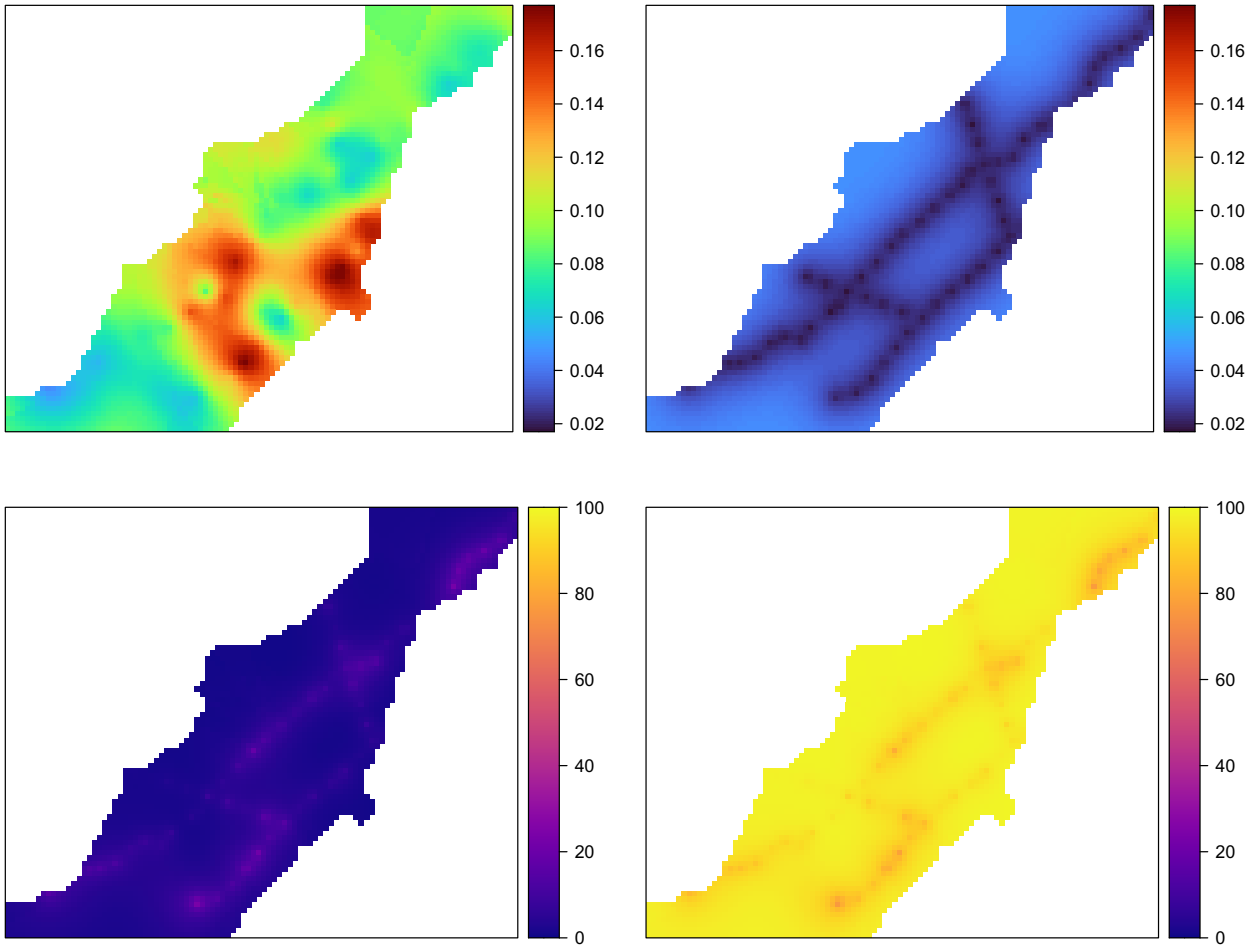

Fig. 5 Uncertainty propagation in the Allier study area. Top left: predicted soil moisture content at wilting point derived from kriged maps of soil moisture at field capacity and soil porosity using a pedotransfer function. Top right: standard deviation of soil moisture at wilting point, obtained using a Taylor series uncertainty propagation analysis where both model uncertainty and input uncertainty were included. Bottom left: Contribution (%) of model uncertainty to overall uncertainty in soil moisture content at wilting point. Bottom right: Contribution (%) of input uncertainty to overall uncertainty in soil moisture content at wilting point. See Heuvelink (2018) for details.

model that needs many more input and calibration data but could predict erosion rates more accurately? These questions can be answered only if the uncertainty is quantified.

4. Geostatistical modelling of uncertainty enables one to see how uncertainty is influenced by a change of support. The larger are the supports at which one predicts the greater is the certainty. But that in itself is not a sufficient reason to predict on large supports. The prediction must depend on the application at hand; what size of support does the user need or demand? The farmer, for example, who wishes to apply fertilizer at a variable rate would want predictions of nutrient concentrations on a support that matched the width of his fertilizer spreader. He or she would be unlikely to want predictions on point supports from which the measurements were made. In like manner, emissions of nitrous oxide from agricultural land might be estimated for fields, farms, large regions, even whole countries, if the aim is to learn how much each is contributing to the total nitrous oxide in the atmosphere. There is a trade-off between accuracy and spatial support, and knowing this relation is indispensable for informed decisions on the spatial scale at which to work and present results.
5. Uncertainty in soil data propagates through models and spatial analyses to output predictions and maps. Quantifying and tracking that uncertainty is valuable because it enables end-users to know how accurate the results are, and that in turn enables them to take into account that uncertainty when making decisions. They can do so, however, only if the uncertainty in the input, i.e. in the data, is quantified. Uncertainty propagation analysis can tell how much each of the uncertainty sources contributes to the output uncertainty. This helps modelers and users to take rational decisions on how to reduce uncertainty about the model output. It is usually best to tackle the main source of uncertainty (i.e., the weakest link in the chain), although this also depends on how easy or difficult it is to reduce input uncertainty. For instance, if an interpolated map of the rootable depth is the main source of uncertainty of an agronomic model that predicts potential yield in West Africa, it may not be easy to collect more point data on rootable depth in the region to improve the map.

6. Quantified uncertainty can be included in decision making, as in risk analysis. Many users want to avoid risk; they would like certainty. They cannot have it, of course, but they can be told what the uncertainty is, and that should help them when they are deciding what to do with the information provided. For instance, a farmer might be unwilling to buy fertilizers if plague, pests or drought are likely seriously to cause loss of his crop even if the expected net-return of fertilization is positive. Breure et al. (2022) showed how farmers can account for uncertainty in spatial decision making. Incorporating uncertainty information in spatial decision making has not yet received the attention it deserves by soil scientists, though as Lark et al. (2022) remarks there is much potential benefit for it.

Finally, note that for some purposes soil maps can be validated and their uncertainty quantified by design-based, probability sampling and statistical analysis and inference (Brus et al., 2011). It has several important and attractive properties. It requires no assumptions about the nature of the spatial distribution, is easily applied, and is backed by well-founded theory. It has its disadvantages, too. It can provide only summary measures of a map's accuracy; it cannot easily estimate the prediction errors for supports larger than those of the data; and it requires an independent probability sample from the population (i.e., the study area). It is a large subject well beyond the scope of this chapter, but it is comprehensively covered by De Gruijter et al. (2006).

Conclusion

Soil information is rarely if ever free of error. The main sources of uncertainty are measurement error, classification error, interpolation error and model error. That uncertainty in soil information can be adequately characterized by probability distributions, although such distributions might be complex and difficult to assess when spatial variation and cross-correlation between soil properties are involved. There are considerable advantages in quantifying uncertainty in soil information: it tells users whether the data are sufficiently accurate for their intended use, it helps producers of soil information to select the best mapping method, it allows analyses of how uncertainty in soil data propagates through environmental models, and it supports decision making under uncertainty.

Many soil databases provide no or only rudimentary information about the accuracy of the data they contain. This is certainly so for laboratory and field measurements, because laboratory errors are often not quantified or not copied into soil databases. Digital soil mappers take pride in routinely quantifying the prediction uncertainty of their maps, but not all of this information is transferred to soil spatial databases. There is also much to be done to communicate uncertainty information better to end-users and help them to use this information in decision making.

Soil science is continuously advancing. New theories and models are proposed and tested, and new techniques for acquiring data help us to improve the accuracy of soil information. The information will never be perfect, however; uncertainty is here to stay. Nevertheless, we have the tools to control and express uncertainty, and soil scientists and practitioners are urged to learn to use them to everyone's advantage.

References

Amelung W, Bossio D, de Vries W, Kogel-Knabner I, Lehmann J, Amundson R, Bol R, Collins C, Lal R, Leifeld J, Minasny B, Pan G, Paustian K, Rumpel C, Sanderman J, van Groenigen JW, Mooney S, van Wesemael B, Wander M, and Chabbi A (2020) Towards a global-scale soil climate mitigation strategy. *Nature Communications* 11: 5427.

Arrouays D, Grundy MG, Hartemink AE, Hempel JW, Heuvelink GBM, Hong SY, Lagacherie P, Lelyk G, McBratney AB, McKenzie NJ, Mendonca-Santos MDL, Minasny B, Montanarella L, Odeh IOA, Sanchez PA, Thompson JA, and Zhang G-L (2014) GlobalSoilMap: Toward a fine-resolution global grid of soil properties. In: Sparks DL (ed.) *Advances in Agronomy. Advances in Agronomy*, vol. 125, pp. 93–134. San Diego, CA: Academic Press.

Benke KK, Norng S, Robinson NJ, Benke LR, and Peterson TJ (2018) Error propagation in computer models: analytic approaches, advantages, disadvantages and constraints. *Stochastic Environmental Research and Risk Assessment* 32: 2971–2985.

Bishop T, Minasny B, and McBratney A (2006) Uncertainty analysis for soil-terrain models. *International Journal of Geographical Information Science* 20: 117–134.

Breure TS, Haefele SM, Hannam JA, Corstanje R, Webster R, Moreno-Rojas S, and Milne AE (2022) A loss function to evaluate agricultural decision-making under uncertainty: A case study of soil spectroscopy. *Precision Agriculture* 23: 1333–1353.

Brown JD and Heuvelink GBM (2005) Assessing uncertainty propagation through physically based models of soil water flow and solute transport. In: Anderson MG (ed.) *Encyclopedia of Hydrological Sciences*, 1181–1195.

Brus DJ, Kempen B, and Heuvelink GBM (2011) Sampling for validation of digital soil maps. *European Journal of Soil Science* 62: 394–407.

Chapagain R, Remenyi TA, Harris RMB, Mohammed CL, Huth N, Wallach D, Rezaei EE, and Ojeda JJ (2022) Decomposing crop model uncertainty: A systematic review. *Field Crops Research* 279: 108448.

Chilès JP and Delfiner P (2012) *Geostatistics. Modeling Spatial Uncertainty*, 1st edn New York: Wiley.

De Gruijter JJ, Brus DJ, Bierkens MFP, and Knotters M (2006) *Sampling for Natural Resources Monitoring*. Berlin: Springer.

Deng H, Ye M, Schaap MG, and Khaleel R (2009) Quantification of uncertainty in pedotransfer function-based parameter estimation for unsaturated flow modeling. *Water Resources Research* 45: W04409.

Diggle PJ and Ribeiro PJ Jr. (2007) *Model-Based Geostatistics*, 1st ed. New York: Springer.

Hengl T, Heuvelink GBM, and Stein A (2004) A generic framework for spatial prediction of soil variables based on regression-kriging. *Geoderma* 12: 75–93.

Heuvelink GBM (2018) Uncertainty and uncertainty propagation in soil mapping and modelling. In: McBratney AB, Minasny B, and Stockmann U (eds.) *Pedometrics*, pp. 439–461. Springer International Publishing AG.

Heuvelink GBM and Webster R (2022) Spatial statistics and soil mapping: A blossoming partnership under pressure. *Spatial Statistics* 50: 100639.

Heuvelink GBM, Brown JD, and van Loon EE (2007) A probabilistic framework for representing and simulating uncertain environmental variables. *International Journal of Geographical Information Science* 21: 497–513.

Kempen B, Brus DJ, and Heuvelink GBM (2012) Soil type mapping using the generalised linear geostatistical model: A case study in a Dutch cultivated peatland. *Geoderma* 189: 540–553.
Kerry R, Goovaerts P, Rawlins BG, and Marchant BP (2012) Disaggregation of legacy soil data using area to point kriging for mapping soil organic carbon at the regional scale. *Geoderma* 170: 347–358.
Lark RM, Lawley RS, Barron AJM, Aldiss DT, Ambrose K, Cooper AH, Lee JR, and Waters CN (2015) Uncertainty in mapped geological boundaries held by a national geological survey: eliciting the geologists' tacit error model. *Solid Earth* 6: 727–745.
Lark RM, Chagumaira C, and Milne AE (2022) Decisions, uncertainty and spatial information. *Spatial Statistics* 50: 100619.
Meerschman E, Van Meirvenne M, Van de Vijver E, De Smedt P, Islam MM, and Saey T (2013) Mapping complex soil patterns with multiple-point geostatistics. *European Journal of Soil Science* 64: 183–191.
Meinshausen N (2006) Quantile regression forests. *Journal of Machine Learning Research* 7: 983–999.
Nosrati K (2013) Assessing soil quality indicator under different land use and soil erosion using multivariate statistical techniques. Environmental Modelling and Assessment 185: 2895–2907.
O'Hagan A (2012) Probabilistic uncertainty specification: Overview, elaboration techniques and their application to a mechanistic model of carbon flux. Environmental Modelling & Software 36: 35–48.
Oliver MA (2010) *Geostatistical Applications for Precision Agriculture*. Dordrecht: Springer.
Orton TG, Pringle MJ, and Bishop TFA (2016) A one-step approach for modelling and mapping soil properties based on profile data sampled over varying depth intervals. *Geoderma* 262: 174–186.
Poggio L, de Sousa LM, Batjes NH, Heuvelink GBM, Kempen B, Ribeiro E, and Rossiter D (2021) SoilGrids 2.0: Producing soil information for the globe with quantified spatial uncertainty. *The Soil* 7: 217–240.
Stenemo F and Jarvis N (2007) Accounting for uncertainty in pedotransfer functions in vulnerability assessments of pesticide leaching to groundwater. *Pest Management Science* 63: 867–875.
Szatmári G, Pásztor L, and Heuvelink GBM (2021) Estimating soil organic carbon stock change at multiple scales using machine learning and multivariate geostatistics. *Geoderma* 403: 115356.
Takoutsing B, Heuvelink GBM, Stoorvogel JJ, Shepherd KD, and Aynekulu E (2022) Accounting for analytical and proximal soil sensing errors in digital soil mapping. *European Journal of Soil Science* 73: e13226.
Truong PN and Heuvelink GBM (2013) Uncertainty quantification of soil property maps with statistical expert elicitation. *Geoderma* 202: 142–152.
Van den Berg F, Tiktak A, Heuvelink GBM, Burgers SLGE, Brus DJ, de Vries F, Stolte J, and Kroes JG (2012) Propagation of uncertainties in soil and pesticide properties to pesticide leaching. *Journal of Environmental Quality* 41: 253–261.
Van Leeuwen MMWJ, Heuvelink GBM, Wallinga J, de Boer IJM, van Dam JC, vanEssen EA, Moolenaar SW, Verhoeven FPM, Stoorvogel JJ, and Stoof CR (2018) Visual soil evaluation: Reproducibility and correlation with standard measurements. *Soil & Tillage Research* 178: 167–178.
Van Leeuwen CE, Mulder VL, Batjes NH, and Heuvelink GBM (2021) Statistical modelling of measurement error in wet chemistry soil data. *European Journal of Soil Science* 73: e13137.
Vaysse K and Lagacherie P (2017) Using quantile regression forest to estimate uncertainty of digital soil mapping products. *Geoderma* 291: 55–64.
Viscarra Rossel RA, McBratney AB, and Minasny B (2010) *Proximal Soil Sensing*. Dordrecht: Springer.
Webster R (2000) Is soil variation random? *Geoderma* 97: 149–163.
Webster R and Oliver MA (2007) *Geostatistics for Environmental Scientists*, 2nd ed. Chicester: Wiley.
WEPAL (2022) International Soil-Analytical Exchange Programme. https://www.wepal.nl/en/wepal/Home/Proficiency-tests/Soil/Proficiency-tests/ISE.htm.