

# Rothamsted Repository Download

## A - Papers appearing in refereed journals

Lark, R. M., Chagumaira, C. and Milne, A. E. 2022. Decisions, uncertainty and spatial information. *Spatial Statistics*. 50, p. 100619.  
<https://doi.org/10.1016/j.spasta.2022.100619>

The publisher's version can be accessed at:

- <https://doi.org/10.1016/j.spasta.2022.100619>
- <https://www.sciencedirect.com/science/article/pii/S2211675322000161?via%3Dihub>

The output can be accessed at:

<https://repository.rothamsted.ac.uk/item/98998/decisions-uncertainty-and-spatial-information>.

© 31 January 2022, Please contact [library@rothamsted.ac.uk](mailto:library@rothamsted.ac.uk) for copyright queries.

Contents lists available at [ScienceDirect](https://www.sciencedirect.com)

# Spatial Statistics

journal homepage: [www.elsevier.com/locate/spasta](http://www.elsevier.com/locate/spasta)

## Decisions, uncertainty and spatial information

R.M. Lark <sup>a,\*</sup>, C. Chagumaira <sup>a,b</sup>, A.E. Milne <sup>b</sup><sup>a</sup> University of Nottingham, School of Biosciences, Sutton Bonington, Nottinghamshire, LE12 5RD, UK<sup>b</sup> Rothamsted Research, Harpenden, Hertfordshire, AL5 2JQ, UK

### ARTICLE INFO

#### Article history:

Received 6 November 2021

Received in revised form 12 January 2022

Accepted 20 January 2022

Available online 31 January 2022

#### Keywords:

Uncertainty

Decision theory

Spatial information

Value of information

### ABSTRACT

In this paper we review how the uncertainty in spatial information has been characterized. This includes both continuous predictions of spatial variables, and thematic maps of landcover classes. We contend that much work in this area has failed to engage adequately with the decision processes of the end-user of information, and that the engagement of spatial statisticians is essential to achieve this. We examine generalized measures of uncertainty, and those focussed on particular decision models. We conclude that the latter are likely to be the most fruitful, particularly if they emerge from a formal decision analysis. We outline the principles of value of information theory, and suggest that this represents an ideal framework in which to develop measures of uncertainty which can support both the rational collection of data and the interpretation of the resulting information.

© 2022 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

That spatial information, derived from data by statistical methods, is uncertain, and that this uncertainty should be accounted for when the information is used, is a piety much repeated in the scientific literature. Spatial statisticians have risen to the challenge, and most studies on spatial prediction would be regarded as incomplete without some attempt to quantify the uncertainty. However, it remains an open question whether the producer of spatial information and the user of that information communicate effectively about uncertainty, and whether the latter is enabled by the former to make decisions from uncertain information which are appropriately robust. Our concerns come from two sources. First, we have formed the strong impression that papers on

\* Corresponding author.

E-mail address: [murray.lark@nottingham.ac.uk](mailto:murray.lark@nottingham.ac.uk) (R.M. Lark).

statistical methods to produce spatial information rarely analyse the end-users' decision process with the same depth and sophistication which they bring to bear on developing the statistical model.

Second, our own experience has illustrated the difficulty of effective communication. One of us (RML) was involved some years ago in a project to support the development of a soil monitoring scheme for the United Kingdom (Black et al., 2008). This entailed interaction with diverse groups who would be users of the information: government departments, regulatory authorities and civil society organizations. One task was to agree how much uncertainty in spatial predictions and regional estimates of soil properties was tolerable for the users. We framed the question in terms of the acceptable width of prediction or confidence intervals, for variables such as soil organic carbon, metals such as zinc and plant nutrients such as available potassium. With this information, and estimates of various statistics for these properties, the plan was then to identify sample intensities, under different designs, which would support predictions and estimates with sufficient precision. Most textbooks on sampling envisage such a process to identify an adequate sample size, and it would be taken as read that if survey effort can be linked to such a measure of uncertainty then the question of how to decide on the sample size has been solved. However, representatives of the stakeholder groups did not feel able to specify the required precision in these terms. Some lacked familiarity with the underlying concepts, but even the more statistically-informed were not willing to weigh the costs of alternative survey proposals against their quality as measured by the expected width of the prediction intervals for soil properties. It occurred to RML that this reluctance was entirely reasonable.

Since then our interactions with, *inter alia*, the agricultural sector (Milne et al., 2015), nutritionists and food systems specialists (Chagumaira et al., 2021; Gashu et al., 2021), agronomists (Marchant et al., 2012), marine conservationists (Lark, 2014) and forensic scientists (Lark and Rawlins, 2008) have convinced us that the problem of effective communication about uncertain information is not simply achieved by better statistical education of stakeholders, but requires rethinking the task. This probably requires that statisticians do more to analyse the process by which information is used. In this paper we look at some approaches to the communication of uncertain information which have been used in the past, and attempt to analyse their strengths and weaknesses. We distinguish 'generalized' uncertainty measures from those focused on decisions made with the information. The former, in short, can be generated with reference only to data and the statistical model used for prediction. The latter require some analysis of the specific decision to be made, and how the information support it. It is probably already clear that we favour the latter general approach, but we think it useful to examine both in some detail to understand better why generalized measures have so often been the default.

This paper, invited to mark the anniversary of the Journal, is more of a cross between an opinion piece and a review than the usual research output which *Spatial Statistics* publishes. However, we offer it because we think that the questions which it raises should challenge readers of the journal with some substantive research questions and also to think about the gaps which may exist between the tools and methods which spatial statisticians develop and the requirements of their users, actual or notional. Many of our examples are taken from developments in soil science, from classical soil survey to recent advances in 'digital soil mapping' (McBratney et al., 2013). This partly reflects our background, but also the wide-ranging influence of the Oxford Soil Science laboratory, under the leadership of Philip Beckett, and its work on the analysis of soil information systems, their structure, costs and utilities.

## 2. Generalized uncertainty measures

### 2.1. Thematic maps: Accuracy, purity and information

A classical soil map represents soil variation in terms of mapped boundaries between parcels of land which are distinguished with respect to topography, vegetation, land use and other features which the surveyor can recognize in the field. These features are part of the surveyor's mental model of the soils and their formation in the mapped region, how soil-forming factors (such as topography)

are expressed in the landscape, and how soil properties influence features such as vegetation. This model is calibrated by observations of the soil in pits or auger borings. A simple soil map unit is dominated by a single soil profile class from a classification, for which it is usually named. When the map-user makes a decision about a location of interest the map unit delineated there provides a prediction of the soil class which is expected. If one examines the memoir which accompanies a soil map it is usual to find that the soil class provides an index to a wealth of information on land capability, for example the periods of the year when farm machinery can be used without damage to soil, whether soil pH is expected to be limiting, the thickness of rootable soil over underlying rock and so on. In addition, analytical data on soil samples can be used to estimate class means, which may serve as predicted values for locations where the delineated soil map unit corresponds to the class.

Because the soil varies at multiple scales, no simple set of boundaries can delineate uniform packages of soil, and it is accepted that one can visit a site where the delineated map unit,  $a$ , is named for dominant class  $A$  only to discover that the soil at that site corresponds to class  $B$ . Let the probability that the soil class at a location selected at random from within map unit  $a$ , after the map has been made, is found to correspond to soil class  $A$  be denoted by  $p(A|a)$ . This quantity is known as the purity of the soil map unit. If map unit  $i$ , covers proportion  $\pi_i$  of the mapped area, then the mean purity of the soil map is defined as

$$\bar{p} = \sum \pi_i p(I|i), \quad \forall i \in \mathcal{M}, \quad (1)$$

where  $\mathcal{M}$  denotes the set of all map units on the map. The mean purity is a measure of the uncertainty of predictions of the soil class made from the map. The mean purity, or the purities of the separate map units, can be estimated from a suitable independent set of validation observations, under which circumstances a confidence interval for the estimates can be obtained. A common target for soil map unit purity is around 0.85 (Young, 1976), but it is not uncommon to find that the actual purity is closer to 0.5, or even less (Beckett and Webster, 1971; Ragg and Henderson, 1980).

Map unit purity has been used as a measure of uncertainty in methodological studies of soil survey. For example, Beckett and Burrough (1971) examined variations in mean purity between soil maps produced in contrasting regions, at different cartographic scales, and by the procedures of 'free survey' and grid survey. Valentine et al. (1971) report the mean purity, and purities of individual map units, achieved in soil mapping by interpretation of airphotographs made with contrasting films (differing in their spectral sensitivity). Brus et al. (2011) discuss the sampling requirements for estimation of map-unit purity, giving examples from digital soil mapping. Western (1978) points out that the purity of a soil map is a measure of uncertainty which can be generally understood by the users of the map as well as the soil surveyor. Furthermore, an experienced soil surveyor will have a reasonable idea of what purity is achievable in different parts of the landscape, and how marginal improvements in purity are likely to decline with increasing survey effort. For this reason, average map purity or purity of individual map units may be reported in a soil survey memoir as a guide to the user about the uncertainties attached to predictions (e.g. Allison and Hartnup, 1981). These may be particularly useful if the composition of the different map units is also described (i.e. not only the proportion of inclusions, but the principle classes which comprise them).

The purity of soil map units is recognized as a practical measure of uncertainty to be used in the appraisal of soil surveys (Landon, 1984; Forbes et al., 1982), in the prescription of standards by survey organizations (Soil Science Division Staff, 2017) and in dialogue between the surveyor and map user (Western, 1978) for quality control and contract specification. As such it may be one statistical measure of the quality of spatial data with the most practical exposure.

Thematic maps of land surface classes extend beyond soil maps. The use of remote sensor data to delineate such classes is now long-established. Specialists in remote sensing have, in some respects, reinvented the ideas of map unit purity for appraising the uncertainty in thematic maps by the use of the so-called 'confusion matrix' to represent the outcome of a validation exercise (Congalton and Mead, 1983; Stehman, 1999). A confusion matrix represents a set of validation points (it is required, although not often stated, that these be selected at random and independently conditional on an appropriate design). The rows and columns of the matrix correspond to the land cover classes of

interest, as observed on the ground and the map units. If for simplicity we assume that there are as many map units as there are classes, then the matrix is square, and if the classes and map-units appear in corresponding order, then observations in cells on the main diagonal of the confusion matrix correspond to validation points where the observed class on the ground corresponds to the map-unit delineated there. For an independent random set of validation points the mean map-unit purity is therefore estimated by the ratio of the trace of the confusion matrix to the sum of its elements (Congalton and Mead, 1983).

If the classes are listed in the rows of the confusion matrix and the map units in the columns, then the  $i$ th element on the main diagonal of the matrix divided by the corresponding column total estimates  $p(I|i)$ , which we can recognize as the purity of map unit  $i$ . Story and Congalton (1986) call this the ‘producer accuracy’ of the map unit, which can be averaged over all units applying the relative areas of the map-units as weights. The  $i$ th element on the main diagonal of the matrix divided by the corresponding row total estimates  $p(i|I)$ , which Story and Congalton (1986) call the ‘user accuracy’.

We take issue with this terminology. On the face of it is laudable to attempt to distinguish between measures of uncertainty of particular interest to the users of information, which Story and Congalton (1986) appear to do. However, it is by no means clear that  $p(I|i)$  should interest the user of a map more than  $p(i|I)$ . Western (1978) and Beckett and Burrough (1971) note that a soil map might be used to predict soil conditions at one or more locations to support land use planning there, but might also be used to inventorize locations where a specified land use is possible. In the former case the purity of the map units, or particular ones, will be of interest to the user. How likely is it that the prediction at the location of interest is correct? In the latter case, however, the user of information might be concerned with both marginal probabilities from the confusion matrix. The so-called ‘producer accuracy’,  $p(i|I)$ , will be of interest to such a user: if a location is suitable for a particular land use, then how likely is it that it will be represented as such on the map? Lark (1995) developed some hypothetical examples, and showed how different users, with contrasting priorities in respect of their tolerance of errors of commission or impurities in map-units corresponding to classes of interest, and errors of commission, failure to represent an instance of a class of interest, might favour different maps produced by discriminant analysis on multispectral data, with prior probabilities for the different classes adjusted to reflect these preferences. Lark (1995) proposed that the terms ‘map-unit purity’ and ‘class representation’ be used in place of ‘user accuracy’ and ‘producer accuracy’ for  $p(I|i)$  and  $p(i|I)$  respectively. Møller et al. (2019) follow this in their assessment of a map of soil drainage classes predicted across a region from multiple spatial covariate processed by machine learning.

One problem with an assessment of uncertainty in thematic spatial information focussed entirely on single measures such as purity or representation is that they fail to account for how the composition of a map unit might be interpreted by a capable user when the map is supported by a sophisticated memoir. The conventional soil survey is, perhaps, the *locus classicus* for such a spatial information product. Bie and Beckett (1971) describe the different forms that the legend to a soil map might take. The most straightforward case is the legend of simple map units, where each map-unit is named for its dominant ‘eponymous’ class. At any location in the mapped area the predicted soil class is the eponymous class of the map-unit delineated there. Because the map unit does not have purity of 1, the class is predicted with uncertainty. However, the purity might underestimate the utility of the map, because an effective memoir should convey sufficient of the surveyor’s mental model of the soil-covered landscape that the user can make a better prediction. First, the dominant inclusions of secondary classes might be named in the memoir, so that the risk entailed in treating a site as if it were the eponymous class when an inclusion occurs there can be assessed. This risk might be small if, for the practical purpose of interest the inclusions and the eponymous class are similar. Second, the memoir might enable the user to identify when rare inclusions might be more likely at a site than the eponymous class of the delineated map unit – perhaps on the basis of physiography (if the site is at a local topographic minimum, for example), or vegetation.

Bie and Beckett (1971) make the point that many maps of soil based on classes comprise a mixture of simple map units and complex ones, the latter consisting of different classes which occur

in a complex spatial pattern which cannot be resolved at the cartographic scale of the map. As with simple map units, it might well be possible to make a better spatial prediction at a location within a complex map unit than to state that it is most likely to correspond to one of the eponymous classes, or an inclusion. This is because, while not cartographically resolved, the co-dominant classes might occur in a regular spatial pattern associated with physiography, it was for this practical reason that Milne (1936) first developed the concept of the catena for mapping the spatial pattern of soils in East Africa. If the memoir presents such a conceptual model adequately, perhaps with the assistance of block diagrams and description of other features through which the catena is expressed (vegetation, outcropping cuirasses, the frost line etc.) then the user of the map might well be able to make a refined local spatial prediction. Bie and Beckett (1971) proposed that map legends could be scored on the basis of how well the soil conditions can be predicted in sites in each legend unit, presenting these values in a type of histogram which shows the relative area of the mapped region in each unit, and the probability with which the correct soil class can be predicted at a site in the unit, depending on all the information available. This approach was adopted by Dalal-Clayton and Robinson (1992) in an appraisal of a reconnaissance soil survey of Zambia. They computed legend scores on the basis of the composition of the units (which were complex), and contrasted these with estimates of map unit purity. The difference between these, consistent with other reported studies, indicates how the complexity of the soil pattern, and the challenges of mapping that pattern, make linked but conceptually distinct contributions to the uncertainty of a thematic map.

The limitations of purity alone as a measure of prediction uncertainty have been noted elsewhere. Consider a situation in which a particular vegetation type occurs over 85% of a region by area, with the remaining area occupied by two other types (10% and 5% respectively). If one were to predict the vegetation class at a site purely as a random guess, with the probability of predicting class  $i$  equal to its relative area, i.e. to  $\pi_i$  as defined for Eq. (1), then the probability of making a correct prediction for the  $m$  items in the set of classes  $\mathcal{C}$  would be  $\sum \pi_i^2, \forall i \in \mathcal{M}$ . In our example the probability is 0.74. Let us imagine that a digital technology for mapping vegetation produces an output which agrees with 80% of a set of independent validation observations. On the face of it this is a good performance, but appears less so when compared with the probability of guessing on the basis of knowledge of the overall frequencies of the class alone. Cohen (1960) proposed an alternative statistic to purity which measures the extent to which a map exceeds this guesswork null model in predictive accuracy. This is certainly a useful statistic for comparative studies of maps produced by different methods, but is perhaps of limited use to the data user who wants to know primarily how reliable the predictions are from some particular output (Stehman, 1999).

An alternative approach is to think about a thematic map as providing information about the likely land cover classes at an unobserved site, and reducing the uncertainty which we had about that outcome before being informed of the map prediction. This might be quantified by Shannon's information. We extend the notation used so far so that  $\pi_I$  denotes the proportion of our region where landcover class  $I$  occurs, equivalently the probability that a randomly selected location in the region will be found to correspond to class  $I$ . Consider the case where, starting from a position of ignorance about the spatial distribution of a set of classes, but knowing the prior probability of finding each class,  $\pi_I; I \in \mathcal{C}$ , we are told without error that class  $J$  occurs at location  $\mathbf{x}$ . We might quantify our 'surprise' at this information from  $\pi_J$ , on the grounds that the more probable a random outcome the less surprising it is. A natural scale on which to make this measurement of information is the negative logarithm of probability, so that our degree of surprise is therefore zero for a certain event (probability of 1), undefined for an impossible event (probability of 0), and positive for some event with probability  $p \in (0, 1)$ . If the base of the logarithm is 2 then the information is measured in bits. The average information provided by an error-free observation of the class at  $\mathbf{x}$  can therefore be computed by

$$\mathcal{I}_0 = - \sum_{I \in \mathcal{C}, \pi_I \neq 0} \pi_I \log_2 \pi_I. \tag{2}$$

Now let us assume that we do not have a perfect prediction, rather we have a delineated map unit, the  $j$ th in  $\mathcal{M}$ , and we know its composition (i.e. in terms of the notation above we know

$p(I|j), \forall I \in \mathcal{C}$ . The mean information provided by an observation of the true class at a site in map unit  $j$  is given by

$$\mathcal{I}_{o,j} = - \sum_{I \in \mathcal{C}, p(I|j) \neq 0} p(I|j) \log_2 p(I|j). \tag{3}$$

The mean information provided over all the map units can be calculated as

$$\mathcal{I}_{o,m} = \sum_{j \in \mathcal{M}} \pi_j \mathcal{I}_{o,j}. \tag{4}$$

If the map is perfect then it is clear that  $\mathcal{I}_{o,m}$  is equal to zero, the actual observation of the class at a site where the delineated map unit is known conveys no additional information; the map has already conveyed it. If, on the other hand, the map units are simply random partitions of space in which, for any class  $I$  and map units  $i$  and  $j$ ,  $p(I|i) = p(I|j) = \pi_I$ , then  $\mathcal{I}_{o,m} = \mathcal{I}_o$ , the information conveyed by a truthful observation is the same whether or not we have the map; the map conveys no information. If the map units are more internally uniform than the land classes of the region as a whole, while not each corresponding entirely to a single class, then  $\mathcal{I}_m < \mathcal{I}$ . That is to say, the information conveyed by the perfect observation at a site is less when we know the delineated map unit than when we do not. The difference can be attributed to the information conveyed by the map:

$$\mathcal{I}_{map} = \mathcal{I}_o - \mathcal{I}_{o,m}. \tag{5}$$

We undertook a simulation to compare information and purity measures of map quality. Consider a confusion matrix normalized so that the column totals sum to 1 to give a map unit composition matrix,  $\mathbf{M}$ . Each column contains the proportions of the corresponding map unit in each land cover class, such that  $\mathbf{M}[g, h]$  contains the probability that a site in map unit  $h$  corresponds to class  $g$ . If the vector  $\phi_c$  contains the overall proportions of the land cover classes over the mapped region, and  $\phi_m$  contains the proportions of the map units, then

$$\phi_c = \mathbf{M}\phi_m. \tag{6}$$

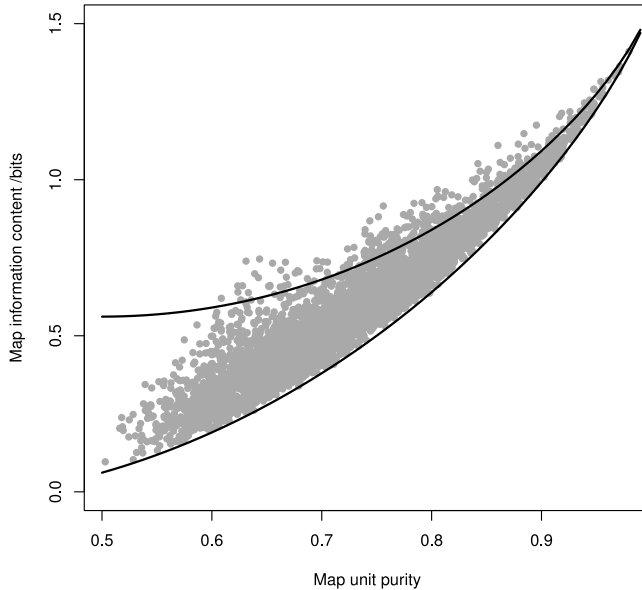
In our simulation we generated a random map unit composition matrix,  $\mathbf{M}$ , for simple map units, specifying a miming purity of 0.5 and ensuring that the column totals sum to 1. For a fixed vector of class proportions,  $\phi_c$ , the relative proportions of the different map units can be obtained from Eq. (6) by

$$\phi_m = \mathbf{M}^{-1}\phi_c. \tag{7}$$

This is a check on the consistency of  $\mathbf{M}$  and  $\phi_c$ , as it is not guaranteed that all elements of  $\mathbf{M}^{-1}\phi_c$  are non-negative. In our procedure we generated multiple realizations of  $\mathbf{M}$  subject to the rules above, for three classes and their corresponding simple map units, following the hypothetical evaluation of a soil map presented by [Giasson et al. \(2000\)](#). Those  $\mathbf{M}$  for which some elements of  $\mathbf{M}^{-1}\phi_c$  were negative were rejected. The simulated matrix, the fixed class composition vector  $\phi_c$  and the corresponding map unit proportion vector obtained as in Eq. (7) were then used to compute, for each realization, the information conveyed by the map,  $\mathcal{I}_{map}$  and the mean map unit purity. In [Fig. 1](#) the information conveyed by each notional map represented by a matrix  $\mathbf{M}$  consistent with the fixed  $\phi_c$  was plotted against the mean purity. The two quantities are clearly related, the maps with largest purity convey most information. However, for any given purity other the range of information values can be quite wide, and vice versa.

Consider a case where, with  $m$  classes delineated in simple map units the, purity is uniform over all map units and equal to  $\tilde{p}$ . Our map will convey the least information if the proportion  $1 - \tilde{p}$  of all map units is equally divided between the  $m - 1$  classes present as inclusions. It can be shown that the information conveyed by the map in these circumstances is

$$\mathcal{I}_o - \tilde{p} \log_2 \tilde{p} + (1 - \tilde{p}) \log_2 \frac{(1 - \tilde{p})}{m - 1}. \tag{8}$$



**Fig. 1.** Plot of mean purity and information conveyed by the map for simulated normalized confusion matrices. The lower and upper lines correspond, respectively, to the lower bound on information content given purity (Eq. (8)), and an upper bound on information content given a purity uniform over all map units (Eq. (9)).

This expression, evaluated as a function of  $\tilde{p}$ , is plotted in Fig. 1 where it forms a lower bound for the plotted points. We do not specify an upper bound for the plot. However, if a uniform purity is assumed, then the minimum disorder within the map unit occurs if proportion  $1 - \tilde{p}$  is occupied by a single inclusion. In this case the information conveyed by the map is

$$\mathcal{I}_0 - \tilde{p} \log_2 \tilde{p} + (1 - \tilde{p}) \log_2 (1 - \tilde{p}). \tag{9}$$

This is shown as the upper line on the plot. While most of the simulated cases fall between the two bounds, some are above the upper bound. These will be cases where the purity of individual map units vary considerably.

This approach to quantifying the information content of map units, explicitly considering their composition and internal ‘disorder’ relative to that of the landscape as a whole, seems to us to be an advance over the more common measures of uncertainty in thematic spatial information. While one might argue that the consideration of map unit composition and information theory has not helped us in our conversation with the general user of spatial information, we suggest that it has served to underline that the value of a spatial map is not entirely characterized by the binary success or failure of a spatial prediction, but rather by the extent to which our uncertainty is reduced by knowing the map unit. As seen, this requires attention to the composition of the map unit. We think that this has been overlooked, and should be given closer attention in the era of digital mapping of land cover.

At least in the field of soil science the increasing use of digital methods to predict soil conditions from multiple covariates has tended to focus on continuous properties rather than the prediction of soil classes. One reason for this could be that soil classes are seen primarily as a mechanism for the prediction of continuous variables (via the class mean, as discussed below). However, this overlooks the complex interpretative information which a classification of soil, or other aspects of land cover, can convey. Beckett and Bie (1978), after examining the application of soil surveys by diverse users in Australia, concluded that the experienced surveyor holds much more information about soil in their head by the end of a survey than is directly extractable from the map and memoir.



It is unfortunate that spatial statisticians have focussed almost exclusively on the development of methodology to predict continuous properties, assuming that the sum total of a set of such predictions would be more informative than a class-based map. A broader approach which included the use of elicitation methodology to formalize and make available the field scientist's experience alongside single property mapping might have been more productive, particularly while active and experienced field surveyors were more numerous. We share the concerns of [Murphy \(2014\)](#) that such information has been overlooked. Perhaps one reason has been insufficient attention to the requirements of the users of spatial information, and how it is required to support decision processes.

## 2.2. Prediction error variance

A first step beyond the evaluation of soil maps in terms of the prediction of discrete classes was to consider the map as a basis for prediction of continuous soil properties on the basis of class or map unit means ([Webster and Beckett, 1968](#)). On the basis of a suitable sample, one may partition the variance of a soil property into within- and between-class or map-unit components,  $\sigma_w^2$  and  $\sigma_b^2$  respectively. The intra-class correlation is the proportion of the total variance accounted for by between-class differences, effectively a correlation among sites within the same class:

$$\rho_c = \frac{\sigma_b^2}{\sigma_w^2 + \sigma_b^2}. \tag{10}$$

If there is evidence for a difference among the classes or map units with respect to a variable of interest, then the class or map-unit mean may be used as a prediction of the variable at unsampled sites. If the mean for class  $I$  is estimated by  $\bar{x}_I$  on the basis of  $n_I$  observations, and the variance within the class is  $\sigma_I^2$ , then the prediction error variance, i.e. the expected squared error of the class mean as a prediction of the variable at a location within the class is given by

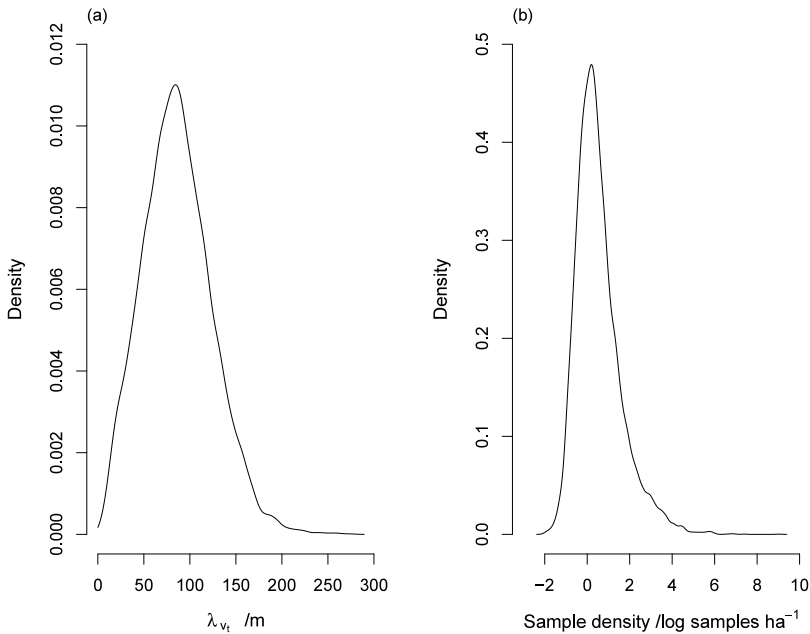
$$\sigma_p^2 = \sigma_I^2 \left( 1 + \frac{1}{n_I} \right). \tag{11}$$

The first term on the right-hand side of the above expression is simply the variability about the class mean, and the second term is the error variance for the sample mean, used as the predictor ([Webster and Lark, 2013](#); [Leenhardt et al., 1994](#)).

This prediction error variance may be computed in other settings. For example, if we have calibrated a simple linear regression to predict vegetation cover from a remotely sensed index, then the prediction error variance can be computed from the statistics of the fit for any location at which the regression is applied. The mean square error of prediction is a quantitative measure of uncertainty and, subject to assumptions about the error distribution, can be used to compute prediction intervals.

The prediction error variance of the predictions made by ordinary kriging, widely used for spatial prediction of variables from limited samples, is known as the kriging variance. Subject to the assumptions of intrinsic stationarity ([Webster and Oliver, 2007](#)), the kriging variance is the quantity minimized for any prediction by the selection of weights to form a linear combination of observations around a site where a prediction is required. In this sense a kriging prediction is optimal, but the kriging variance is also a quantification of the uncertainty. It is common to see maps of the kriging variance alongside the predictions themselves, indicating both the magnitude of the uncertainty, and also how it varies in space depending on the distribution of sample points (e.g. [Holmes et al., 2007](#); [Hatvani et al., 2021](#)). Beyond ordinary kriging, the empirical best linear unbiased predictor, from a linear mixed model which can incorporate spatial trends, or relationships to exhaustive covariates in the spatial prediction, also generates a prediction error variance alongside the predictions.

In the ordinary kriging case the kriging variance depends only on the distribution of sampling points around the location of interest and the variogram model which characterizes the spatial dependence of the random variable. For this reason, if an estimate of the variogram is available, it is possible to compute the grid spacing required to achieve a target prediction error variance, for



**Fig. 2.** Empirical probability density function for grid spacing that achieves the target kriging variance (0.18) and corresponding PDF for the logarithm of the sample density which achieves the same. From Lark et al. (2017) under CC-BY Licence <https://creativecommons.org/licenses/by/3.0/>.

example at the centre of a cell of the sampling grid where the uncertainty will generally be largest. This was proposed by McBratney et al. (1981), and has been applied, for example, in precision agriculture (Kerry et al., 2010) and soil sampling to support management of nitrogen fertilizer (Ruffo et al., 2005). When the variogram is obtained from a reconnaissance survey the error in its estimated parameters can be accounted for when planning the more detailed spatial survey through Bayesian modelling which allows a posterior distribution of the grid spacing required to achieve a target minimum precision (Lark et al., 2017). Fig. 2 shows (a) the empirical PDF for samples from the posterior distribution of sampling densities that achieve a target maximum kriging variance (0.18) for predictions of uranium concentration in topsoil of a part of the Copperbelt province in Zambia with legacy effects from copper mining (Lark et al., 2017). Fig. 2(b) shows the corresponding distribution of log sampling densities.

In many respects the kriging variance, or its relations in the wider set of prediction error variances, is an ideal measure of the uncertainty of spatial information. It can be mapped so as to visualize the uncertainty in the kriging map of the target variable. It is the variance of the prediction distribution of the variable at an unsampled site so, subject to distributional assumptions, it can be used to compute prediction intervals. Because we can express a maximum or mean prediction error variance over a region as a function of sample density, which is an important determinant of survey costs, the survey sponsor can be helped to identify an acceptable trade off between the cost of a survey and the precision of its product.

However, all this is based on the assumption that a prediction error variance can be effectively communicated to the users and sponsors of spatial information. This is by no means certain. Chagumaira et al. (2021) conducted an evaluation of alternative methods to express the uncertainty in spatial information about the concentration of micronutrients in staple grains grown in Malawi and the Amhara Region of Ethiopia. This was done with two panels, one in Malawi and one in Ethiopia, and each comprising specialists in nutrition and public health, agronomy and soil science. The panels were provided maps of micronutrient concentration along with alternative methods

to express the uncertainty. These included the kriging variance, prediction intervals, and different methods to represent the probability that the concentration fell below a nutritionally-relevant threshold. The panel were then asked a series of questions about the different presentations, and asked whether the methods helped them to understand the uncertainty, and finally to rank the methods. While a majority of respondents agreed that every method helped them to understand the uncertainty, the proportion who agreed was smallest in the case of the kriging variance (just over 60%) than for all the others (all over 80% and most over 85%). Furthermore, the overall ranking of methods across the panel appeared significantly different from random, with kriging variance ranked lowest.

### 2.3. Uncertainty 'indices'

The kriging variance is a generalized measure of the uncertainty of spatial information in the sense that it refers only to the variability of the target variable and the intensity with which it has been sampled, irrespective of the particular use to which the information is put. It can be given a specific meaning through a statistical model (e.g. a prediction interval). However, for the user unaccustomed to interpreting statistical measures of variation, it may be difficult to interpret the kriging variance with respect to a question such as whether an intervention should be approved or not at a location, or whether the variance is sufficiently large to warrant further investment in sampling to reduce it. Other general measures of uncertainty are used, often for specific types of spatial information where a formal statistical model is not used, and so a prediction variance is not produced.

For example, geologists increasingly present information inferred from boreholes, exposures and geophysical data in the form of three-dimensional models known as framework models (Lark et al., 2013). In some cases these are produced by implicit modelling methods, which entail statistical modelling. However, with limited available data, explicit modelling may be more feasible. This involves initial expert interpretation of the available data, perhaps by the production of interlocking cross-sections where point observations, at the locations of boreholes along a transect, are interpolated visually, guided by knowledge of geological processes, to produce a continuous line (perhaps with discontinuities at inferred faults) representing the height of a particular geological contact along the transect. Multiple such cross-sections may then be used to infer surfaces, perhaps with the help of an algorithm for interpolation. The final outputs of such a process comprise spatial information (the inferred elevation of a contact between geological units at locations of interest), and might be used for a range of applications such as planning engineering work or modelling the behaviour of aquifers. The uncertainty about the true position of a particular contact might have serious implications for the success or failure of a project, and we discuss this below for the case of the Channel Tunnel between France and England where geostatistical modelling was used to support decision making. In the case of explicit modelling, however, the uncertainties introduced by, for example, expert interpretation of the shape of a contact along a section, based on limited boreholes, cannot be quantified directly as part of the modelling procedure, neither can the propagation of errors of interpretation and other sources of uncertainty (such as location uncertainty for boreholes, and uncertainty about the true elevation of the ground level relative to borehole records) through the final step of interpolation to structures in 3-D. Empirical studies have examined the interpretative errors (Lark et al., 2014a,b) and the overall error in the final model surface (Lark et al., 2013), but this is a post-hoc validation which cannot be generally integrated into modelling workflows to provide the information-user with a direct measure of uncertainty.

For this reason some general indices of uncertainty have been developed, which aim to represent spatially the effects of known sources of uncertainty. Lelliott et al. (2009) proposed an uncertainty index synthesized from factors expected to contribute to the uncertainty of a framework model. The sources of uncertainty include the local geological complexity and the distance from a location under consideration to neighbouring boreholes which were interpreted as part of the modelling process. Their index was consistent with experts' expectations in a number of case studies. Although it did not directly quantify uncertainty, it did show where the user might have greatest confidence in the model and where it might be less reliable.

In other studies where uncertainty is quantifiable from a statistical model it may be visualized for the general reader with some uncertainty index. In the field of geological modelling [Zhang et al. \(2021\)](#) did this for models based on the interpretation of borehole observations. In another application area [Osgood-Zimmerman et al. \(2018\)](#), used a Bayesian geostatistical model to quantify measures of child growth failure across Africa. Their analysis provided samples from posterior distributions of interest, and they used these to produce prediction intervals based on quantiles of the posterior distribution. They presented these as measures of ‘relative uncertainty’, related to the width of this prediction interval and visualized them in a colour plot.

#### 2.4. The drawbacks of generalized measures

A mapped uncertainty index, such as that of [Lelliott et al. \(2009\)](#), may show where in a region spatial information can be treated with most confidence and where with least. Other indices, such as the width of the prediction interval, or, indeed, the kriging variance, may be used in general and visualizable measures of uncertainty (e.g. [Osgood-Zimmerman et al., 2018](#)) which reflect an underlying statistical model. However, we are sceptical that the problem of effectively communicating uncertainty is thereby solved.

This scepticism comes, in part, from anecdotal experience, but also from some established principles about decision making under uncertainty. We start with two anecdotes, offered to provoke thought, not as hard evidence. The first is from a meeting which RML attended in the Netherlands with an agency responsible for providing spatial information. In the discussion a scientist from the agency said that many of their customers confessed that measures of spatial uncertainty were regarded as *de rigueur* on any spatial information products purchased, used or presented to the public by their organizations. However, they were not aware of any cases of the customers actually using uncertainty measures in a decision-making process.

The second anecdote is from a meeting which RML and CC attended in Ethiopia with a stakeholder group interested in micronutrients and human health in Africa. The meeting was a formal consultation for co-production of a project, now underway, to construct a unified information system on micronutrient supply. One question for discussion was how uncertainty in this information should be presented. A piece of feedback was that statements about uncertainty should be treated with care because they typically do not help with decision-making, but can paralyse the process because it is far from clear how the uncertainty should be accounted for. Uncertainty might not matter if the plausible range of values are all consistent with one decision, for example.

We think that these anecdotes are informative, and are consistent with what is known by psychologists. Cognitive scientists have found a tendency to what they call regressive interpretation of information where uncertainty is indicated. [Budescu et al. \(2014\)](#) examined the ‘verbal phrases’ which the Intergovernmental Panel on Climate Change use and require to indicate the level of certainty in factual statements. For example, an outcome with a probability in the interval  $[\frac{1}{3}, \frac{2}{3}]$  is said to be *as likely as not*, while an outcome with probability larger than 0.9 is said to be *Very likely*. [Budescu et al. \(2014\)](#) found that interpretations of outcomes described as *very likely* were actually more consistent with their probability being close to 0.5. In short, drawing attention to uncertainty in information may result in the interpreter’s being over-cautious in its interpretation, underestimating its genuine information content.

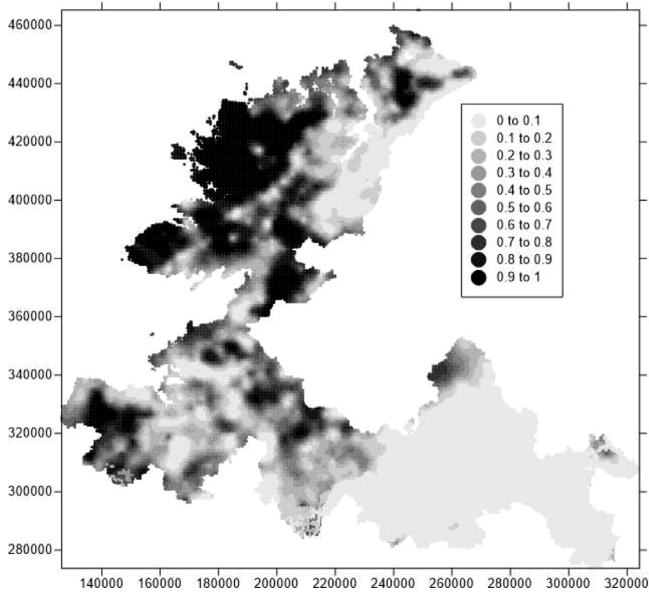
Of course the uncertainty in spatial information should not be concealed because of these difficulties. However, we do suggest that the producers of information should take note of research which suggests that psychological strategies to make decisions with uncertain information account not only for the logistical costs and benefits, but also the cognitive load of interpreting new information ([Petitet et al., 2021](#)). Generalized measures of uncertainty which do not further the decision-making process may actually inhibit it, squandering ‘cognitive bandwidth’ of the interpreter, and the reputation of the provider. A further finding of [Chagumaira et al. \(2021\)](#) may be pertinent here. It was found that presentations of uncertainty in information couched in terms of a specific decision process, in this case, the probability that the concentration of a micronutrient in grain is less than a nutritionally-significant threshold, were consistently regarded as clearer and easier to interpret than were general measures of uncertainty such as kriging variance or the prediction intervals. We therefore turn our attention to such ‘decision-focused’ measures of uncertainty for the remainder of this paper.

### 3. Decision-focused uncertainty measures

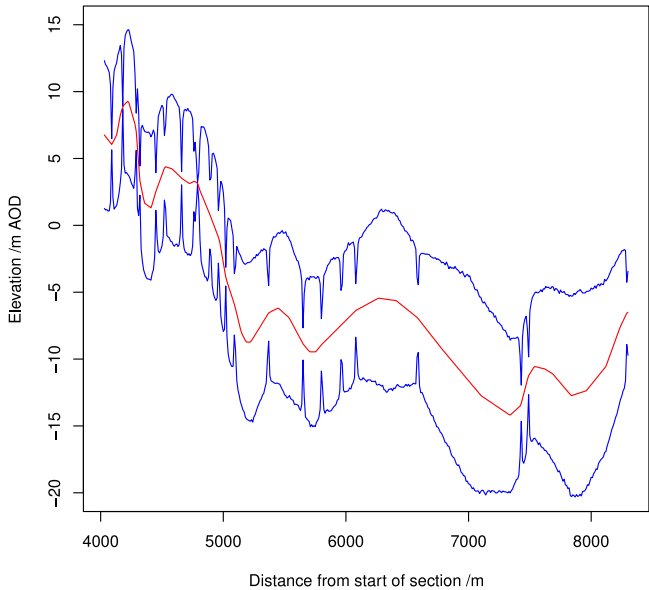
#### 3.1. Probability relative to a threshold

A common decision model for interpretation of spatial information entails comparing values with thresholds. For example, fertilizer requirement in agriculture may be based on index values of a soil analysis, while decisions on the use of land might depend on threshold concentrations of potentially harmful elements or other pollutants. Wood et al. (1990) give an example from soil management. A soil sufficiently saline that the conductivity in winter months exceeds  $4 \text{ mS cm}^{-1}$  is unsuitable for wheat production as germination and growth of seedlings will be impaired. When making regional decisions about land use, information on salinity is therefore useful. Wood et al. (1990) obtained data from Bet Shean in Israel. The point measurements of salinity had a mean value of  $4.8 \text{ mS cm}^{-1}$ . They used the disjunctive kriging method due to Matheron (1976) to obtain point predictions of conductivity across the surveyed region, but the method also estimates the probability that the value at a location exceeds some threshold to be estimated. This is the quantile of the prediction distribution which corresponds to the threshold, also known as the conditional probability. Estimates of this probability, relative to thresholds related to decisions, have also been made by indicator kriging (Journal, 1983) in which the variable is divided into an indicator,  $I \in \{0, 1\}$  at the threshold value of interest. For example, Liu et al. (2004) used the method to evaluate the probability that arsenic concentrations exceeded drinking water regulatory thresholds at locations across an aquifer in Taiwan. Conditional geostatistical simulation, in which multiple realizations of the geostatistical model, matching values at the sample locations, are drawn, may also be used for the same purpose (Goovaerts, 1997), although it is most useful when the joint probability at more than one location is required. Geostatistical methods can also be used when the decision depends on the joint values of more than one variable. For example, Lark et al. (2014a) used cokriging to estimate the joint conditional probability at sites across the north of Ireland that the value of cobalt and manganese concentration in the soil fell within a range where intervention to avoid cobalt deficiency among grazing sheep is warranted. This map is shown in Fig. 3. The same approach can be taken to presenting outputs from models which provide samples from the posterior distribution of a quantity of interest on the basis of a Bayesian model. Steinbuch et al. (2018) give an example of the prediction of the probability of finding ripened clay at locations in part of the Netherlands where marine clays are partially still subject to waterlogging which presents this process. The status of these clays has direct implications for land use decisions, and so the probabilities were presented as a map.

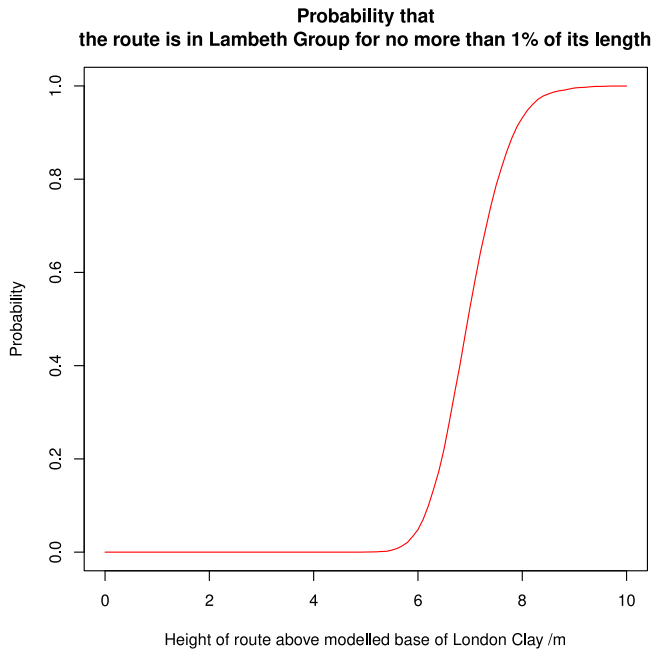
Geostatistical modelling was used for decision support during construction of the Channel Tunnel between France and the south of England (Blanchin and Chilès, 1993). The tunnel engineers aimed to keep the construction within the Chalk Marl, avoiding underlying Gault Clay and overlying altered and fractured Grey Chalk. The output of a geostatistical analysis of borehole observations of the depth to the top of the Gault Clay allowed engineers to quantify, for proposed tunnel routes, the risk of tunnelling into the Gault and to adjust their plans to reduce this risk to acceptable levels. This approach to the uncertainty in spatial information is not generic, but draws on the particular model and the particular decision which it is to be used for. Lark et al. (2014b) gave a similar hypothetical example. A modelling study was undertaken to quantify the uncertainties in expert interpretation of geological boreholes along a cross section, with respect to the depth below the surface of a contact between particular units. Fig. 4 shows a generic presentation of this uncertainty, as the 95% prediction interval for the height of the base of the London Clay relative to a datum. They then went on to consider the challenge of planning a tunnel along this section. Here engineers would want to keep the tunnel within the London Clay, and avoid the underlying Lambeth Group. If it is specified that the tunnel should intrude into the Lambeth Group over no more than 1% of its length, then we may examine the probability that this is achieved if the tunnel is planned to stay some specified height,  $k$  m, above the modelled base of the London Clay. The probability evaluated from multiple conditional simulations of the error in the model is shown in Fig. 5 as a function of  $k$ . Again, the underlying statistical model can provide the generic uncertainty measure (Fig. 4) and the specific measure of uncertainty computed for a bespoke engineering challenge and quality measure (Fig. 5). We suggest that the latter will be of greater use to the interpreter of the information.



**Fig. 3.** Probability at sites across the north of Ireland that soil analysis would indicate that an application of Co fertilizer at a rate of  $3 \text{ kg ha}^{-1}$  is required. Coordinates are in metres relative to the origin of the Irish National Grid. From Lark et al. (2014a) with a change of colour scale under CC-BY Licence <https://creativecommons.org/licenses/by/3.0/>.



**Fig. 4.** A geologist's interpretation of the base of the London Clay with 95% prediction intervals. From Lark et al. (2014b) under CC-BY Licence <https://creativecommons.org/licenses/by/3.0/>.



**Fig. 5.** How close to the modelled base of the London Clay could you build a tunnel (over the last 4 km of the cross-section) and have a specified probability (ordinate) that the tunnel strays into the underlying Lambeth Group for no more than 1% of its length? From Lark et al. (2014b) under CC-BY Licence <https://creativecommons.org/licenses/by/3.0/>.

We have shown examples in which, given a specific decision, and threshold-based decision model, measures of uncertainty attached to outcomes of the decision can be expressed as a probability that the objective of the decision model is achieved. While this is more targeted to the requirements of the user than a general uncertainty measure, it might be reasonably asked whether the communication of this uncertainty has been satisfactorily achieved. The two examples in the last paragraph are in settings where engineers are using spatial information to make decisions. One might expect engineers, through their professional background, to be competent in handling probabilistic accounts of uncertainty, although a recent study (Kaplar et al., 2021) found that misconceptions about probability are surprisingly frequent among engineering students. It is certainly the case that misconceptions about probability are widespread (Galesic and Garcia-Retamero, 2010; Spiegelhalter et al., 2011). For this reason the Intergovernmental Panel on Climate Change (Mastrandrea et al., 2010) requires that report authors use a standard verbal scale to attach probabilistic information to uncertain outcomes. If the probability of an outcome is  $P_o$  and  $P_o \geq 0.66$  then the outcome is described as 'Likely'. This may be intensified to 'Very likely' if  $P_o \geq 0.9$  and to 'Virtually certain' if  $P_o \geq 0.99$ . If  $0.33 \leq P_o < 0.66$  then the outcome is described as 'About as likely as not'. If  $P_o < 0.33$  then the outcome is described as 'Unlikely'. This may be intensified to 'Very unlikely' if  $P_o < 0.1$  and to 'Exceptionally unlikely' if  $P_o < 0.01$ . Lark et al. (2014a,b) used this scale to present the probabilistic information about soil cobalt and manganese content, and decisions on pasture soil management (presented as simple probabilities in Fig. 3 of this paper).

The IPCC verbal scale has been investigated by psychologists, and criticized as a result. Harris and Corner (2011) showed that the interpretation of a verbal uncertainty for some outcome was prone to 'severity bias', that is to say, the severity of a possible outcome from a decision is not separated in the interpretation of information from the uncertainty that the outcome would result. We have already noted the work of Budescu et al. (2014) on regressive interpretation; they showed that this effect was smaller when numeric information was combined with the verbal 'calibrated

phrase'. Lark et al. (2014a,b) used this in their presentation. However, the question of how best to use and communicate probability-based uncertainty measures remains open. Jenkins et al. (2018) found that mixed-formats (verbal and numerical) were prone to the 'extremity' effect as were purely verbal ways of expressing probability. By 'extremity' effect is meant the tendency to over-interpret a statement that an outcome is 'unlikely' by placing it either entirely outwith the range of a histogram of outcomes, or at least in very extreme tails. This was mitigated to some extent by a mixed format for communication in which the numeric part of the information is emphasized, e.g. by placing it before the verbal expression. Jenkins et al. (2019) also found that emphasizing the numeric value of probabilities in communication made the credibility of an information source more robust to events where an outcome, said to be 'unlikely' (but not impossible), actually happens. We also note that Chagumaira et al. (2021) found no evidence that the stakeholders they engaged with preferred formats in which probabilistic information was communicated verbally, or with pictograms, to direct presentation of the numbers. It is clear that more work needs to be done before the question can be settled.

One issue addressed by Chagumaira et al. (2022), was how stakeholders would use probabilistic information to make a decision. They asked stakeholders to select a critical probability (that micronutrient supply falls below a significant threshold) above which they would recommend an intervention. It was found that this threshold was prone to a 'framing' effect (Tversky and Kahneman, 1981), such that if the question is put 'negatively' (*the probability that the micronutrient supply is less than the threshold is  $p$* ) that interpreters would be more cautious and likely to recommend intervention than if presented with the equivalent information in a 'positive' framing (*the probability that the micronutrient supply exceeds the threshold is  $1 - p$* ). It was found that a negative framing also resulted in more consistent interpretation over different professional groups. The critical value of  $p_t$  under a negative framing, was about 0.3, indicating a tendency to a conservative interpretation. One way to interpret  $p_t$  is in terms of the losses that a user perceives under errors of commission,  $L_c$ , (intervene where not necessary) or omission,  $L_o$ , (fail to intervene where it is necessary). It is straightforward to show that, if these two losses are assumed to be equal that setting  $p_t = 0.5$  minimizes the expected loss. If the loss under an error of omission exceeds that under an error of commission (e.g. because the public health costs of a failure to address nutrient deficiency outweigh the opportunity costs of an unnecessary campaign), then the expected loss is minimized by recommending an intervention if the probability that the nutrient supply is inadequate exceeds some  $p_t < 0.5$ , in fact, a result in decision theory is that

$$p_t = \frac{L_c}{L_o + L_c}. \quad (12)$$

In short, decision theory links the interpretation of probabilistic information about a threshold to, at least tacit, measures of loss under different outcomes from the decision. This invocation of the notion of a loss or cost, which it is assumed is minimized in rational decision making, as described in standard texts on decision theory such as those of Berger (1980) and Peterson (2017), takes us to the final section of this paper.

### 3.2. Expected loss and value of information

We have already referred to the Soil Survey of England and Wales's Sheet SE39 (Soils of Northallerton) made by Allison and Hartup (1981). The largest single map unit on this sheet is a compound unit, the Clifton–Dunkswick map unit, comprised 40% and 35% respectively of the two eponymous soil series, and with small inclusions of five other soil series. If we assume that the 25% of the area in this map unit, belonging to minor soil series is equally distributed among them, then the expected information content of a random observation of the soil series somewhere in the map unit, in the sense of Eq. (2), is 1.92 bits. If, through considerable effort, this compound map unit were replaced by two simple map units, each of purity 0.8, and with the remaining area of the map unit occupied by 6 soil series, of equal area, then the expected information content of an observation is reduced to 1.24 bits. This tells us something about the uncertainties of predicting soil series. However, if we inspect the analytical data for the dominant series in the map unit, presented by



Allison and Hartnup (1981), and also by Jarvis et al. (1984), we see that the mechanical properties of the topsoil are very similar, so that the period of time in a typical year when the soil can be worked by machinery differs little between them. Similarly, both have the same (moderate) organic content in the topsoil. However, there is a difference in pH, such that the Dunkeswick series (topsoil pH 6.3) would be recommended for liming for optimal performance of clover-based pasture systems, whereas the Clifton series (topsoil pH 6.9) would not – see Lark et al. (2016) for a discussion of the threshold. The point that this example makes is that, whilst there is genuine information obtained by producing the hypothetical simple map units for the Clifton and Dunkeswick series (0.68 bits, indeed), this information is immaterial to a map user concerned with estimating soil carbon stocks or evaluating impacts on land trafficability under climate change. However, the additional information is valuable to a map user who wants to calculate lime requirement to improve pasture over a region. This example illustrates the difference between information conceived as an abstract reduction of uncertainty, and information conceived as a way of reducing the costs arising from uncertainty. Howard (1966), building on the statistical decision theory of Raiffa and Schlaiffer (1961) introduced the idea of the value of information as an explicit alternative to Shannon's information theory. He stated that, under Shannon's approach, or others based on the probabilistic structure of the communicated variable, 'if losing all your assets in the stock market and having whale steak for supper have the same probability then the information associated with the occurrence of either event is the same'. For this reason, he said, '[a]ttempts to apply Shannon's information theory to problems beyond communications, have, in the large, come to grief.'

Bie and Ulph (1972) evaluated the potential economic benefit from the information provided by soil survey on the basis of the expected pay-off, both positive and negative, to different agricultural practices on the constituent soils of a region, and the benefit to be obtained if the area under the different soils can be distinguished. This allowed them to identify an optimal purity for map units, on the assumption that the marginal costs of making a map of the optimal purity equal the marginal benefits from basing decisions on it. By putting the costs of information in the same terms as the benefits from its use, Bie and Ulph (1972) were able, in effect, to push the question of how one interprets some measure of information uncertainty into the background, and replace it with a more tractable model of information as an input to an activity which, much like fertilizer, labour or fuel, which contributes to the net return to activities which it supports.

Assessments of the economic value of spatial information have been made in the context of geostatistical surveys of the soil. Viscarra Rossel et al. (2001) considered the case of site-specific lime application to a wheat crop, and showed that the rational decision in the case of their study site was to apply lime uniformly to a field rather than attempting to obtain the information needed to apply lime at a variable rate. We use the basic model of wheat responses to pH to illustrate the concept of value of information in statistical decision analysis.

In Fig. 6(a) we show the function which Viscarra Rossel et al. (2001) used to relate soil pH to yield of the wheat crop. The maximum yield is small in global terms, but not for the extensive dryland systems of Australia. It is a quadratic function, and so is symmetrical, with a loss of yield at excessive pH (attributable, among other factors to the 'locking up' of micronutrients such as iron), and yield losses at more acidic pH due to effects on microbial activity, nitrogen cycling and root development. We assume that the goal of soil management is to achieve the optimum pH (7.4), and that the actions taken to correct excess acidity are done on the basis of soil pH information at field scale. We assume that the management variable is continuous (e.g. applying some specified rate of lime). We therefore assume that if, for example, pH is over-estimated, lime will be underestimated accordingly, and so the pH achieved will be somewhere below the optimum. Similarly, if pH is underestimated, lime will be over-applied, and the pH of the soil will be correspondingly too high. We recognize that this simplifies real management practice.

We use the response function in Fig. 6(a) to obtain a loss-function, Fig. 6(b) expressing the cost of an error in the estimation of soil pH on the assumption that an error results in adjustment of soil pH to a non-optimal value, resulting in a loss of potential yield, and assuming that a tonne of wheat is worth A\$370 (Australian dollars, based on information from the <https://www.indexmundi.com/commodities/> website for September 2021). This is an oversimplification, because, among other considerations it does not account for how liming the soil will have benefits, or disbenefits,

for several cropping seasons after application. However, for illustrative purposes we work with this simple loss function. Let us assume that, in the absence of information about the pH of the soil in a field where we are managing a crop of wheat, we use a regional mean value. If the soil pH is overestimated by 0.5 pH unit by the regional mean then the loss is about A\$ 23 ha<sup>-1</sup>. This means that perfect information about the pH of our particular field would be worth A\$ 23 ha<sup>-1</sup>, it would save the manager that loss. Because this ignores error in the information it is sometimes called the ‘value of perfect information’ or ‘value of clairvoyance’ (VoC), and we use the latter term here. If we specify the error by  $\varepsilon$ , and the loss function by  $L(\varepsilon)$ , then, if the mean pH values of soil in a region has a normal distribution with standard deviation  $\sigma_R$ , then the expected loss over the region, and so the expected VoC, can be computed as

$$E[\text{VoC}] = E[L|\sigma_R] = \int_{-\infty}^{\infty} L(\varepsilon) f_{\mathcal{N}}(\varepsilon|0, \sigma_R) d\varepsilon. \tag{13}$$

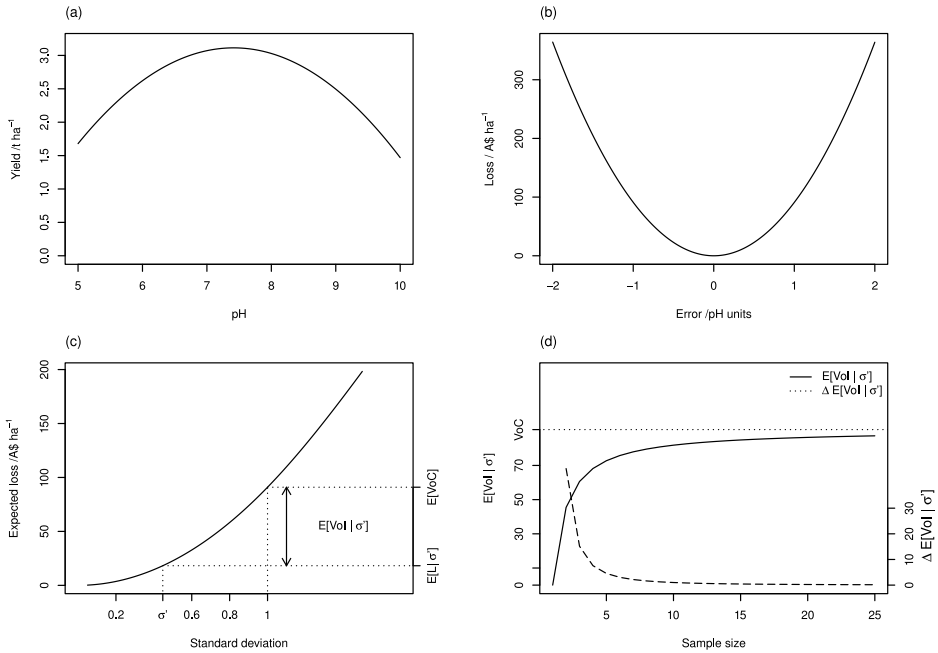
Now, rather than clairvoyance, let us assume that the soil pH in a field of our region is determined by a sample. If the sample is unbiased, with standard error  $\sigma'$ , then we can compute the expected loss when the estimate is used to manage soil pH, and this will be obtainable from Eq. (13) as  $E[L|\sigma']$ . Whereas clairvoyance reduces the expected loss from  $E[L|\sigma_R]$  to zero, so the value of clairvoyance is equated with the expected loss, with sampling the expected loss is reduced from  $E[L|\sigma_R]$  to  $E[L|\sigma']$ . This is illustrated in Fig. 6(c) which shows  $E[L|\sigma]$ , as a function of the standard deviation,  $\sigma$ , and the expected value of the imperfect information, or  $E[\text{VoI}|\sigma']$  as the difference between  $E[\text{VoC}]$  and  $E[L|\sigma']$ .

In Fig. 6(d) we compute the expected value of imperfect information as a function of the sample size used to obtain that information. We assume that the standard deviation of the field mean pH,  $\sigma_R$  is 1 pH unit, and the within-field standard deviation of soil pH,  $\sigma_W$ , is also 1. These values are compatible with coefficients of variation and variograms for soil pH presented by Beckett and Webster (1971) and by Patterson et al. (2018). If we assume simple random sampling with  $n$ , then the standard error of the sample mean is

$$\sigma' = \frac{\sigma_W}{\sqrt{n}}.$$

The solid line in Fig. 6(b) shows the value of imperfect information increasing as the sample size increases, and so the information becomes more precise, but also more costly. The dashed line in the plot shows the change in the value of expected information for successive unit increases in sample size, with diminishing returns apparent. The best sample size to be used would be the one where the increment in the value of the imperfect information just equals the marginal unit cost of the soil information. In this case, for example, if we assume that the marginal cost of an additional soil pH sample from a field is A\$3, then the best sample size for a field is 6 samples. If the marginal cost is reduced to A\$0.5, then a sample size of about 14 would be preferred. Note, again, that these examples are given for illustrative purposes, and do not constitute specific advice (see Fig. 6).

The paragraphs above introduce the elementary ideas of the theory of value of information, as developed since the work of Howard (1966). The key ideas are the loss function, the value of clairvoyance (or perfect information) in some particular case, the expected value of clairvoyance, and the expected value of imperfect information. The study of value of information can generally be developed in one of two directions. The first is a detailed decision analysis, following the principles of decision theory (e.g. Peterson, 2017) in which decisions or acts have different outcomes in uncertain states – e.g. a *decision/act* whether or not to spray a field crop against a fungal disease has *outcomes* (costs incurred, or saved, yield lost, or saved) according to the *state* (a pathogen is present or not). Our example above is a very simplified illustration, but more complete analyses related to sampling decisions on contaminated land have been produced, notably by Ramsey et al. (2002), Boon et al. (2011). Here the decision is how intensively to sample a site proposed for development. The outcome of the sampling action will be a choice either to proceed to development of the site, or first to remediate it. Depending on the underlying state of the site, there may be substantial unnecessary costs associated with remediation where this is not required, or high costs from remediation post-development, with possible associated fines and legal liabilities. A loss function



**Fig. 6.** (a) The effect of soil pH on wheat yields in Australia, following [Viscarrá Rossel et al. \(2001\)](#). (b) The corresponding loss function if attempts are made to adjust field mean soil pH to the optimum on the basis of an estimated soil pH with specified error. (c) Expected loss as a function of standard error of the estimate of soil pH and showing the expected Value of Clairvoyance (VoC) if the standard deviation of field mean pH is 1 unit and for a specified standard error of the estimate of the mean soil pH of a field ( $\sigma'$ ). (d) Expected value of information (solid line) for the field mean soil pH determined by simple random sampling as a function of sample size assuming a within-field standard deviation of soil pH of 1 unit. The dotted line shows the VoC. The dashed line shows the increment in expected VoI for successive additional samples.

can be produced from this analysis, and then the expected loss, on the basis of samples of different size (and cost) can be computed as in our illustration.

We suggest that there is a need for more worked examples of detailed decision analysis to allow the interpretation of uncertain information by assessing its value, and by extension the possible additional value of further information in particular settings. However, it might not always be possible to form a sufficiently complete loss function. Consider, for example, the case of a public research organization proposing to map the soil carbon stocks of a region. This information will be invaluable in decisions on carbon accounting, and policies to promote carbon sequestration (see [Lark and Knights, 2015](#)). However, it is difficult fully to analyse the decision because the losses due to errors in the information may be hard to quantify, reflecting intangibles such as political capital, confidence of the general public, and knock-on effects on the credibility of other interventions. For this reason, [Lark and Knights \(2015\)](#) proposed the implicit loss function as a tool for exploring sampling decisions. The implicit loss function is a concept used to analyse decisions that have been made. For example, [Pierdzioch et al. \(2016\)](#) developed implicit loss functions to quantify how members of the Federal Open Market Committee (FOMC) of the US Federal Reserve appeared to weight different outcomes of decisions on interest rates (unemployment, balance of payments, inflation etc.). [Lark and Knights \(2015\)](#) proposed a framework based on piece-wise linear loss functions, with asymmetry reflecting an elicited preference for policy outcomes emphasizing environmental goods over against those which emphasize business or other goods which must be traded off against soil quality. With this established, a cost model for survey and a proposed sampling budget the implicit loss function, the loss function which makes the sampling budget

rational, can be computed and exhibited. It is proposed that this could be used to discuss whether the costs associated with erroneous information seem plausible, or perhaps imply that a larger or smaller budget should be considered. In particular it was suggested that implicit loss functions could be compared between competing projects in order to decide on a partition of resources between them.

So the assessment of value of information through statistical decision analysis can be a frontal attack on a problem about how to assess the uncertainty in environmental information. Given that statisticians can commonly relate measures of information quality (e.g. kriging variance) to measures of the effort required to obtain that information (e.g. sample grid spacing), we can find specific answers to questions about how much information it is rational to collect in a particular case. It can be turned on its head in the implicit loss function to try to illuminate decisions about environmental information when the full decision analysis appears to be intractable.

A third approach, which may be generally illuminating, is the formal study of value of information to identify general conclusions which may be instructive, and may highlight where close attention to decision analysis in practice might be focussed. For example, three general conclusions of value of information theory are that the value of clairvoyance is an upper bound for the value of imperfect information; that the value of information is zero if decisions are never changed on the basis of information, and that the expected value of information is non-negative (Clemen, 1996). However, it has been pointed out that these results assume that the sponsor of information and the decision-maker are either the same agency, or that the latter, in effect, bases their decisions on the same model (e.g. set of threshold concentrations acceptable for a pollutant) which the sponsor assumes when making decisions on information collection. This has been called, formally, the case of value of information with Control (Vol-C), i.e. control of subsequent decisions. In practice, however, the collection of information may be sponsored by one agency, (which makes decisions on information collection), and subsequent decisions are made by another. From the perspective of the sponsor any assessment must consider what von Winterfeldt et al. (2012) call value of information without control (VOI-NC). Peck and Richels (1987) distinguished these cases, considering a situation where a scientific agency evaluates evidence for damage to the environment by acid deposition, whilst a second agency, the regulator, make decisions on how to limit activities which cause the deposition. These two agencies may value environmental goods and costs to industry identically, but hold different prior probabilities with respect to the question of whether deposition causes damage. In these circumstances the Vol-NC for the science agency can be negative, a positive value of Vol-NC is most likely if the quality of the research is large (i.e. if the probability of providing evidence for damage in the case that such damage is occurring is large, in frequentist terms if the power to detect the effect is large). von Winterfeldt et al. (2012) went on to show that, in environmental studies, the Vol-NC can be analysed as the sum of a component due to possible reduction of environmental harm and a component due to policy costs. Both components are positive when research quality (as measured above) is large and possible environmental impacts are small, but can be negative in other conditions. Even when the science agency's prior probability that there is an environmental impact is zero or one, Vol-NC can be positive (despite the research costs), because high-quality research still influences the decision maker. In short, the Vol-NC or Vol-C frameworks can provide a basis for understanding how uncertain information is used, and so its value.

#### 4. Conclusion: Next steps

We started with the observation that a statistical model may permit a sophisticated and robust quantification of the uncertainty in spatial information. However, if we stop with this quantification then it is unlikely that users of information, and the wider stakeholder groups which they serve, will benefit from it. The statistical model is only the start.

Furthermore, the challenge is about more than communication. While there are very real questions about how the user of a map might be helped by legends, colour schemes, pictograms and other tools to grasp, for example, probabilistic measures of uncertainty, these generally take their meaning only in the setting of a particular decision system, with particular decisions to be made, dependent on particular unknown states, and with outcomes which may vary from one context to

another. The generalized measure of uncertainty is unlikely to help the user with this, and indeed may, at best, be purely decorative and at worst an obstacle to the effective use of spatial information.

In fact, there is no substitute for the hard work, for all involved, of a proper analysis of the decision system which users of spatial information employ, whether explicitly or tacitly. On the basis of the work reviewed above, the key steps are as follows

1. To identify who is sponsoring the collection of information, who is making decisions with it, and who the other stakeholders are.
2. To assess whether these parties can be treated as a single representative stakeholder, with common prior beliefs about the states, and shared interests in the outcomes (be these regulatory costs, social costs, environmental costs), or whether different stakeholders must be recognized. In this latter case the sponsor, who will make decisions about the effort put into collection of information, has a Vol-NC, but the decision maker may value the information differently.
3. To characterize the states, decisions and outcomes which constitute the decision problem. It is likely that many decisions depend on states of multiple variables, and a careful analysis of how these should be considered in combination is needed.
4. To identify the decision-making process, the relevant stakeholders, what risks are carried and by whom.
5. To characterize the decision-making process by one or more loss functions for errors in information, or, if this is not possible, to identify simpler proxies to characterize acceptable information quality.
6. To base on the information collected above, and on any existing data, a robust sampling and mapping strategy which should provide information of acceptable quality for the decision-making process.
7. To evaluate the quality of the information *post-hoc*, and to assess how far the requirements of different stakeholders have been met.

We suggest that the success of this process rests in large part on the engagement of statisticians throughout; much as the statistical profession might prefer to draw sharp demarcation boundaries, such that its responsibilities end with the successful and rigorous completion of a statistical model and computation of its predictions and their uncertainties. We also accept that such a complete and closed decision system might often not occur, for example when data are collected to establish national or regional capability and without a complete view as to who might be beneficiaries of the information further down the line. That said, we agree with [von Winterfeldt \(2014\)](#) that the best efforts to characterize a decision system will not be wasted, even if they do not succeed in delivering a full decision analysis, insofar as they allow stakeholders and scientists to sharpen their understanding of their objectives and assumptions. Some seventy years ago [Yates \(1952\)](#) recognized the urgent need to develop best practices to determine rational levels of investment in the collection of information, it is high time that this was fully addressed in respect of spatial information.

## Acknowledgements

CC is supported by a studentship in the University of Nottingham Future Food Research Beacon and Rothamsted Research programme in International Agricultural Development.

## References

- Allison, J.W., Hartnup, R., 1981. Soils in North Yorkshire VI. Sheet 39 Northallerton. Soil Survey Record No. 68, Soil Survey of England and Wales, Harpenden.
- Beckett, P.H.T., Bie, S.W., 1978. Use of Soil and Land-System Maps to Provide Soil Information in Australia. Division of Soils Technical Paper No. 33, CSIRO, Melbourne.
- Beckett, P.H.T., Burrough, P.A., 1971. The relation between cost and utility in soil survey. IV. Comparison of the utilities of soil maps produced by different survey procedures and to different scales. *J. Soil Sci.* 22, 460–480.
- Beckett, P.H.T., Webster, R., 1971. Soil variability: a review. *Soils Fertil.* 34, 1–15.
- Berger, J.O., 1980. *Statistical Decision Theory, Foundations, Concepts and Methods*. Springer-Verlag, New York.

- Bie, S.W., Beckett, P.H.T., 1971. Quality control in soil survey introduction: I. The choice of mapping unit. *J. Soil Sci.* 22, 32–49.
- Bie, S.W., Ulph, A., 1972. The economic value of soil survey information. *J. Agric. Econ.* 23, 285–297.
- Black, H., Bellamy, P., Creamer, R., Elston, D., Emmett, B., Frogbrook, Z., Hudson, G., Jordan, C., Lark, M., Lilly, A., Marchant, B., Plum, S., Potts, J., Reynolds, B., Thompson, P., Booth, P., 2008. Design and Operation of a UK Soil Monitoring Network. Science Report – SC060073, Environment Agency, Bristol.
- Blanchin, R., Chilès, J.P., 1993. The Channel Tunnel: geostatistical prediction of the geological conditions and its validation by the reality. *Math. Geol.* 25, 963–974.
- Boon, K.A., Rostron, P., Ramsey, M.H., 2011. An exploration of the interplay between the measurement uncertainty and the number of samples in contaminated land investigations. *Geostand. Geoanal. Res.* 35, 353–367.
- Brus, D.J., Kempen, B., Heuvelink, G.B.M., 2011. Sampling for validation of digital soil maps. *Eur. J. Soil Sci.* 62, 392–407.
- Budescu, D.V., Por, S.B., Smithson, M., 2014. The interpretation of IPCC probabilistic statements around the world. *Nature Clim. Change* 4, 508–512.
- Chagumaira, C., Chimungu, J.C., Gashu, D., Nalivata, P.C., Broadley, M.R., Milne, A.E., Lark, R.M., 2021. Communicating uncertainties in spatial predictions of grain micronutrient concentration. *Geosci. Commun.* 4, 245–265.
- Chagumaira, C., Nalivata, P.C., Chimungu, J.G., Gashu, D., Broadley, M.R., Milne, A.E., Lark, R.M., 2022. Stakeholder interpretation of probabilistic representations of uncertainty in spatial information: an example on the nutritional quality of staple crops. *Int. J. Geogr. Inf. Sci.* in press.
- Clemen, R.T., 1996. *Making Hard Decisions: An Introduction to Decision Analysis*. Duxbury Press, Belmont.
- Cohen, J.A., 1960. Coefficient of agreement for nominal scales. *Educ. Psychol. Meas.* 20, 37–46.
- Congalton, R.G., Mead, R.A., 1983. A quantitative method to test for consistency and correctness in photointerpretation. *Photogramm. Eng. Remote Sens.* 49, 69–74.
- Dalal-Clayton, D.B., Robinson, D.A., 1992. An assessment of the success of a geomorphologically based reconnaissance soil survey in eastern Zambia. *Geoderma* 54, 213–230.
- Forbes, T.R., Rossiter, D.G., VanWambeke, A., 1982. Guidelines for Evaluating the Adequacy of Soil Resource Inventories. Cornell University Department of Agronomy, New York.
- Galesic, M., Garcia-Retamero, R., 2010. Statistical numeracy for health: a cross-cultural comparison with probabilistic national samples. *Arch. Internal Med.* 170, 462–468.
- Gashu, D., Nalivata, P.C., Amede, T., Ander, E.L., Bailey, E.H., Botoman, L., Chagumaira, C., Gameda, S., Haeefe, S.M., Hailu, K., Joy, E.J.M., Kalimbara, A.A., Kumssa, D.B., Lark, R.M., Ligowe, I.S., McGrath, S.P., Milne, A.E., Mossa, A.W., Munthali, M., Towett, E.K., Walsh, M.G., Wilson, L., Young, S.D., Broadley, M.R., 2021. Cereal micronutrient quality varies geospatially in Ethiopia and Malawi. *Nature* 594, 71–76.
- Giasson, E., van Es, C., van Wambeke, A., Bryant, R.B., 2000. Assessing the economic value of soil information using decision analysis techniques. *Soil Sci.* 165, 971–978.
- Goovaerts, P., 1997. *Geostatistics for Natural Resource Evaluation*. Oxford University Press, New York.
- Harris, A.J.L., Corner, A., 2011. Communicating environmental risks: Clarifying the severity effect in interpretations of verbal probability expressions. *J. Exp. Psychol.: Learn. Mem. Cogn.* 37, 1571–1578.
- Hatvani, I.G., Szatmári, G., Kern, Z., Erdélyi, D., Vreča, P., Kanduč, T., Czuppon, G., Lojen, S., Kohán, B., 2021. Geostatistical evaluation of the design of the precipitation stable isotope monitoring network for Slovenia and Hungary. *Environ. Int.* 146, 106263.
- Holmes, K.W., Van Niel, K.P., Kendrick, G.A., Radford, B., 2007. Probabilistic large-area mapping of seagrass species distributions. *Aquat. Conserv.* 17, 385–407.
- Howard, R.A., 1966. Information value theory. *IEEE Trans. Syst. Sci. Cybern.* 1, 22–26.
- Jenkins, S.C., Harris, A.J.L., Lark, R.M., 2018. Understanding ‘unlikely (20% likelihood)’ or ‘20% likelihood (unlikely)’ outcomes: The robustness of the ‘extremity effect’. *J. Behav. Decis. Mak.* 31, 572–586.
- Jenkins, S.C., Harris, A.J.L., Lark, R.M., 2019. When unlikely outcomes occur: the role of communication format in maintaining communicator credibility. *J. Risk Res.* 22, 537–554.
- Journal, A.G., 1983. Non-parametric estimation of spatial distributions. *Math. Geol.* 15, 445–468.
- Kaplar, M., Lužanin, Z., Verbić, S., 2021. Evidence of probability misconception in engineering students—why even an inaccurate explanation is better than no explanation. *Int. J. STEM Educ.* 8 (18).
- Kerry, R., Oliver, M.A., Frogbrook, Z.L., 2010. Sampling in precision agriculture. In: Oliver, M.A. (Ed.), *Geostatistical Applications for Precision Agriculture*. Springer, Dordrecht, pp. 35–63.
- Landon, J.R., 1984. *Booker Tropical Soil Manual*. Longman.
- Lark, R.M., 1995. Components of accuracy of maps with special reference to discriminant analysis on remote sensor data. *Int. J. Remote Sens.* 16, 1461–1480.
- Lark, R.M., 2014. Mapping Seabed Sediments of the Swallow Sand and South-West Deeps (West) MCZs. British Geological Survey, Nottingham, UK, p. 18pp, (OR/14/015).
- Lark, R.M., Ander, E.L., Cave, M.R., Knights, K.V., Glennon, M.M., Scanlon, R.P., 2014a. Mapping trace element deficiency by cokriging from regional geochemical soil data: a case study on cobalt for grazing sheep in Ireland. *Geoderma* 226–227, 64–78.
- Lark, R.M., Ander, E.L., Knights, K.V., 2016. Information for agriculture from regional geochemical surveys: the example of soil pH in the Tellus and Tellus Border data. In: Young, M.E. (Ed.), *Unearthed: Impacts of the Tellus Surveys of the North of Ireland*. Royal Irish Academy, Dublin.
- Lark, R.M., Hamilton, E.M., Kanninga, B., Maseka, K.K., Mutondo, M., Sakala, G.M., Watts, M.J., 2017. Planning spatial sampling of the soil from an uncertain reconnaissance variogram. *Soil* 3, 235–244.
- Lark, R.M., Knights, K.V., 2015. The implicit loss function for errors in soil information. *Geoderma* 251–252, 24–32.

- Lark, R.M., Mathers, S.J., Thorpe, S., Arkley, S.L.B., Morgan, D.J., Lawrence, D.J.D., 2013. A statistical assessment of the uncertainty in a 3-D geological framework model. *Proc. Geol. Assoc.* 124, 946–958.
- Lark, R.M., Rawlins, B.G., 2008. Can we predict the provenance of a soil sample for forensic purposes by reference to a spatial database?. *Eur. J. Soil Sci.* 59, 1000–1006.
- Lark, R.M., Thorpe, S., Kessler, H., Mathers, S.J., 2014b. Interpretative modelling of a geological cross section from boreholes: sources of uncertainty and their quantification. *Solid Earth* 5, 1189–1203.
- Leenhardt, D., Voltz, M., Bornand, M., Webster, R., 1994. Evaluating soil maps for prediction of soil water properties. *Eur. J. Soil Sci.* 45, 293–301.
- Lelliott, M.R., Cave, M.R., Wealthall, G.P., 2009. A structured approach to the measurement of uncertainty in 3D geological models. *Quart. J. Eng. Geol. Hydrogeol.* 42, 95–105.
- Liu, C.-W., Jang, C.-S., Liao, C.-M., 2004. Evaluation of arsenic contamination potential using indicator kriging in the yun-lin aquifer (Taiwan). *Sci. Total Environ.* 321, 173–188.
- Marchant, B.P., Dailey, A.G., Lark, R.M., 2012. Cost-Effective Sampling Strategies for Soil Management. Home-Grown Cereals Authority Research and Development Project Report No. 485, HGCA, London.
- Mastrandrea, M.D., Field, C.B., Stocker, T.F., Edenhofer, O., Ebi, K.L., Frame, D.J., Held, H., Kriegler, E., Mach, K.J., Matschoss, P.R., Plattner, G.-K., Yohe, G.W., Zwiwers, F.W., 2010. Guidance Note for Lead Authors of the IPCC Fifth Assessment Report on Consistent Treatment of Uncertainties. Intergovernmental Panel on Climate Change (IPCC), <http://www.ipcc.ch/pdf/supporting-material/uncertainty-guidance-note.pdf>.
- Matheron, G., 1976. A simple substitute for conditional expectation: the disjunctive kriging. In: Guarascio, M., David, M., Huijbregts, C. (Eds.), *Advanced Geostatistics in the Mining Industry*. D. Reidel, Dordrecht, pp. 221–236.
- McBratney, A.B., Mendonca-Santos, M.de L., Minasny, B., 2013. On digital soil mapping. *Geoderma* 117, 3–52.
- McBratney, A.B., Webster, R., Burgess, T., 1981. The design of optimal sampling schemes for local estimation and mapping of regionalized variables—I: Theory and method. *Comput. Geosci.* 7, 331–334.
- Milne, G., 1936. Normal erosion as a factor in soil profile development. *Nature* 138, 548–549.
- Milne, A.E., Glendinning, M.J., Lark, R.M., Perryman, S.A.M., Taylor, G., Whitmore, A.P., 2015. Communicating the uncertainty in estimated greenhouse gas emissions from agriculture. *J. Environ. Manag.* 160, 139–153.
- Møller, A.B., Iversen, B.V., Beucher, A., Greve, M.H., 2019. Prediction of soil drainage classes in Denmark by means of decision tree classification. *Geoderma* 352, 314–329.
- Murphy, B.W., 2014. Estimating information levels—using interpretation to convert data into useable information. In: Arrouays, D., McKenzie, N., Hempel, J., Richer de Forges, A., McBratney, A.B. (Eds.), *GlobalSoilMap: Basis of the Global Spatial Soil Information System*. CRC Press, pp. 379–383.
- Osgood-Zimmerman, A., Millea, A.L., Stubbs, R.W., Shields, C., Pickering, B.V., Earl, L., Graetz, N., Kinyoki, D.K., Ray, S.E., Bhatt, S., Browne, A.J., Burstein, R., Cameron, E., Casey, D.C., Deshpande, A., Fullman, N., Gething, P.W., Gibson, H.S., Henry, N.J., Herrero, M., Krause, L.K., Letourneau, I.D., Levine, A.J., Liu, P.Y., Longbottom, J., Mayala, B.K., Mosser, J.F., et al., 2018. Mapping child growth failure in Africa between 2000 and 2015. *Nature* 555, 41–47.
- Patterson, S., McBratney, A.B., Minasny, B., Pringle, M.J., 2018. Variograms of soil properties for agricultural and environmental applications. In: McBratney, A.B., Minasny, B., Stockmann, U. (Eds.), *Pedometrics: A System of Quantitative Soil Information*. Springer, pp. 623–667, (Chapter 21).
- Peck, S.C., Richels, R.G., 1987. The value of information to the acidic deposition debates. *J. Bus. Econom. Statist.* 5, 205–217.
- Peterson, M., 2017. *An Introduction to Decision Theory*. Cambridge University Press, Cambridge.
- Petit, P., Attaallah, B., Manohar, S.G., Husain, M., 2021. The computational cost of active information sampling before decision-making under uncertainty. *Nat. Hum. Behav.* 5, 935–946.
- Pierdzioch, C., Rülke, J.-C., Tillmann, P., 2016. Using forecasts to uncover the loss function of Federal Open Market Committee members. *Macroecon. Dyn.* 20, 791–818.
- Ragg, J.M., Henderson, R., 1980. A reappraisal of soil mapping in an area of southern Scotland, I. The reliability of four soil mapping units, and the morphological variability of their dominant taxa. *J. Soil Sci.* 31, 559–572.
- Raiffa, H., Schlaiffer, R., 1961. *Applied Statistical Decision Theory*. Harvard University, Boston.
- Ramsey, M.H., Taylor, P.D., Lee, J.C., 2002. Optimized contaminated land investigation at minimum overall cost to achieve fitness-for-purpose. *J. Environ. Monit.* 4, 809–814.
- Ruffo, M.L., Bollero, G.A., Hoef, R.G., Bullock, D.G., 2005. Spatial variability of the illinois soil nitrogen test: implications for soil sampling. *Agron. J.* 97, 1485–1492.
- Soil Science Division Staff, 2017. Soil survey manual. In: Ditzler, C., Scheffe, K., Monger, H.C. (Eds.), *USDA Handbook 18*. Government Printing Office, Washington, D.C.
- Spiegelhalter, D., Pearson, M., Short, I., 2011. Visualizing uncertainty about the future. *Science* 333, 1393–1400.
- Stehman, S.V., 1999. Comparing thematic maps based on map value. *Int. J. Remote Sens.* 20, 2347–2366.
- Steinbuch, L., Brus, D.J., Heuvelink, G.B.M., 2018. Mapping the probability of ripened subsoils using Bayesian logistic regression with informative priors. *Geoderma* 316, 56–69.
- Story, M., Congalton, R., 1986. Accuracy assessment: a user's perspective. *Photogramm. Eng. Remote Sens.* 52, 397–399.
- Tversky, A., Kahneman, D., 1981. The framing of decisions and the psychology of choice. *Science* 211, 453–458.
- Valentine, K.W.G., Lord, T.N., Watt, W., Bedway, A.L., 1971. Soil mapping accuracy from black and white, color, and infrared aerial photography. *Can. J. Soil Sci.* 51, 461–469.
- Viscarra Rossel, R.A., Goovaerts, P., McBratney, A.B., 2001. Assessment of the production and economic risks of site-specific liming using geostatistical uncertainty modelling. *Environmetrics* 12, 699–711.
- Webster, R., Beckett, P.H.T., 1968. Quality and usefulness of soil maps. *Nature* 219, 680–682.
- Webster, R., Lark, R.M., 2013. *Field Sampling for Environmental Science and Management*. Routledge, Abingdon.
- Webster, R., Oliver, M.A., 2007. *Geostatistics for Environmental Scientists*. John Wiley & Sons, Chichester.

- Western, S., 1978. *Soil Survey Contracts and Quality Control*. Oxford University Press, Oxford.
- von Winterfeldt, D., 2014. Bridging the gap between science and decision making. *Proc. Natl. Acad. Sci. USA* 110, 14055–14061.
- von Winterfeldt, D., Kavet, R., Peck, S., Mohan, M., Hazen, G., 2012. The value of environmental information without control of subsequent decisions. *Risk Anal.* 12, 2113–2132.
- Wood, G., Oliver, M.A., Webster, R., 1990. Estimating soil salinity by disjunctive kriging. *Soil Use Manage.* 6, 97–104.
- Yates, F., 1952. Principles governing the amount of experimentation in development work. *Nature* 170, 138–140.
- Young, A., 1976. *Tropical Soils and Soil Survey*. Cambridge University Press, Cambridge.
- Zhang, JZ., Huang, HW., Zhang, DM., Phoon, KK., Liu, ZQ., Tang, C., 2021. Quantitative evaluation of geological uncertainty and its influence on tunnel structural performance using improved coupled Markov chain. *Acta Geotech.* 170, 3709–3724.